DTIC FILE COPY

AD-A223 126

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH UNITED STATES AIR FORCE RESEARCH INITIATION PROGRAM CONDUCTED BY UNIVERSAL ENERGY SYSTEMS U.E.S.

DTIC ELECTE JUN 26 1990 S D D

1988

TECHNICAL REPORT

VOLUME 4 OF 4

RODNEY C. DARRAH
PROGRAM DIRECTOR, UES

SUSAN K. ESPY
PROGRAM ADMINISTRATOR, UES

LT. COL. CLAUDE CAVENDER
PROGRAM MANAGER, AFOSR

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 30 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		Technical Report

4. TITLE AND SUBTITLE

1988 Vol. 4 of 4

United States Air Force Research
Initiation Program

5. FUNDING NUMBERS

61102F
3484/D5

6. AUTHOR(S)

Rodney C. Darrah
Lt Col Claude Cavender

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Universal Energy Systems

8. PERFORMING ORGANIZATION REPORT NUMBER

AFOSR-TR- 90 0710

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/XOT
Bld 410
Bolling AFB DC 20332-6448

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

F49620-88-C-0053

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unlimeted

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

See Attached

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	N/A

NSN 7540-01-280-5500

Standard Form 298 (890104 Draft)
Prescribed by ANSI Std. 239-18
298-01

90 06 25 028

AFOSR-TR-90-0710

INTRODUCTION

Research Initiation Program - 1988

AFOSR has provided funding for follow-on research efforts for the participants in the Summer Faculty Research Program. Initially this program was conducted by AFOSR and popularly known as the Mini-Grant Program. Since 1983 the program has been conducted by the Summer Faculty Research Program (SFRP) contractor and is now called the Research Initiation Program (RIP). Funding is provided to establish RIP awards to about half the number of participants in the SFRP.

Participants in the 1988 SFRP competed for funding under the 1988 RIP. Participants submitted cost and technical proposals to the contractor by 1 November 1988, following their participation in the 1988 SFRP.

Evaluation of these proposals was made by the contractor. Evaluation criteria consisted of:

1. Technical Excellence of the proposal
2. Continuation of the SFRP effort
3. Cost sharing by the University

The list of proposals selected for award was forwarded to AFOSR for approval of funding. Those approved by AFOSR were funded for research efforts to be completed by 31 December 1989.

The following summarizes the events for the evaluation of proposals and award of funding under the RIP.

A. Rip proposals were submitted to the contractor by 1 November 1988. The proposals were limited to $20,000 plus cost sharing by the universities. The universities were encouraged to cost share since this is an effort to establish a long term effort between the Air Force and the university.

3. Proposals were evaluated on the criteria listed above and the final award approval was given by AFOSR after consultation with the Air Force Laboratories.

C. Subcontracts were negotiated with the universities. The period of performance of the subcontract was between October 1988 and December 1989.

Copies of the Final Reports are presented in Volumes I through IV of the 1988 Research Initiation Program Report. There were a total of 92 RIP awards made under the 1988 program.

UNITED STATES AIR FORCE

1988 RESEARCH INITIATION PROGRAM

Conducted by

UNIVERSAL ENERGY SYSTEMS, INC.

under

USAF Contract Number F49620-88-C-0053

RESEARCH REPORTS

VOLUME IV OF IV



Submitted to

Air Force Office of Scientific Research

Bolling Air Force Base

Washington, DC

By

Universal Energy Systems, Inc.

April 1990

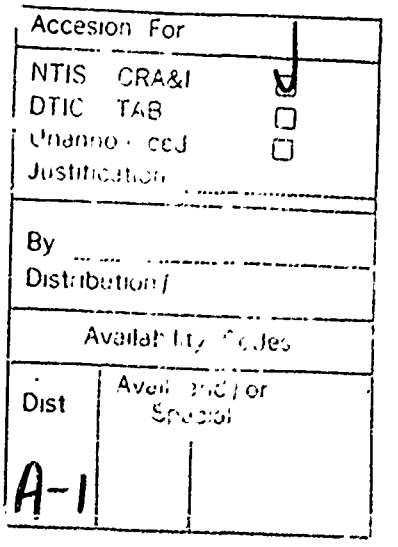


TABLE OF CONTENTS

SECTION	PAGE
INTRODUCTION	i
STATISTICS	ii
PARTICIPANT LABORATORY ASSIGNMENT	vii
RESEARCH REPORTS	xvii

INTRODUCTION

Research Initiation Program - 1988

AFOSR has provided funding for follow-on research efforts for the participants in the Summer Faculty Research Program. Initially this program was conducted by AFOSR and popularly known as the Mini-Grant Program. Since 1983 the program has been conducted by the Summer Faculty Research Program (SFRP) contractor and is now called the Research Initiation Program (RIP). Funding is provided to establish RIP awards to about half the number of participants in the SFRP.

Participants in the 1988 SFRP competed for funding under the 1988 RIP. Participants submitted cost and technical proposals to the contractor by 1 November 1988, following their participation in the 1988 SFRP.

Evaluation of these proposals was made by the contractor. Evaluation criteria consisted of:

1. Technical Excellence of the proposal
2. Continuation of the SFRP effort
3. Cost sharing by the University

The list of proposals selected for award was forwarded to AFOSR for approval of funding. Those approved by AFOSR were funded for research efforts to be completed by 31 December 1989.

The following summarizes the events for the evaluation of proposals and award of funding under the RIP.

A. Rip proposals were submitted to the contractor by 1 November 1988. The proposals were limited to $20,000 plus cost sharing by the universities. The universities were encouraged to cost share since this is an effort to establish a long term effort between the Air Force and the university.

B. Proposals were evaluated on the criteria listed above and the final award approval was given by AFOSR after consultation with the Air Force Laboratories.

C. Subcontracts were negotiated with the universities. The period of performance of the subcontract was between October 1988 and December 1989.

Copies of the Final Reports are presented in Volumes I through IV of the 1988 Research Initiation Program Report. There were a total of 92 RIP awards made under the 1988 program.

STATISTICS

PROGRAM STATISTICS

Total SFRP Participants	153
Total RIP Proposals submitted by SFRP	121
Total RIP Proposals submitted by GSRP	5
Total RIP Proposals submitted	126
Total RIP's funded to SFRP	85
Total RIP's funded to GSRP	3
Total RIP's funded	88
Total RIP's Proposals submitted by HBCU's	8
Total RIP's Proposals funded to HBCU's	4

LABORATORY PARTICIPATION

Laboratory	SFRP Participants	RIP's Submitted	RIP's Funded
AAMRL	10	8 (1 GSRP)	5
AFWAL/APL	8	8	4
ATL	8	9 (1 GSRP)	8 (1 GSRP)
AEDC	5	5 (1 GSRP)	4 (1 GSRP)
AFWAL/AL	8	8	4
ESMC	1	0	0
ESD	2	2	2
ESC	8	7	5
AFWAL/FDL	10	9 (1 GSRP)	6 (1 GSRP)
FJSRL	7	5	4
AFGL	12	7	5
HRL	14	13	9
AFWAL/ML	12	9	6
OEHL	4	3	3
AL	8	7	6
RADC	12	8	8
SAM	16	9	8
WL	6	7 (1 GSRP)	4
WHMC	2	2	1
Total	153	126	88

LIST OF UNIVERSITIES THAT PARTICIPATED

Akron, University of	- 1	Louisiana Tech. University	- 1
Alabama, University of	- 1	Lowell, University of	- 2
Albany College	- 1	Maine, University of	- 1
Arizona State University	- 1	Meharry Medical College	- 1
Arizona, University of	- 1	Miami University	- 1
Arkansas State University	- 1	Miami, University of	- 1
Arkansas, University of	- 2	Michigan State University	- 1
Auburn University	- 1	Michigan Tech. University	- 1
Austin Peay State Univ.	- 1	Michigan, University of	- 2
Ball State University	- 1	Minnesota, University of	- 1
Boston College	- 1	Missouri Westerr State Coll.	- 1
California State Univ.	- 2	Missouri, University of	- 2
California, Univ. of	- 1	Montana, University of	- 1
Calvin College	- 1	Montclair State College	- 1
Carnegie Mellon University	- 1	Morehouse College	- 1
Central State University	- 3	Muhlenberg College	- 1
Central Wesleyan College	- 1	Murray State University	- 1
Cincinnati, University of	- 3	Nebraska, University of	- 1
Clarkson University	- 2	New Hampshire, Univ. of	- 3
Clemson University	- 1	New Mexico, University of	- 1
Colorado State University	- 2	New York State University	- 2
Columbia Basin College	- 1	New York, City College of	- 1
Dayton, University of	- 5	North Carolina State Univ.	- 1
Delta State University	- 1	North Carolina, Univ. of	- 2
East Texas State University	- 1	Northern Illinois Univ.	- 1
Eastern New Mexico Univ.	- 1	Ohio State University	- 2
Fairleigh Dickinson Univ.	- 1	Oklahoma State University	- 1
Fayetteville State Univ.	- 1	Oral Roberts University	- 1
Florida Inst. of Technology	- 1	Oregon Inst. of Technology	- 2
Florida, University of	- 1	Oregon State University	- 1
Francis Marion University	- 1	Pennsylvania State Univ.	- 1
George Mason University	- 1	Polytechnic University	- 1
Georgia Inst. of Technology	- 2	Prairie View A&M Univ.	- 2
Georgia, University of	- 1	Presbyterian College	- 1
Gonzaga University	- 1	Purdue University	- 1
Hampton University	- 1	Redlands, University of	- 1
Illinois Inst. of Technology	- 1	Rennsselaer Polytechnic Inst	- 1
Indiana University	- 1	Rice University	- 1
Iowa State University	- 1	Rochester Inst. of Tech.	- 1
Jackson State University	- 3	Rose-Hulman Inst. of Tech.	- 1
Jacksonville State Univ.	- 1	Saint Paul's College	- 1
Jarvis Christian College	- 1	San Francisco State Univ.	- 1
Kentucky, University of	- 1	Santa Clara University	- 1
LaVerne, University of	- 1	Southeast Oklahoma State U.	- 1
Louisiana State University	- 2	Southern Mississippi, Univ.	- 1

LIST OF UNIVERSITIES THAT PARTICIPATED
Continued

Southern University	- 2	Tuskegee University	- 1
Southwest Missouri State U.	- 1	Virginia Polytechnic Inst.	- 1
St. Norbert College	- 1	Warren Wilson College	- 1
Staten Island, College of	- 1	Wayne State University	- 1
Syracuse University	- 1	Wesleyan College	- 1
Taylor University	- 1	West Florida, University of	- 1
Tennessee Space Inst., Univ.	- 1	West Texas State Univ.	- 1
Tennessee Tech. University	- 2	West Virginia Tech.	- 1
Tennessee, University of	- 1	Western Illinois University	- 1
Texas A&I University	- 1	Western Michigan University	- 1
Texas Lutheran College	- 1	Widener University	- 1
Texas, University of	- 4	Wilberforce University	- 1
Towson State University	- 1	Wisconsin-Madison, Univ. of	- 1
Trinity University	- 1	Wright State University	- 5
		Total	153

PARTICIPANTS LABORATORY ASSIGNMENT

PARTICIPANT LABORATORY ASSIGNMENT

AERO PROPULSION LABORATORY
(Wright-Patterson Air Force Base)

Dr. Suresh K. Aggarwal (1987)
University of Illinois at Chicago
Specialty: Aerospace Engineering

Dr. Mingking K. Chyu
Carnegie Mellon University
Specialty: Heat Transfer

Dr. Derek Dunn-Rankin
University of California
Specialty: Laser Diagnostics (combustion)

Dr. Wayne A. Eckerle
Clarkson University
Specialty: Experimental Fluid Mechanics

Dr. Arthur A. Mason (1986)
University of TEnnessee Space Institute
Specialty: Physics

Dr. Douglas G. Talley
University of Michigan
Specialty: Combustion

Dr. Richard Tankin (1987)
Northwestern University
Specialty: Mechanical Engineering

Dr. Cheng-Hsiao Wu (1987)
University of Missouri
Specialty: Solid State Physics

ARMAMENT LABORATORY
(Eglin Air Force Base)

Dr. Ibrahim A. Ahmad
Northern Illinois University
Specialty: Statistics and Operations

Dr. Charles Bell (1987)
Arkansas State University
Specialty: Mechanical Engineering

Dr. Stephen J. Dow
Univ. of Alabama in Huntsville
Specialty: Discrete Mathematics

Dr. Joseph J Feeley (1987)
University of Idaho
Specialty: Electrical Engineering

Dr. Manuel A. Huerta
University of Miami
Specialty: Plasma Physics

Prof. Anastas Lazaridis
Widener University
Specialty: Ablation, Solar Energy

Dr. Kwang S. Min
East Texas State University
Specialty: Signal Processing

Dr. Joseph J. Molitoris
Muhlenberg College
Specialty: Nuclear Physics

Prof. Wafa E. Yazigi
Columbia Basin College
Specialty: Solid Mechanics

Harry G. Armstrong Aerospace Medical Research Laboratory
(Wright-Patterson Air Force Base)

Dr. Charles D. Covington
University of Arkansas
Specialty: Digital Signal Processing

Dr. Barry P. Goettl
Clemson University
Specialty: Engineering Psychology

Dr. David G. Payne
SUNY Binghamton
Specilaty: Human Memory

Dr. Donald Robertson (1987)
Indiana University of PA
Specialty: Psychology

Dr. Joseph E. Saliba
University of Dayton
Specialty: Engineering Mechanics

Dr. Sanford S. Singer
University of Dayton
Specialty: Enzymology

ARNOLD ENGINEERING DEVELOPMENT CENTER
(Arnold Air Force Base)

Mr. Ben A. Abbott (GSRP)
Vanderbilt University
Specialty: Electrical Engineering

Dr. Eustace L. Dereniak
University of Arizona
Specialty: Infrared Physics

Prof. William M. Grissom
Morehouse College
Specialty: Combustion Diagnostics

Dr. William Sutton (1985)
University of Oklahoma
Specialty: Heat Transfer

Dr. Ahmad D. Vakili
Univ. of Tennessee Space Inst.
Specialty: Unsteady Flows

ASTRONAUTICS LABORATORY
(Edwards Air Force Base)

Dr. Gurbux S. Alag (1987)
Western Michigan University
Specialty: Systems Engineering

Dr. Clarence Calder
Oregon State University
Specialty: Stress Wave Propagation

Mr. David B. Chenault (GSRP)
University of Alabama
Specialty: Physics

Dr. David W. Jensen
Pennsylvania State University
Specialty: Advanced Composite Materials

Dr. John Kenney (1987)
Eastern New Mexico University
Specialty: Phsyical Chemistry

Dr. Mark A. Norris
Virginia Polytechnic Inst. & State Univ.
Specialty: Structural Dynamics & Controls

ASTRONAUTICS LABORATORY
(Edwards Air Force Base)
(continued)

Dr. Phillip A. Christiansen
Clarkson University
Specialty: Physical Chemistry

Dr. Susan T. Collins
California State University
Specialty: Matrix Isolation Spectroscopy

Dr. Rameshwar P. Sharma
Western Michigan University
Specialty: Fluid Mechanics

Dr. Siavash H. Sohrab (1986)
Northwestern University
Specialty: Engineering Physics

AVIONICS LABORATORY
(Wright-Patterson Air Force Base)

Prof. William K. Curry (1987)
Rose-Hulman Inst. of Technology
Specialty: Computer Science

Dr. Gerald W. Grams
Georgia Tech.
Specialty: Atmospheric Physics

Dr. David Hemmendinger
Wright State University
Specialty: Logic Programming

Dr. Feriasamy K. Rajan
Tennessee Tech. University
Specialty: Digital Signal Processing

Dr. Mateen M. Rizki
Wright State University
Specialty: Modeling and Simulation

ENGINEERING AND SERVICES CENTER
(Tyndall Air Force Base)

Dr. Wayne A. Charlie
Colorado State University
Specialty: Geotechnical Engineering

Dr. David H. DeHeer
Calvin College
Specialty: Molecular Biology

Dr. Deanna S. Durnford
Colorado State University
Specialty: Groundwater

Dr. Neil J. Hutzler
Michigan Tech. University
Specialty: Environmental Engineering

Dr. Peter Jeffers (1987)
S.U.N.Y.
Specialty: Chemistry

Dr. Richard S. Myers
Delta State University
Specialty: Experimental Physical Chem.

Dr. William Schulz (1987)
Eastern Kentucky University
Specialty: Chemistry

Dr. Dennis Truax (1987)
Mississippi State University
Specialty: Civil Engineering

ELECTRONIC SYSTEMS DIVISION
(Hanscom Air Force Base)

Mr. George N. Bratton
Austin State Peay State Univ.
Specialty: Statistics

Dr. John F. Dalphin
Towson State University
Specialty: Computer Science

Dr. Stephan Kolitz (1986)
University of Massachusetts
Specialty: Operations Reserach

FLIGHT DYNAMICS LABORATORY
(Wright-Patterson Air Force Base)

Dr. Peter J. Disimile (1986)
University of Cincinnati
Specialty: Fluid Mechanics

Mr. Thomas Enneking (GSRP), (1987)
University of Notre Dame
Specialty: Civil Engineering

Dr. Awatef Hamed
University of Cincinnati
Specialty: Engineering

Dr. Yulian B. Kin
Purdue University Calumet
Specialty: Stress Analysis

Dr. Oliver McGee (1987)
Ohio State University
Specialty: Engineering Mechanics

Dr. James A. Sherwood
University of New Hampshire
Specialty: Solid Mechanics

Dr. Gary Slater (1987)
University of Cincinnati
Specialty: Aerospace Engineering

Dr. Kenneth M. Sobel
The City College of New York
Specialty: Eigenstructure

Dr. Forrest Thomas (1987)
University of Montana
Specialty: Chemistry

Mr. David F. Thompson (GSRP)
Purdue University
Specialty: Computer Information

Dr. William E. Wolfe
Ohio State University
Specialty: Geotechnical Engineering

FRANK J. SEILER RESERACH LABORATORY
(United States Air Froce Academy)

Dr. Richard Bertrand (1985)
University of Colorado
Specialty: NMR Spectroscopy, Atomic Spectroscopy

Dr. Tammy J. Melton
St. Norbert College
Specialty: Inorganic Synthesis

FRANK J. SEILER RESERACH LABORATORY
(United States Air Froce Academy)
(continued)

Dr. Dan R. Bruss
Albany College of Pharmacy
Specialty: Physical Organic Chemistry

Dr. Charles M. Bump (1987)
Hampton University
Specialty: Organic Chemistry

Dr. Michael L. McKee
Auburn University
Specialty: Molecular Orbital Theory

Dr. Patricia L. Plummer
Columbia Univ. of Missouri
Specialty: Quantum Chemistry

Dr. Howard Thompson (1987)
Purdue University
Specialty: Mechanical Engineering

Dr. Melvin Zandler (1987)
Wichita State University
Specialty: Physical Chemistry

GEOPHYSICS LABORATORY
(Hanscom Air Force Base)

Dr. Lucia M. Babcock
Louisiana State University
Specialty: Gas Phase Ion-Molecule Chem.

Dr. Pradip M. Bakshi
Boston College
Specialty: Quantum Theory

Dr. Donald F. Collins
Warren Wilson College
Specialty: Optics, Image Processing

Dr. Lee Flippin (1987)
San Francisco State University
Specialty: Organic Chemisty

Dr. Janet U. Kozyra
University of Michigan
Specialty: Space Physics

Dr. Steven Leon (1987)
Southeastern Massachusettes
Specialty: Mathematics

Dr. John P. McHugh
University of New Hampshire
Specialty: Fluid Mechanics

Dr. Timothy Su (1987)
Southeastern Massachusetts Univ.
Specialty: Physical Chemistry

HUMAN RESOURCES LABORATORY
(Brooks, Williams and Wright-Patterson Air Force Base)

Dr. Ronna Dillion (1987)
Southern Illinois University
Specialty: Educational Psychology

Dr. J. Kevin Ford
Michigan State University
Specialty: Industrial/Organ. Psychology

Dr. Jorge L. Mendoza (1986)
Texas A&M University
Specialty: Psychology

Dr. Philip D. Olivier (1986)
University of Texas
Specialty: Electrical Engineering

HUMAN RESOURCES LABORATORY
(Brooks, Williams and Wright-Patterson Air Force Base)
(continued)

Dr. Hugh. P. Garraway, III
Univ. of Southern Mississippi
Specialty: Computer Based Learning

Dr. Douglas E. Jackson
Eastern New Mexico University
Specialty: Math/Statistical Information

Dr. Charles E. Lance
University of Georgia
Specialty: Industrial/Organizational Psy.

Dr. Thomas L. Landers
University of Arkansas
Specialty: Reliability & Maintainability

Dr. Mufit H. Ozden
Miami University
Specialty: Operations Research

Dr. Dharam S. Rana
Jackson State University
Specialty: Quantitative Techniques

Dr. Jonathan M. Spector
Jacksonville State University
Specialty: Logic

Dr. Charles Wells (1987)
University of Dayton
Specialty: Management Science

Dr. Robert K. Young
University of Texas
Specialty: Experimental Psychology

LOGISTICS COMMAND
(Wright-Patterson Air Force Base)

Dr. Ming-Shing Hung (1986)
Kent State University
Specialty: Business Administration & Management Science

MATERIALS LABORATORY
(Wright-Patterson Air Force Base)

Dr. Bruce Craver (1987)
University of Dayton
Specialty: Physics

Dr. Parvis Dadras
Wright State University
Specialty: Mechanics of Materials

Dr. David A. Grossie
Wright State University
Specialty: X-ray Crystallography

Dr. Gordon Johnson (1987)
Walla Walla College
Specialty: Electrical Engineering

Dr. L. James Lee
The Ohio State University
Specialty: Polymer & Composite Processing

MATERIALS LABORATORY
(Wright-Patterson Air Force Base)
(continued)

Dr. Barry K. Fussell
University of New Hampshire
Specialty: Systems Modeling & Controls

Dr. John W. Gilmer (1987)
Penn State University
Specialty: Physical Chemistry

Dr. Michael Sydor
University of Minnesota
Specialty: Optics, Material Science

Dr. Richard S. Valpey
Wilberforce University
Specialty: Organic Synthesis

OCCUPATIONAL AND ENVIRONMENT HEALTH LABORATORY
(Brooks Air Force Base)

Dr. Steven C. Chiesa
Santa Clara University
Specialty: Biological Waste Treatment

Dr. Larry R. Sherman
University of Akron
Specialty: Organotin Chemistry

Dr. Gary R. Stevens
Oklahoma State University
Specialty: Stochastic Processes

Dr. Shirley A. Williams (1986)
Jackson State University
Specialty: Physiology

ROME AIR DEVELOPMENT CENTER
(Griffiss Air Force Base)

Dr. Keith A. Christianson
University of Maine
Specialty: Electronic Materials

Dr. Hugh K. Donaghy
Rochester Inst. of Technology
Specialty: Natural Language Processing

Dr. Oleg G. Jakubowicz
State University of New York
Specialty: Neural Nets

Dr. Louis Johnson (1987)
Oklahoma State University
Specialty: Electrical Engineering

Dr. Samuel P. Kozaitis
Florida Institute of Tech.
Specialty: Optics, Computer Architecture

Dr. David Sumberg (1987)
Rochester Institute of Tech.
Specialty: Physics

Dr. Donald R. Ucci
Illinois Inst. of Technology
Specialty: Adaptive Arrays

Dr. Peter J. Walsh
Fairleigh Dickinson University
Specialty: Superconductivity

Dr. Kenneth L. Walter
Prairie View A&M University
Specialty: Chemical Engineering Process

Dr. Gwo-Ching Wang
Rensselaer Polytechnic Inst.
Specialty: Surface Sciences

SCHOOL OF AEROSPACE MEDICINE
(Brooks Air Force Base)

Dr. Ronald Bulbulian
University of Kentucky
Specialty: Exercise Physiology

Dr. John A. Burke, Jr.
Trinity University
Specialty: Inorganic Compounds

Dr. Hoffman H. Chen (1986)
Grambling State University
Specialty: Mechanical Engineering

Dr. Frank O. Hadlock (1986)
Florida Atlantic University
Specialty: Mathematics

Dr. Eric R. Johnson
Ball State University
Specialty: Protein Biochemistry

Dr. Harold G. Longbotham
Univ. of Texas - San Antonio
Specialty: Nonlinear Digital Filtering

Dr. Mohammed Maleque (1987)
Meharry Medical College
Specialty: Pharmacology

Dr. Parsottam J. Patel (1986)
Meharry Medical College
Specialty: Microbiology

Dr. William Z. Plachy
San Francisco State University
Specialty: Physical Chemistry

Dr. Ralph Peters (1987)
Wichita State University
Specialty: Zoology

Dr. Thomas R. Rogge
Iowa State University
Specialty: Finite Element Analysis

Prof. Sonia H. Sawtelle-Hart
Univ. of Texas - San Antonio
Specialty: Exercise Physiology

Dr. Wesley Tanaka (1987)
University of Wisconsin
Specialty: Biochemistry

Dr. John R. Wright
Southeast Oklahoma State Univ.
Specialty: Biochemistry

WILFORD HALL MEDICAL CENTER
(Lackland Air Force Base)

Dr. David R. Cecil
Texas A&I University
Specialty: Algebra (Finite Fields)

Dr. Donald Welch (1986)
Texas A&M University
Specialty: Microbiology

WEAPONS LABORATORY
(Kirtland Air Force Base)

Dr. Albert W. Biggs (1986)
Unviersity of Alabama
Specialty: Electrical Engineering

Dr. William M. Jordan
Lousiana Tech. University
Specialty: Composite Materials

WEAPONS LABORATORY
(Kirtland Air Force Base)
(continued)

Dr. Lane Clark
University of New Mexico
Specialty: Graph Theory

Dr. David A. Dolson
Murray State University
Specialty: Laser Spectroscopy

Dr. Arkady Kheyfets
North Carolina State Univ.
Specialty: Mathematical Physics

Dr. Barry McConnell (1987)
Florida A&M University
Specialty: Computer Science

Dr. William Wheless (1987)
New Mexico State University
Specialty: Electrical Engineering

RESEARCH REPORTS

MINI-GRANT RESEARCH REPORTS

1988 RESEARCH INITIATION PROGRAM

Technical Report Number	Title and Mini-Grant No.	Professor
Volume I		
Armament Laboratory		
1	Statistical Analysis of Residual Target Performance and for Measures of Target Partial Availability Pending Approval 210-9MG-010	Dr. Ibrahim A. Ahmad
2	Synergistic Effects of Bomb Cratering, Phase II 760-7MG-025	Dr. Charles Bell (1987)
3	Automated Motion Parameter Determination from an Image Sequence 210-9MG-025	Dr. Stephen J. Dow
4	Modeling and Simulation on Microcomputers, 1989 760-7MG-070	Dr. Joseph J. Feeley (1987)
5	Two Dimensional MHD Simulation of Accelerating Arc Plasmas 210-9MG-090	Dr. Manuel A. Huerta
6	Modeling Reactive Fragments 210-9MG-011	Prof. Anastas Lazaridis
7	Target-Aerosol Discrimination for Active Optical Proximity Sensors 210-9MG-016	Dr. Kwang S. Min
8	The Dynamics of Impact 210-9MG-008	Dr. Joseph J. Molitoris
9	Report Not Acceptable at this Time 210-9MG-015	Prof. Wafa E. Yazigi

Arnold Engineering Development Center

10	Multigraph Kernel for Transputer Based Systems 21-9MG-087	Mr. Ben A. Abbott (GSRP)
11	MTF Studies of IR Focal Plane Arrays at Low Flux Levels 210-9MG-020	Dr. Eustace L. Dereniak
12	Droplet Size Distributions and Combustion Modeling in a Pintle Injector Spray 210-9MG-069	Prof. William M. Grissom
13	Multiple Scattering in Solid Fuel Rocket Plumes 760-0MG-091	Dr. William Sutton (1985)
14	Influence of Forced Disturbances on the Vortex Core and the Vortex Burst 210-9MG-056	Dr. Ahmad D. Vakili

Astronautics Laboratory

15	Large Space Structure Parameter Estimation 760-7MG-042	Dr. Gurbux S. Alag (1987)
16	Integrated Strain Measurement in Composite Members Using Embedded Constantan Wire Pending Approval 210-9MG-076	Dr. Clarence Calder
17	Calibration of the Infrared Spectro-polarimeter 210-9MG-026	Mr. David B. Chenault (GSRP)
18	Computer Code to Include Core Polarization in Effective Potential Basis Set Expansion Studies 210-9MG-092	Dr. Phillip A. Christiansen
19	Fluorescence Spectra of Matrix-isolated Lithium 210-9MG-115	Dr. Susan T. Collins

20	Calibration of Composite-Embedded Fiber-Optic Strain Sensors Pending Approval 210-9MG-052	Dr. David W. Jensen
21	Energy-And Time-Resolved Photophysics and Photochemistry of High Energy Cryogenic Metal-Containing Rocket Fuels 760-7MG-019	Dr. John Kenney (1987)
22	Experimental Verification and Development of Structural Identification Techniques on a Grid 210-9MG-045	Dr. Mark A. Norris
23	Report Not Available at this Time 210-9MG-103	Dr. Rameshwar P. Sharma
24	Experimental Investigation of the Stability of Jets Near the Critical Point 760-6MG-110	Dr. Siavash H. Sohrab (1986)
Electronics Systems Division		
25	HF Network Evaluation 210-9MG-012	Mr. George N. Bratton
26	Report Not Available at this Time 210-9MG-023	Dr. John F. Dalphin
27	Reliability in Satellite Communication Networks Pending Approval 760-6MG-094	Dr. Stephan Kolitz (1986)
Engineering and Services Center		
28	High Intensity Compressive Stress Wave Propagation Through Unsaturated Sands 210-9MG-075	Dr. Wayne A. Charlie
29	Decontamination and Elisa Analysis of Blood Group Substances from Human Tissue Pending Approval 210-9MG-112	Dr. David H. DeHeer

30	Estimation of Jet Fuel Contamination in Soils 210-9MG-074	Dr. Deanna S. Durnford
31	Extraction of Volatile Organic Chemicals from Unsaturated Soil: Experimental Results and Model Predictions 210-9MG-059	Dr. Neil J. Hutzler
32	Homogeneous Hydrolysis Rate Constants for Selected Chlorinated Methanes, Ethanes, Ethenes, and Propanes 760-7MG-038	Dr. Peter Jeffers (1987)
33	Sorption Kinetics of Volatile Organic Compounds on Aquifer Materials 210-9MG-047	Dr. Richard S. Myers
34	Report will be Submitted Under Mini-Grant 210-10MG-095	Dr. William Schulz (1987)
35	Report Not Available at this Time 760-7MG-105	Dr. Dennis Truax (1987)

Volume II
Frank J. Seiler Research Laboratory

36	NMR Studies of Alkylammonium-Chloroaluminate Room-Temperature Electrolytes 760-0MG-095	Dr. Richard Bertrand (1985)
37	Mechanistic Studies on the Thermal Decomposition of NTO by High Performance Liquid Chromatography 210-9MG-111	Dr. Dan R. Bruss
38	Aromatic Nitrations in Chloroaluminate Melts 760-7MG-076	Dr. Charles M. Bump (1987)
39	Calculated C-NO_2 Bond Dissociation Energies (PartI) and A MCSCF Study of the Rearrangement of Nitromethane to Methyl Nitrite (Part II) 210-9MG-054	Dr. Michael L. McKee
40	Sodium as an Electrode for Chloroaluminate Melts 210-9MG-098	Dr. Tammy J. Melton

41	Report Not Available at this Time 210-9MG-097	Dr. Patricia L. Plummer
42	Transient Shock Waves in a Mach 3 Flow 760-7MG-071	Dr. Howard Thompson (1987)
43	Ab-initio and Semi-Empirical Molecular Orbital Studies of Energetic Materials (Nitrogen Heterocyclics) and Polymers 760-7MG-092	Dr. Melvin Zandler (1987)
Geophysics Laboratory		
44	Radiative Association in Ion-Molecule Reactions: Reactions of Some Carbon Cations 210-9MG-086	Dr. Lucia M. Babcock
45	Impulse Approximation Fromalism for Atom Molecule Collisions 210-9MG-109	Dr. Pradip M. Bakshi
46	Stellar Photometry, Vehicle Glow, and Advanced Image Analysis 210-9MG-100	Dr. Donald F. Collins
47	Synthesis of Organometallic Reagents for SIFT Studies of Electron Attachment Reactions 760-7MG-056	Dr. Lee Flippin (1987)
48	Theoretical and Observational Studies of Geomagnetic Storm-Related Ion and Electron Heating in the Subauroral Region 210-9MG-084	Dr. Janet U. Kozyra
49	Algorithms for Generalized Exponential Inversion 760-7MG-036	Dr. Steven Leon (1987)
50	Report Not Available at this Time 210-9MG-125	Dr. John P. McHugh
51	Trajectory Calculations of High Temperature and Kinetic Energy Dependent Ion-Polar Molecule Collision Rate Constants 760-7MG-040	Dr. Timothy Su (1987)

Rome Air Development Center

52	Aging Studies of GaAs Schottky Barriers 210-9MG-073	Dr. Keith A. Christianson
53	Report Not Available at this Time 210-9MG-094	Dr. Hugh K. Donaghy
54	Neural Network for Aiding Intelligent Analysis Pending Approval 210-9MG-124	Dr. Oleg G. Jakubowicz
55	Supply Line Testing in CMOS Digital Circuits 760-7MG-050	Dr. Louis Johnson (1987)
56	Characterization of Detectors for Optical Pattern Recognition 210-9MG-018	Dr. Samuel P. Kozaitis
57	Fiber Optic Distribution for Phased Array Antennas Pending Approval 760-7MG-113	Dr. David Sumberg (1987)
58	Continuation Study of the Effect of Nonlinearities of High Speed Analog-to-Digital Converters on Digital Beamforming Arrays 210-9MG-040	Dr. Donald R. Ucci
59	Analysis of Microwave Surface Impedance of High Temperature Superconductors 210-9MG-072	Dr. Peter J. Walsh
60	Report Not Available at this Time 210-9MG-113	Dr. Kenneth L. Walter
61	X-Ray Pole-Figure Analysis of $YBa_2Cu_3O_{7-x}$ Thin Film on $SrTiO_3(100)$ Prepared by RF Diode Sputtering 210-9MG-077	Dr. Gwo-Ching Wang

Weapons Laboratory

62	Slow Wave Transmission Line Transformers 760-6MG-072	Dr. Albert W. Biggs (1986)

63	Report Not Available at this Time 210-9MG-119	Dr. Lane Clark
64	Vibrational Energy Transfer in Sulfur Monoxide 210-9MG-101	Dr. David A. Dolson
65	Development of an Experimental Program to Evaluate Laser Composite Material Damage Models 210-9MG-034	Dr. William M. Jordan
66	Report Not Available at this Time 210-9MG-114	Dr. Arkady Kheyfets
67	Report Not Available at this Time 760-7MG-047	Dr. Barry McConnell (1987)
68	Slow to Fast Wave Transition Analysis 760-7MG-068	Dr. William Wheless (1987)

Volume III
Air Force Wright Aeronautical Laboratories
Aero Propulsion Laboratory

69	Vaporization Behavior of Pure and Multicomponent Fuel Droplets in a Hot Air Stream 760-7MG-061	Dr. Suresh K. Aggarwal (1987)
70	Effects of Injection-To-Mainstream Density Ratios on Film Cooling Heat Transfer Pending Approval 210-9MG-096	Dr. Mingking K. Chyu
71	Accurate Temperatures Using Cars in Droplet Laden Flows Pending Approval 210-9MG-055	Dr. Derek Dunn-Rankin
72	Report Not Available at this Time 210-9MG-019	Dr. Wayne A. Eckerle
73	Report Not Available at this Time 760-6MG-099	Dr. Arthur A. Mason (1986)
74	Report Not Available at this Time 210-9MG-022	Dr. Douglas G. Talley

75	Vortical Structures in 2-D Slot Burner-Cold Flow 760-7MG-051	Dr. Richard Tankin (1987)
76	Calculations of Interface-State Occupation Function and GaAs/Ge Heterostructure Solar Cell Efficiency 760-7MG-093	Dr. Cheng-Hsiao Wu (1987)

Avionics Laboratory

77	Computer Simulation of Adaptive Resource Management in Real-Time 760-7MG-081	Prof. William K. Curry (1987)
78	Study of Sky Backgrounds and Subvisual Cirrus Pending Approval 210-9MG-120	Dr. Gerald W. Grams
79	Proving Equivalence of High-and Low-Level Architectural Descriptions in VHDL 210-9MG-108	Dr. David Hemmendinger
80	Report Not Available at this Time 210-9MG-051	Dr. Periasamy K. Rajan
81	Applications of Evolutionary Learning Strategies to Pattern Recognition Tasks 210-9MG-058	Dr. Mateen M. Rizki

Flight Dynamics Laboratory

82	The Effect of a Roughened Surface on Turbulent Boundary Layer Separation at Mach 6.0 760-6MG-075	Dr. Peter J. Disimile (1986)
83	A Stochastic Model of Fatigue Crack Growth Due to Random Loading for Application to Aircraft Wheels 760-7MG-124	Mr. Thomas Enneking (GSRP) (1987)
84	An Investigation of the Flow Field in Shock Wave/Boundary Layer/Bleed Interactions Pending Approval 210-9MG-061	Dr. Awatef Hamed

85	Fatigue Characteristics of F-16 Composite Transparency Material Determined by Long-Term and Accelerated Methods 210-9MG-038	Dr. Yulian B. Kin
86	Convergence of Upper-Bound Optimum Design of Large-Scale Structures with Specified Frequency Bands 760-7MG-115	Dr. Oliver McGee (1987)
87	Report Not Available at this Time 210-9MG-088	Dr. James A. Sherwood
88	Robustness with Positive Real Controllers for Large Space Structures 760-7MG-088	Dr. Gary Slater (1987)
89	Robust Eigenstructure Assignment for Flight Control Design 210-9MG-035	Dr. Kenneth M. Sobel
90	Comparative Burning Rates and Duplex Loads of Solid Propellants 760-7MG-080	Dr. Forrest Thomas (1987)
91	Optimal and Sub-Optimal Loop Shaping in Quantitative Feedback Theory Pending Approval 210-9MG-106	Mr. David F. Thompson (GSRP)
92	Low Velocity Impact of Composite Materials 760-7MG-102 and 210-9MG-082	Dr. William E. Wolfe
Logistics Command		
93	Aircraft Availabilty Model: Feasibility Study for POM Forecasting 760-6MG-105	Dr. Ming S. Hung (1986)
Materials Laboratory		
94	Tunable Absorption in Doping Superlattices 760-7MG-097	Dr. Bruce Craver (1987)
95	Joining of Carbon-Carbon Composites 210-9MG-004	Dr. Parviz Dadras
96	Report Not Available at this Time 210-9MG-064	Dr. Barry K. Fussell

97	Characterization of the Phase Separation Behavior of Poly(p-phenylene benzobisthiazole)/Amorphous Nylon Molecular Composites by Small Angle Light Scattering 760-7MG-013	Dr. John W. Gilmer
98	Structural Analysis of Model Compounds with Potential Second and Third Order Nonlinear Optical Properties 210-9MG-080	Dr. David A. Grossie
99	Liquid Crystal Biomolecules for use as Optical Filters 760-7MG-075	Dr. Gordon Johnson (1987)
100	Knowledge Development for the Rule Based Process Automation of Resin Transfer Molding 210-9MG-063	Dr. L. Jmaes Lee
101	Photoreflectance Measurements of the Quality of Undoped GaAs 210-9MG-031	Dr. Michael Sydor
102	Synthesis of 2, 6-Diformyl Pyridobisimidazoles 210-9MG-029	Dr. Richard S. Valpey

Volume IV
Human Systems Division Laboratories
Harry G. Armstrong Aerospace Medical Research Laboratory

103	Auditory Modeling 210-9MG-060	Dr. Charles D. Covington
104	Assessing the Cognitive Demands of Tracking Strategies 210-9MG-078	Dr. Barry P. Goettl
105	Report Not Available at this Time 210-9MG-121	Dr. David G. Payne
106	Effect of System Reliability on Probabilistic Inference Pending Approval 760-7MG-094	Dr. Donald Robertson (1987)

107	Optimization of the Nonlinear Discrete Parameter Model of the Seated Human Spine 210-9MG-071	Dr. Joseph E. Saliba
108	In Vitro Modeling of Perfluoro-N-Decanoate Effects on Enzymes of Fatty Acid Metabolism 210-9MG-002	Dr. Sanford S. Singer

Human Resources Laboratory

109	Report Not Acceptable at this Time 760-7MG-100	Dr. Ronna Dillion (1987)
110	An Investigation of Training Content Validity and Training Efficiency in the Air Force Airmen Basic-In-Residence Training Course 210-9MG-066	Dr. J. Kevin Ford
111	An Intelligent Tool to Facilitate the Development of Qualitative Process Models in Novice Programmers 210-9MG-007	Dr. Hugh P. Garraway, III
112	On the Effect of Range Restriction on Correlation Coefficient Estimation 210-9MG-027	Dr. Douglas E. Jackson
113	Validation of an Enlisted Air Force Specialty Task Taxonomy and Cross-AFS Ease-of-Movement Predictions 210-9MG-017	Dr. Charles E. Lance
114	Proportional Intensity Reliability Analysis for Repairable Items 210-9MG-104	Dr. Thomas L. Landers
115	A Monte Carlo Comparison of Validity Generalization Procedures 760-6MG-136	Dr. Jorge L. Mendoza (1986)
116	A Network Tutor Based on the Heuristic of Polya 760-6MG-032	Dr. Philip D. Olivier (1986)
117	Graphical Programming of Simulation Models in an Object-Oriented Environment 210-9MG-028	Dr. Mufit H. Ozden

118	Report Not Available at this Time 210-9MG-043	Dr. Dharam S. Rana
119	Refinement Considerations for an Advanced Instructional Design Advisor 210-9MG-021	Dr. Jonathan M. Spector
120	Engineering Design with Decision Support: An Application of Goal Decomposition 760-7MG-046	Dr. Charles Wells (1987)
121	Report Not Available at this Time 210-9MG-099	Dr. Robert K. Young

Occupational and Environment Health Laboratory

122	Solvent Extraction of Boron from Industrial Wastewaters 210-9MG-102	Dr. Steven C. Chiesa
123	Comparison of Asbestos Analysis by SEM-EDXA and TEM-SAED 210-9MG-122	Dr. Larry R. Sherman
124	An Examination of Kriging Techniques for Ground Water Monitoring 210-9MG-070	Dr. Gary R. Stevens
125	Cortisol Prevention of Chronic Beryllium Disease in Postpartum Rats: A Pilot Study 760-6MG-078	Dr. Shirley A. Williams (1986)

School of Aerospace Medicine

126	Blood Flow Distribution in the Non-Working Forearm During Exercise 210-9MG-057	Dr. Ronald Bulbulian
127	Photophysics and Photochemistry of Transition Metal Complexes 210-9MG-091	Dr. John A. Burke, Jr.
128	Serum Squalene and Cholesterol Ratio as Risk Predictor for Coronary Artery Disease 760-6MG-118	Dr. Hoffman H. Chen (1986)
129	A Feasibility Study for a Computerized ECG Database 760-6MG-073	Dr. Frank O. Hadlock (1986)

130	Development of a New Ultrasensitive Cholesterol Assay System for the Determination of Free Cholesterol in Biological Fluids 210-9MG-105	Dr. Eric R. Johnson
131	Application of Nonlinear Filters to VEP Data 210-9MG-033	Dr. Harold G. Longbotham
132	Effects of Low Dose Soman on CNS Neurotransmitters 760-7MG-078	Dr. Dr. Mohammed Maleque (1987)
133	Cleansing of Bone-Marrow by Lymphokine Activated Killer Cells (LAK-Cells) 760-6MG-131	Dr. Parsottam J. Patel (1986)
134	Transcutaneous Oxygen Delivery 210-9MG-042	Dr. William Z. Plachy
135	Report Not Acceptable at this Time 760-7MG-091	Dr. Ralph Peters (1987)
136	A Computer Model of the Human Systemic Arterial Tree 210-9MG-003	Dr. Thomas R. Rogge
137	The Effect of Age, Family Status, and Physical Activity on Select Dietary Components of TAC Pilots 210-9MG-095	Prof. Sonia H. Sawtell-Hart
138	Comprehensive Lipoprotein Analysis by High-Performance Molecular Exclusion Chromatography 760-7MG-043	Dr. Wesley Tanaka (1987)
139	Nmr and Temperature-Dependence Studies of the Metal-Ion Catalyzed Chemiluminescence of Luminol 210-9MG-037	Dr. John R. Wright
Wilford Hall Medical Center		
140	Enhancements to PC-Mainframe Interface for Data Entry 210-9MG-048	Dr. David R. Cecil

141	Effect of Hyperoxia on the Permeability of the Blood-Brain Barrier in Several Laboratory Species and on Organotypic Explant Tissue Cultures of Hamster Brain 760-6MG-091	Dr. Donald Welch (1986)

1989 FOLLOW-ON GRANT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Universal Energy Systems, Inc.

FINAL REPORT

Auditory Modeling

Prepared by:	C. David Covington, Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Department of Electrical Engineering University of Arkansas, Fayetteville
Research Location:	University of Arkansas Department of Electrical Engineering BEC 3217 Fayetteville Arkansas 72701
USAF Researcher:	Timothy R. Anderson
Date:	1 Jun 1989
Contract No:	F49620-88-C-0053

Auditory Modeling

by

C. David Covington

ABSTRACT

Payton [J. Acoust. Soc. Am. 83, 145-162, (1988)] and Lyon and Mead [IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-36, pp. 1119-1134, July 1988] have constructed auditory models based on two entirely different assumptions regarding the behavior of biological auditory systems. Payton incorporated model components from a variety of sources in order to create a complete auditory model predicting neural firing rates based on a pressure wave input at the tympanic membrane. In particular, Payton called upon basilar membrane modeling by Allen and Sondhi. We have reimplemented the middle ear, basilar membrane, and response sharpening filter models from Payton in previous work. Lyon and Mead proposed a cascade of 480 analog second order filters on a single integrated circuit as means of modeling the behavior of waves on the basilar membrane. In this report we transform the specifications given by Lyon and Mead for an analog electronic cochlea into an approximate equivalent discrete model. We then compare the resulting magnitude and phase responses at 20 membrane locations. We compare the discretized Lyon-Mead model first with published behavior of the analog Lyon-Mead model and then with responses we observed for the Payton model.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Universal Energy Systems must be mentioned for their concern and help to me in all administrative aspects of this program. This work began with a Summer Faculty Research Program appointment during 1988 at the Armstrong Aerospace Medical Research Laboratory of Wright-Patterson Air Force Base. I was also fortunate to have my graduate student Michael Ellis accepted in the Graduate Student Research Program to work with me on the project. Special thanks go to Dr. Karen Payton for providing the FORTRAN source code from which we derived the lisp version used to produce the results stated in this report. Timothy Anderson of the lab has also provided continuing support and direction to the research without which this report would not be possible.

I. INTRODUCTION

In order to better understand how humans process acoustical signals in the process we call hearing, researchers over the years have studied the results of medical experiments on human and animal auditory systems. Using the results of such investigations, they have proposed increasingly effective models which predict acoustical and neural activity at various points along the hearing apparatus. These models offer hope of revealing the underlying processes involved in the various auditory subsystems.

Two auditory components have received a great deal of attention in recent years, specifically the basilar membrane and hair cell systems. This research focuses primarily on the behavior of the basilar membrane. Some models however incorporate adaptive effects in the inner and/or outer hair cells in order to account for the shape of tuning curves and observed time constants for transient inputs.

The inclusion of adaptive amplification phenomena gives cochlear modeling a new appeal as means to improved speech processing and represents acoustical transformations in the inner ear that cannot be adequately captured in fixed coefficient filter bank models. In general, speech recognition developments should benefit significantly from insights gained from the analysis of both time proven and novel cochlear models.

This research compares the published work of Payton [16] and Lyon and Mead [11]. The two models proposed rest on substantially different assumptions however regarding the propagation of waves in the cochlea. The first model by Payton assumes that all adaptive behavior resides entirely in the hair cell transduction mechanisms or at least that adaptation can be adequately represented in the hair cell model only. In order to accomodate the apparent narrowly tuned response at a particular place along the basilar membrane, Payton incorporated a second order response sharpening filter at each membrane tap of interest.

The second model by Lyon and Mead departs from this thinking, using broadly tuned filter sections and explains the peak sharpening by incorporating an adaptive level-dependent negative damping phenomena. This negative damping takes place at the outer hair cells where a hair fiber deflection is accompanied by a reinforcing pull in the direction of motion. This reinforcement adds energy to the wave traveling along the membrane and increases the effective Q of a local section of the membrane. Lyon and Mead consider the basilar membrane as a nonuniform transmission line which can be modeled as a cascade of filters each representing a small section of the membrane.

This report briefly reviews the mathematics underlying each of the models investigated and discusses the details of computer implementations of the two models. Extensive discussions and figures are given which describe the output of the two auditory models when the input is an impulse.

II. OBJECTIVES OF THE RESEARCH EFFORT

In our Summer Faculty Research appointment [7], we ported two auditory models to the Symbolics lisp machines in the Biodynamics/Bioacoustics laboratory. Specifically we installed the Spire resident SAM system by Seneff of MIT and converted the FORTRAN based system written by Payton to run on the Symbolics as well. The objective of this research effort was to extend this effort to the model by Lyon and Mead and compare and contrast simulation results between the models. In particular we concentrated on comparing the Payton and Lyon-Mead models.

Of particular interest was the impulse response at selected locations along the basilar membrane. We then proposed to perform Fourier analysis of the impulse responses in order to compare magnitude and phase responses in the frequency domain. This could then be related to published results for a variety of auditory models, particularly the models by Allen and Sondhi.

III. THE PAYTON MODEL

Payton [15,16] constructed a model of the entire auditory periphery using previously developed models to simulate the individual subsystems: middle ear (Guinan and Peake [8]), basilar membrane (Allen and Sondhi [2]), and hair-cell transduction (Brachman [6]).

Specifically regarding the cochlear model, the system simulates basilar membrane mechanics based on the numerical solution of a differential equation relating membrane displacement to stapes velocity, membrane compliance, and membrane damping. For purposes of analyzing the behavior of the model up to the membrane/hair-cell level, we have included the response sharpening filter model but not the hair cell transduction model. This facilitates comparison with the Lyon-Mead model described in the next section.

Payton in the introduction of [16] comments on the limitations of linear filter models: "Using vowels as stimuli, Sinex and Geisler (1984) compared the predicted responses of linear filters, designed to match threshold tuning curves, with those of auditory-nerve fibers. Their results indicated that, although linear filters were able to predict some features in neural responses, they were unable to predict significant neural response characteristics." Payton indicates that this shortcoming should be addressed by adding a nonlinear synaptic mechanism.

The middle ear model consists of an analog linear filter with one zero at the origin and four broadly tuned poles. This creates a transfer function with a 6dB/octave rise up to 1100Hz, a 6dB/octave fall between 1100Hz and 9400Hz, and an 18dB/octave rolloff above 9400Hz. This serves to remove DC and high frequencies from the model which improves the numerical simulation.

Payton used the basilar membrane model described in Allen [1] and Allen

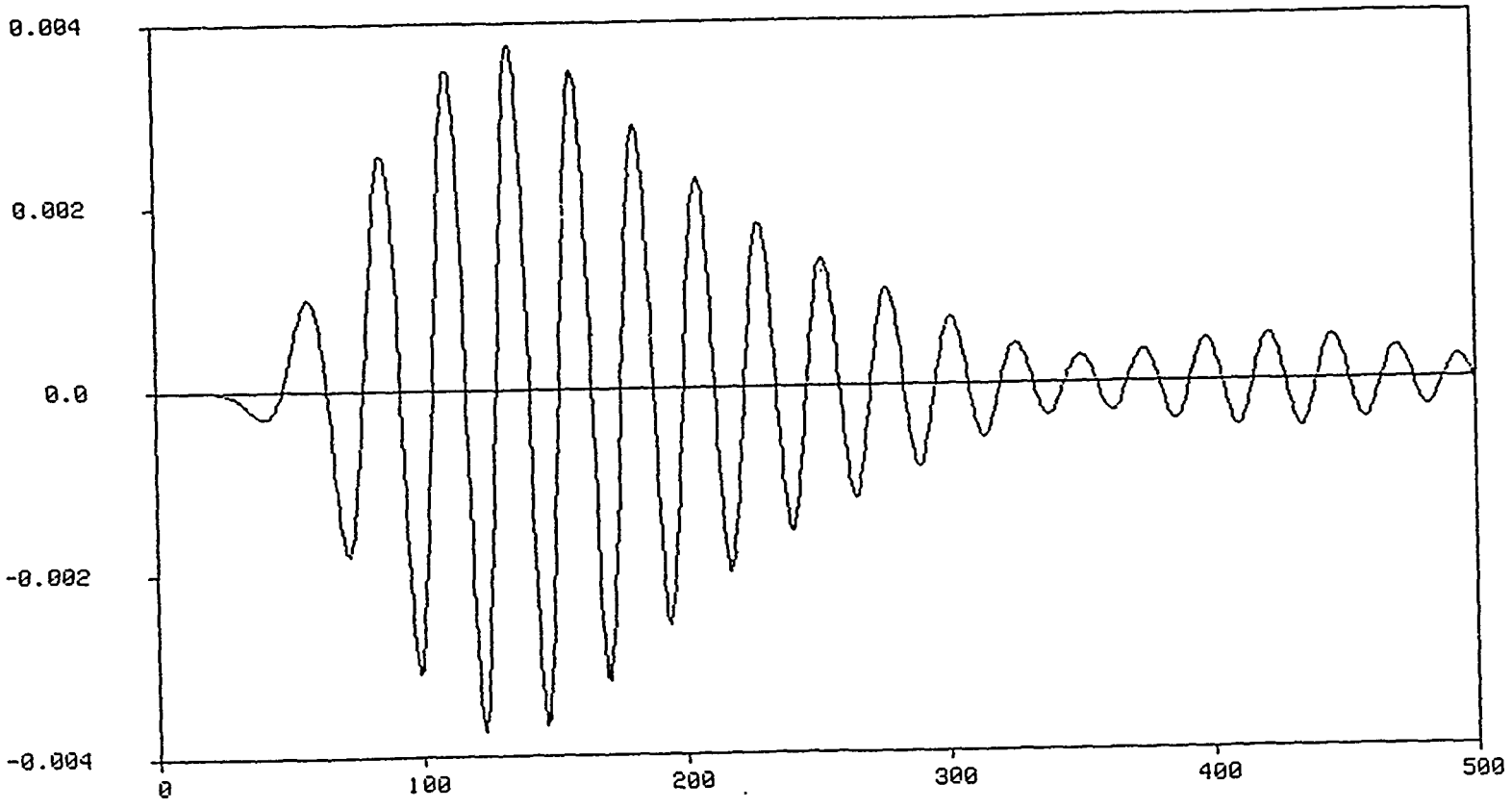


Figure 1(a). Impulse response of the Payton model for the basilar membrane tap corresponding to the pseudoresonance frequency of 7000Hz. Horizontal axis corresponds to samples at a 160kHz sampling rate. The entire interval is then 3.125ms long.

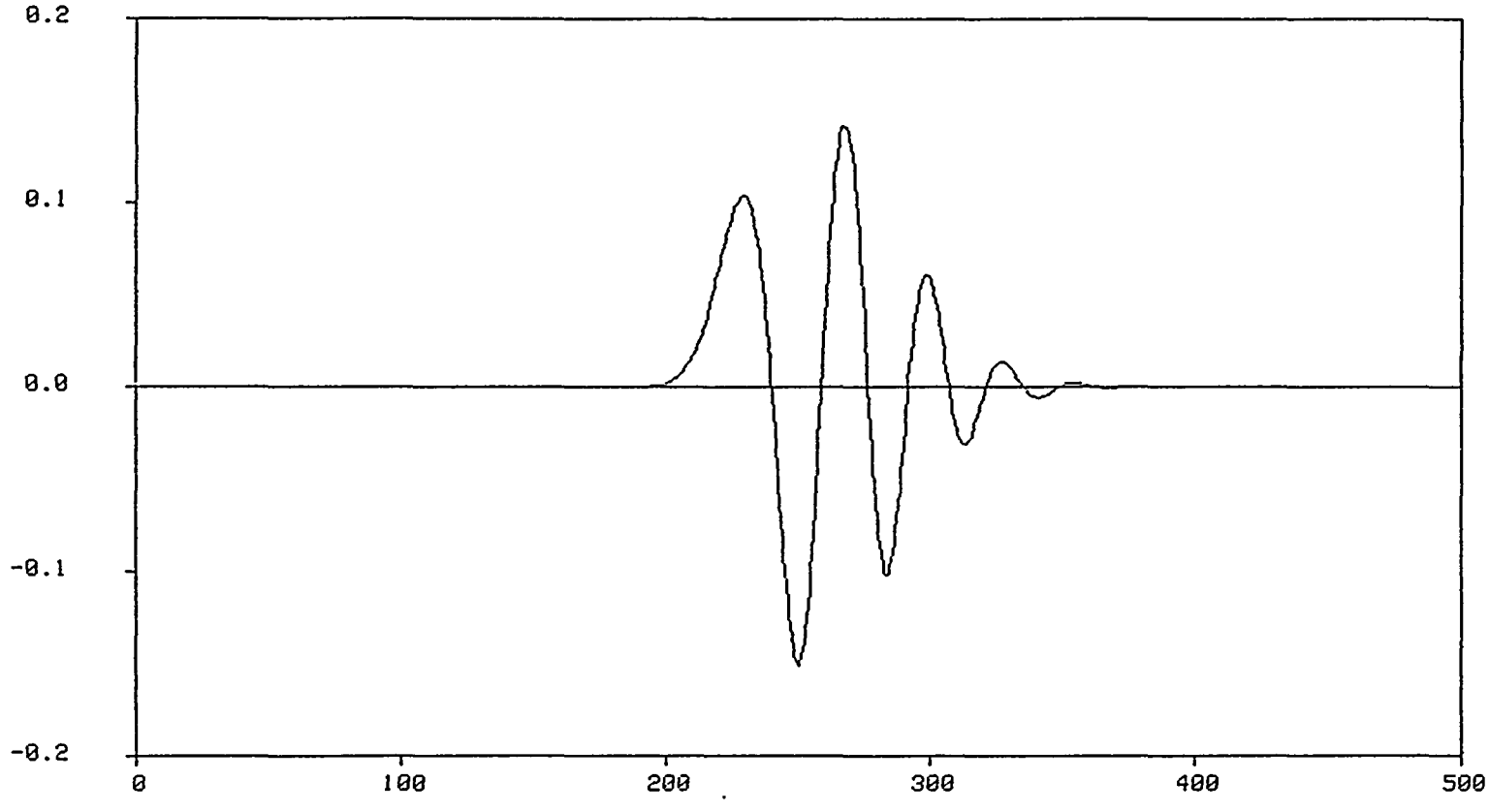


Figure 1(b). Impulse response of the Lyon-Mead model for the basilar membrane tap corresponding to the pseudoresonance frequency of 7000Hz.

and Sondhi [2] which solves the membrane plate equation

$$b(x,t) = -[D_z\pi^4/\omega^4(x)]\xi(x,t) - R(x)\dot{\xi}(x,t) - 2(L-|x|)\rho\dot{I}_2(t) \qquad (1)$$

where $\xi(x,t)$ is membrane displacement, x is the distance along the membrane, $D_z\pi^4/\omega^4(x)=K(x)$ is the membrane compliance, R(x) is the membrane damping term, L is the length of the membrane, ρ is the fluid density, and I_2 is the stapes velocity. Then the acceleration of the membrane at a particular location is given by

$$\ddot{\xi}(x,t_0) = F^{-1}\left[\frac{F[b(x,t_0)]}{F[Q(x)]}\right] \qquad (2)$$

from which we can derive membrane displacement with a double time integration. Q(x) represents the boundary conditions. Allen and Sondhi compute the solution in the spacial frequency domain using forward and inverse fast Fourier transforms (FFTs) as shown in (2).

In order to match physiological data, Payton incorporated a second order sharpening mechanism normally presumed to describe a tectorial membrane resonance [21]. The sharpening mechanism is described by a two-pole, two-zero filter which has coefficients "chosen to provide qualitative fits to published neural tuning curves (Kiang et. al., 1965; Allen 1983)."

Catenating the middle ear, basilar membrane, and sharpening mechanism models gives a system which we characterize as a bank of linear filters. Each filter then has an impulse response with associated magnitude and phase frequency responses. Figure 1(a) shows a typical impulse response of the Payton membrane model corresponding to a

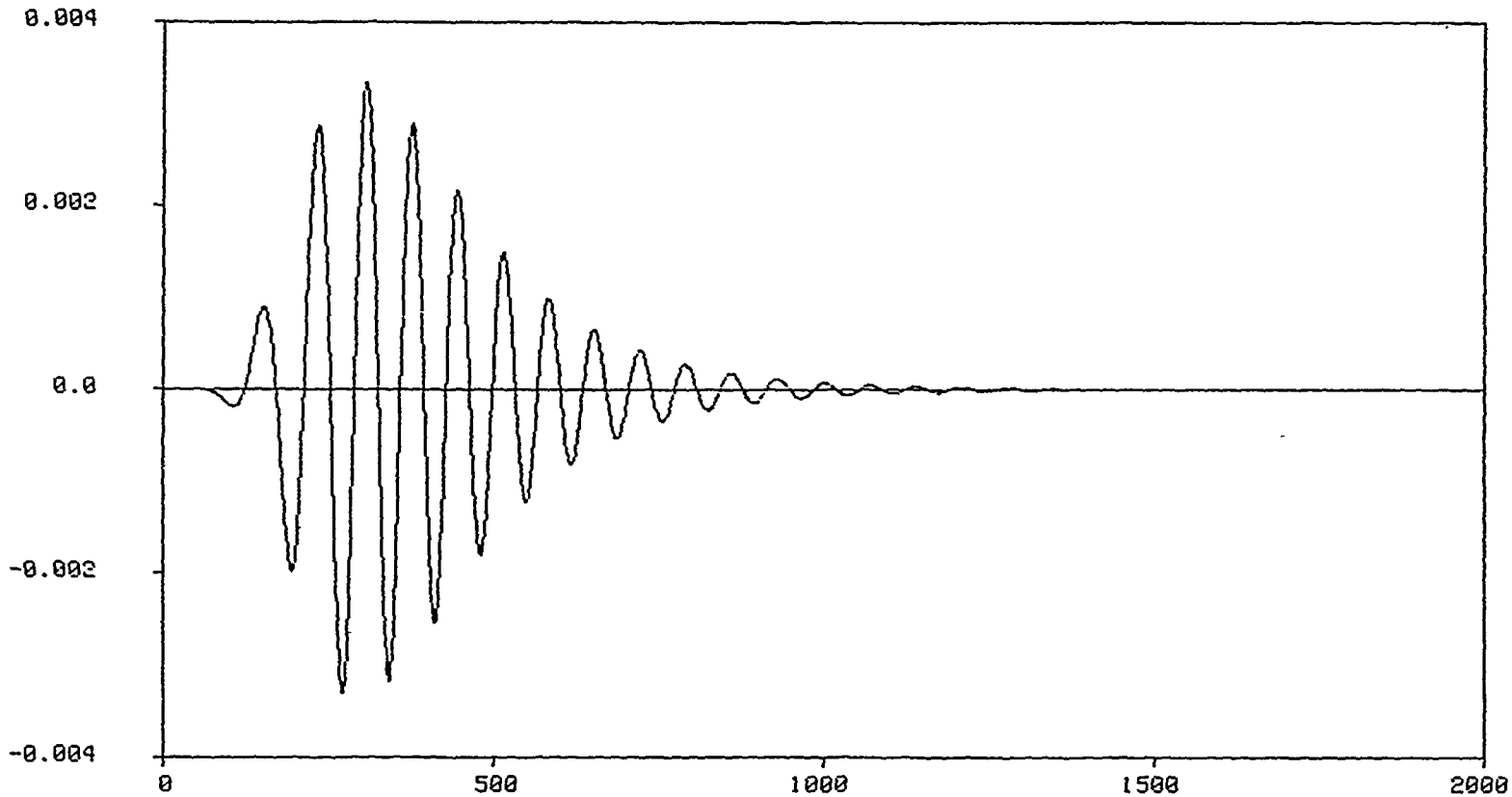


Figure 2(a). Impulse response of the Payton model for the basilar membrane tap corresponding to the pseudoresonance frequency of 2100Hz. Horizontal axis correspons to samples at a 160kHz sampling rate. The entire interval is then 12.5ms long.

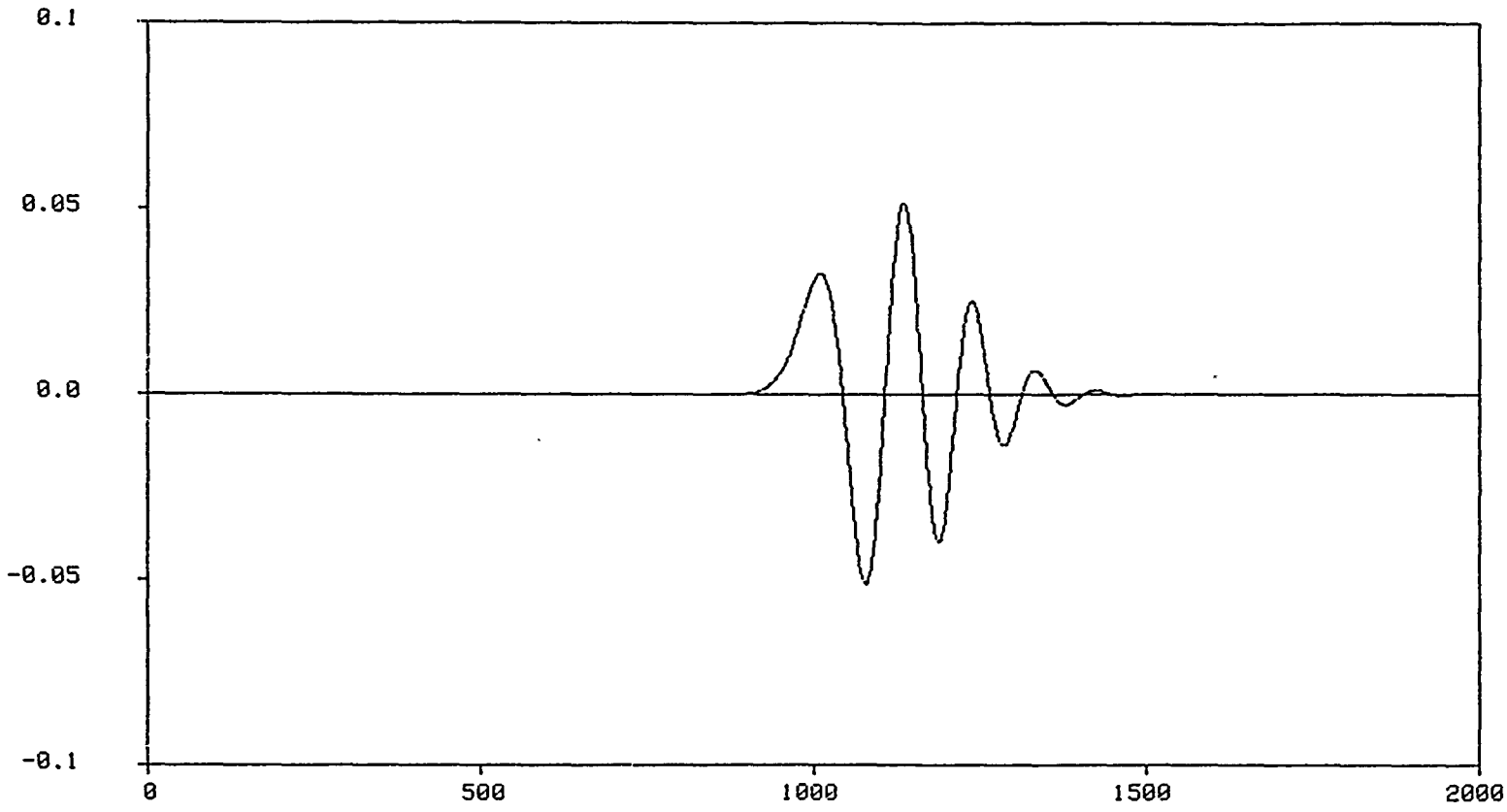


Figure 2(b). Impulse response of the Lyon-Mead model for the basilar membrane tap corresponding to the pseudoresonance frequency of 2100Hz.

pseudoresonance frequency of about 7000Hz. The narrow bandwidth causes the large number of oscillations at the center frequency of the filter. The response also indicates a group delay of approximately 0.8ms. Figure 2(a) shows the impulse response for a location on the basilar membrane corresponding to a pseudoresonance frequency of about 2100Hz. The group delay for this response appears to be about 1.8ms.

IV. THE LYON-MEAD MODEL

Lyon and Mead [11,12] have interpreted the basilar membrane as a nonuniform transmission line with adaptive active undamping. In this model, the membrane is divided along its length into a cascade of 480 short sections each of uniform characteristic impedance. The signal at any point in the model represents an acoustical pressure wave. Each section is then modeled by a simple second order, two-pole analog lowpass filter with cutoff frequencies starting at 20kHz for the first section and logarithmically dropping to 20Hz for the last section.

The Q of the filters can be varied such that the peak amplitude response of a given filter will vary from 0dB to about +1dB from the DC response. A cascade of such filters reinforces the peak up beyond +10dB to a level dependent upon the Q value selected. The reinforcement ceases when successive filters have shifted their cutoff frequency below the frequency of measurement, at which time all successive filters attenuate the signal each at a 6dB/octave rate giving rise to the sharp cutoff observed in physiological membrane

measurements.

Lyon and Mead have departed from the philosophy that the sharp resonances observed in basilar membrane responses should be modeled by a linear filter sharpening model. They note that this phenomena can be explained by using a very broadly tuned membrane model but allowing the Q of the model to adapt with signal level. Payton touted nonlinear synaptic mechanisms as the way to improved auditory modeling. Lyon and Mead fit biological data not with a nonlinear model, but with an adaptive linear model. The Lyon-Mead model assumes an actively amplifying membrane which potentially can adapt to the signal level at all points along the basilar membrane.

Allen [5] comments in reference to multidimensional differential equation models of the basilar membrane system, "more time is needed to fully evaluate the significance of these more complicated calculations and models, but it presently appears that they do not close the gap, as was originally hoped, between the mechanical model and neural measurements of cochlear frequency response. Thus the most important problem which remains unsolved in cochlear theory is explaining the sharpness of tuning of the neural response." Allen then advances two theories, the tectorial resonance theory (Zwislocki 1979; Allen 1980) and the active undamping theory (Neely and Kim 1983; Neely 1981).

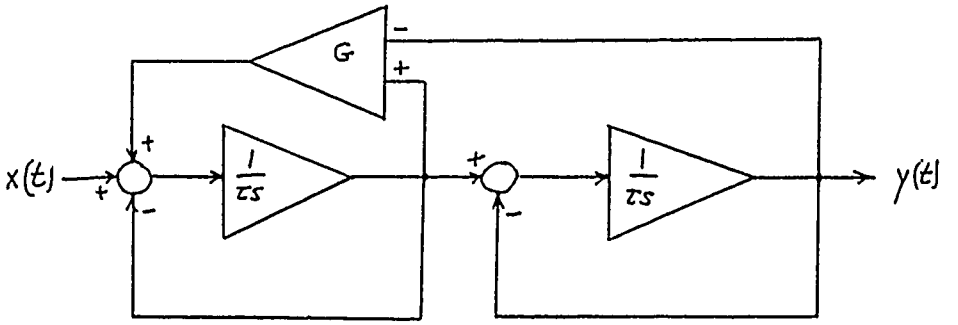
The problem with fitting neural tuning curves arises from the measurement techniques involved. A nerve fiber in the auditory system responds to stimulus level in a very nonlinear way. In addition the

characteristics of neural responses limit the dynamic range to 10-15dB in terms of the neural firing rate. The actual range depends on the stimulus level which affects neural sensitivity.

In order to counteract this effect, researchers create neural tuning curves using an iso-output strategy. Instead of maintaining the input level constant while varying the input frequency, they adjust the input level so as to maintain a constant output level. The curve representing the input level as a function of frequency indicates the selectivity of the system from tympanic membrane to nerve fiber. The problem arises in the fact that the input level is allowed to vary. According to the active undamping theory, this invokes adaptive mechanisms on the basilar membrane which make small changes in the Q for each local section. The overall effect however is quite pronounced due to the gain multiplying effect discussed previously. This gives the appearance of sharply tuned resonances on the basilar membrane.

V. DISCRETIZING THE LYON-MEAD MODEL

In order to execute the Lyon-Mead model on a digital computer it is necessary to transform the analog model into an equivalent digital form. This cannot be done exactly except as the sampling frequency approaches infinity. We have chosen the sampling rate to be 160kHz which is enough higher than the highest frequency component to yield results very close to those reported for the analog model. This particular frequency is also the sampling frequency chosen by Payton



$$\frac{Y(s)}{X(s)} = \frac{1}{\tau^2 s^2 + (2-G)\tau s + 1}$$

Figure 3(a). Prototype analog second order section of the Lyon-Mead model with corresponding analog transfer function. This is a simplified version of figure 7 in [10] with G1=G2=1. G=G3 can then vary between 0 and 2 with stable results.

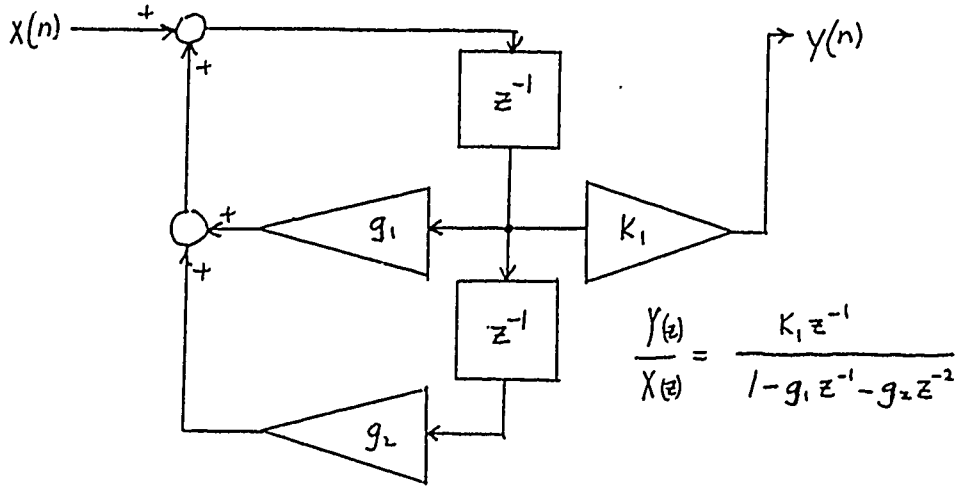


Figure 3(b). Prototype digital second order section of the Lyon-Mead model. The values of the constants are given by (6). Note the built-in unit delay.

and thus represents a natural choice to facilitate comparison.

In the interest of maintaining the characteristics of the analog Lyon-Mead model as closely as possible, we have selected the impulse invariant technique [9,17] to convert the analog impulse response into an equivalent digital response. This maintains a linear frequency relationship between the analog and digital models. The bilinear transform is used extensively since it will not alias filter responses. However it distorts the frequency axis to accomodate the finite digital frequency range. In addition, the bilinear transform will generate two additional transmission zeros at z=-1 as part of the frequency warping. This creates the need for one more multiplication and two more additions per section. The sampling rate could be lowered for the bilinear transform though, so there is a tradeoff to be investigated. We chose the impulse invariant technique to match the analog Lyon-Mead model as closely as possible even though it takes longer to execute.

Lyon uses a simple two-pole lowpass filter to model each membrane section. The transfer function selected is

$$H(s) = \frac{1}{\tau^2 s^2 + 1/Q\ \tau s + 1} \qquad (3)$$

where Q is the quality factor and τ is the reciprocal of the cutoff frequency in radians per second. Figure 3(a) depicts the flow graph section corresponding to this transfer function. To discretize this transfer function using the impulse invariant technique, we separate H(s) into a sum of two separate one-pole systems using the partial

fraction expansion

$$H(s) = \frac{1}{\tau^2 s^2 + 2(1-\alpha)\tau s + 1}$$

$$= \frac{j}{2\tau\sqrt{2\alpha-\alpha^2}} \left[\frac{1}{s-\alpha/\tau+1/\tau+(j/\tau)\sqrt{2\alpha-\alpha^2}} - \frac{1}{s-\alpha/\tau+1/\tau-(j/\tau)\sqrt{2\alpha-\alpha^2}} \right]. \quad (4)$$

Each single pole function in the analog domain can then be translated into an equivalent discrete transfer function having impulse response samples identical to samples of the analog impulse response, hence the name impulse invariant transformation. Each pole of the form

$$\frac{1}{s-\alpha}$$

is then replaced with the z-domain transfer function

$$\frac{1}{1-e^{aT}z^{-1}} . \quad (5)$$

Each single-pole function is a simple exponential and can be added to form H(z). It is also necessary to multiply by T in order to equate frequency responses instead of impulse responses. Adding the z-domain poles for our transfer function results in

$$H(z) = \frac{Tj}{2\tau\sqrt{2\alpha-\alpha^2}} \left[\frac{1}{1-e^{\left[\alpha/\tau-1/\tau-(j/\tau)\sqrt{2\alpha-\alpha^2}\right]T}z^{-1}} \right.$$

$$\left. - \frac{1}{1-e^{\left[\alpha/\tau-1/\tau+(j/\tau)\sqrt{2\alpha-\alpha^2}\right]T}z^{-1}} \right]. \quad (6)$$

After simplification we have a digital transfer function of the form

$$H(z) = \frac{k_1 z^{-1}}{1-2r\cos\theta z^{-1}+r^2 z^{-2}} \qquad (7)$$

where $r = \exp(-T/2Q\tau)$

$$\theta = \frac{T}{\tau\sqrt{1 - 1/4Q^2}}$$

$$\tau = 1/2\pi f_0$$

$$k_1 = \frac{T\ r\ \sin\theta}{\tau\sqrt{1 - 1/4Q^2}}$$

Figure 3(b) depicts the flow graph section corresponding to the discretized model.

Table I. Lyon-Mead model membrane tap locations in terms of filter section index (corresponding to membrane place) and pseudoresonance frequency. This Lyon-Mead model implementation contains 480 sections as did the CMOS integrated circuit built by Lyon and Mead.

Lyon-Mead tap number	section index	cutoff frequency in Hertz	
0	73	6979	first tap
1	83	6042	
2	94	5155	
3	104	4463	
4	115	3808	
5	125	3297	
6	135	2854	
7	146	2435	
8	156	2108	
9	167	1799	
10	177	1557	
11	188	1329	
12	198	1150	
13	209	981	
14	219	850	
15	229	735	
16	240	627	
17	250	543	
18	261	463	
19	271	401	last tap

Table I gives the frequencies corresponding to the 20 membrane taps chosen to agree with the frequencies selected by Payton in her model. Once the signal has propagated to the first tap at section 79, the effect of cumulative gains and losses has nearly achieved steady state. Prior to this tap a gain accumulating processs takes place. Careful inspection of figure 3(b) indicates a small rise from the earliest tap (highest frequency) to the remaining taps (lower frequencies).

VI. SIMULATION RESULTS

Figures 1(b) and 2(b) depict the impulse response of the Lyon-Mead model for the same pseudoresonance frequencies of 7000Hz and 2100Hz used to analyze the Payton model in figures 1(a) and 2(a). The number of significant cycles drops from more than 10 down to 3 or 4. This indicates a substantial decrease in the Q of the system. Bear in mind that the Lyon model recovers the apparent narrow tuning by adapting the Q of each section with signal level.

Table II. Filter group delays for the Payton and Lyon-Mead models at 7000Hz and 2100Hz in time and phase. Physiological data by Rhode indicates that 9π is a typical total phase at pseudoresonance.

frequency	Payton Model		Lyon-Mead Model	
	seconds	radians	seconds	radians
at 7000Hz	0.8ms	11.2π	1.6ms	22.4π
at 2100Hz	1.8ms	7.5π	7.2ms	30.2π

Note that the group delay of the membrane taps has increased greatly. The delay for figures 1(b) and 2(b) appear to be about 1.6ms and 7.2ms respectively. Comparing to the results for figures 1(a) and 2(a), this represents a factor of 2-4 increase. Table II compares the time delays of the two models at these two taps in seconds and radians. The delays computed here for the Lyon-Mead model agree with the delays observed in figure 15 in Lyon and Mead [11].

According to Allen [1] in reference to Rhode the target phase is around 9π. Thus the Payton model comes closer to the physiological phase data than the Lyon model. This poses no great threat to the digital Lyon-Mead model since the delay incurred in each section can be easily adjusted. Specifically, one could increase or decrease the output phase by systematically reconnecting some of the output connections. Normally the output is attached between the two filter delays blocks in

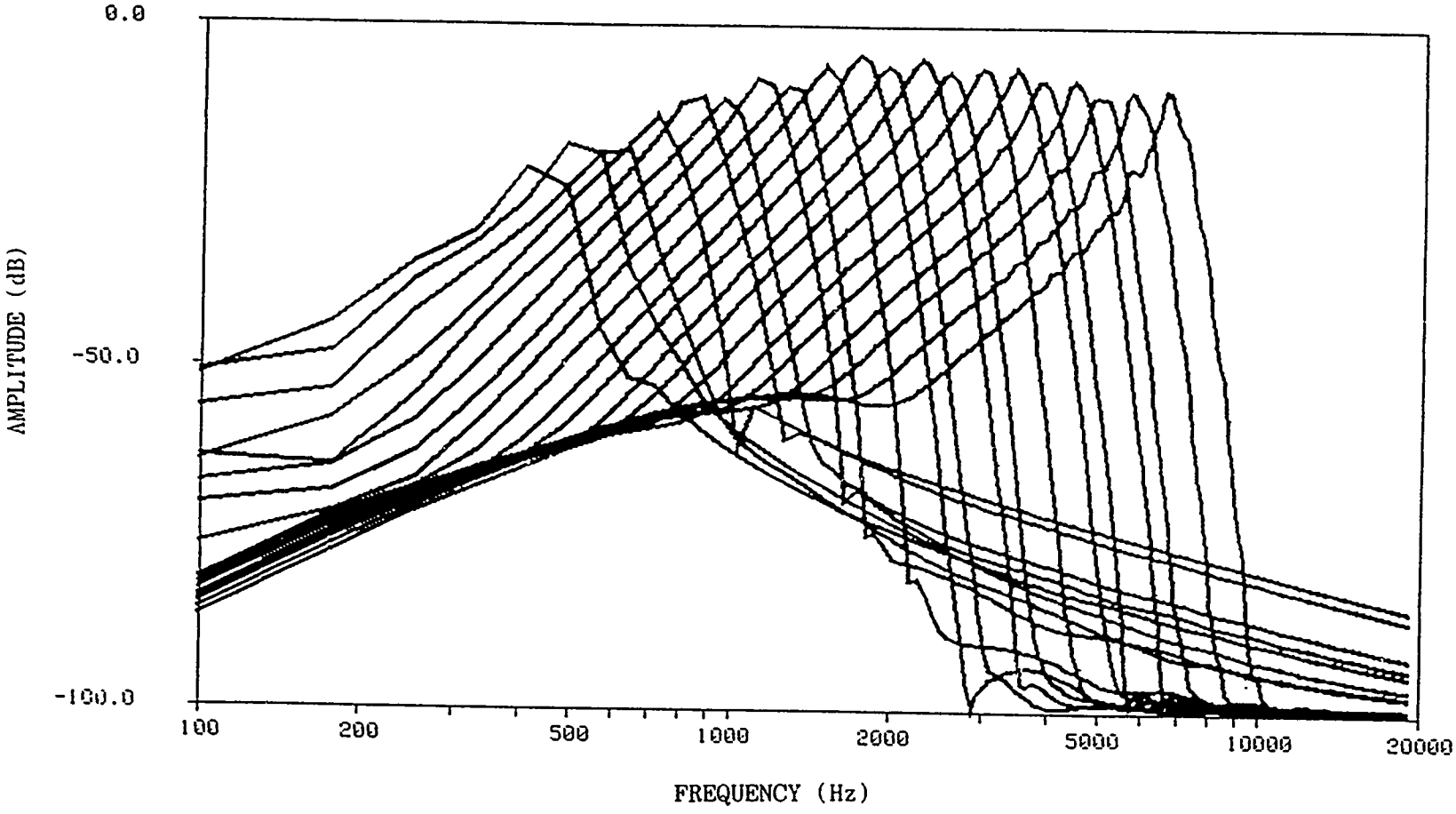


Figure 4(a). Composite of 20 magnitude frequency responses of the Payton model for the 20 membrane taps chosen by Payton. This set of curves agrees with Payton [16].

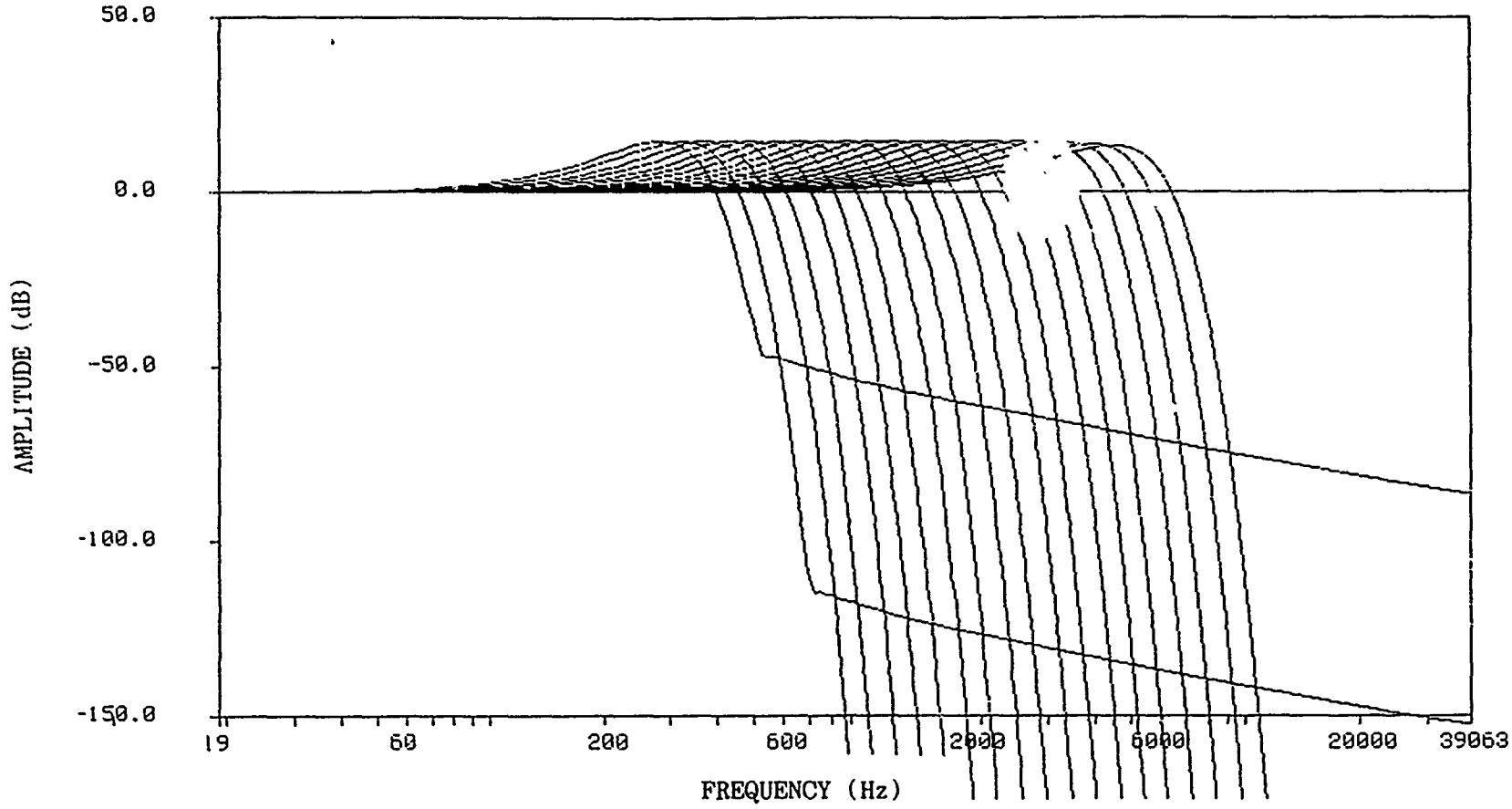


Figure 4(b). Composite of 20 magnitude frequency responses of the Lyon-Mead model with Q=0.8 for all sections and the 20 membrane taps as chosen by Payton. The two lowest frequency responses are inaccurate below -50dB and -100dB respectively as the responses were truncated slightly due to the finite analysis window length.

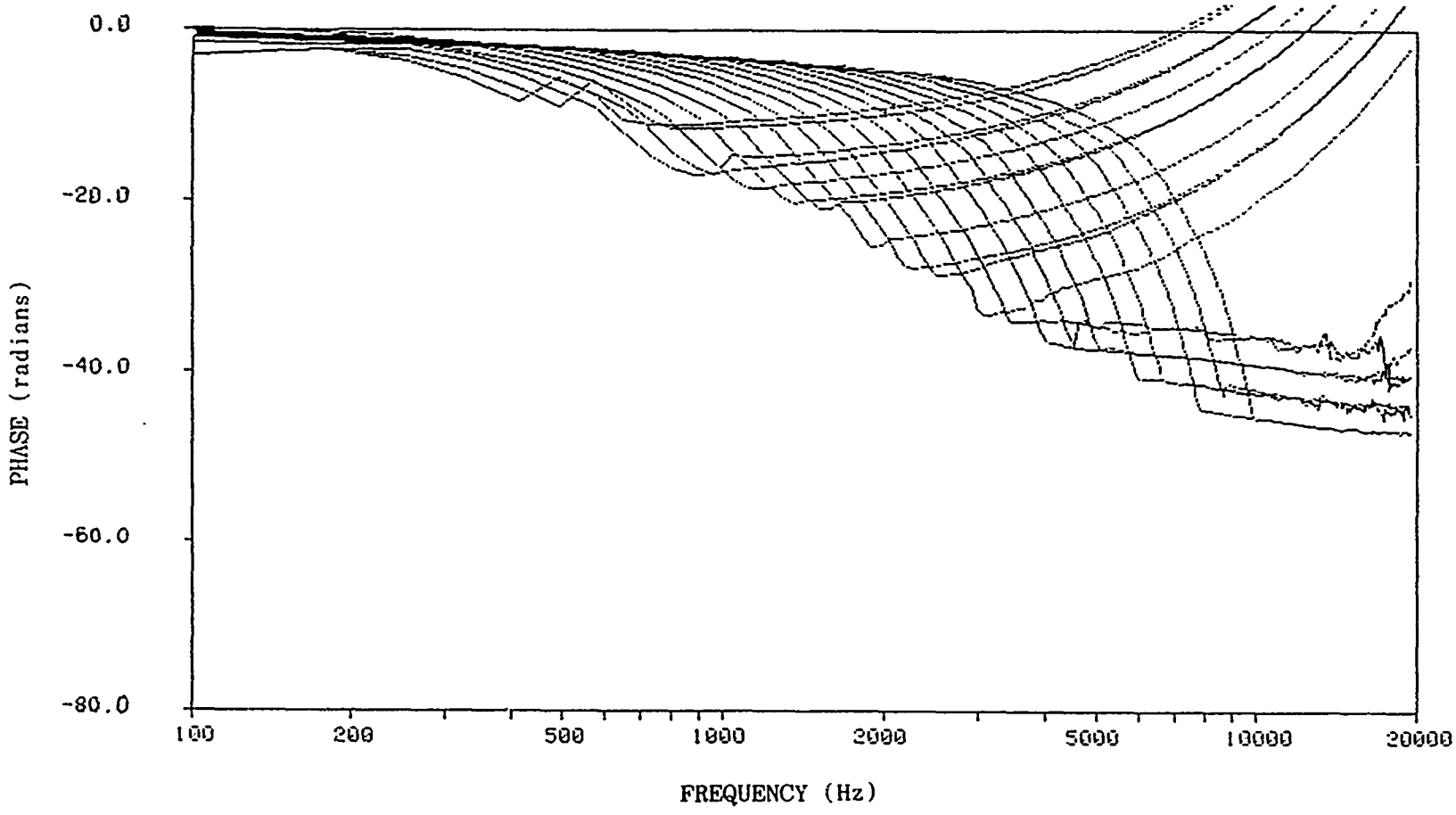


Figure 5(a). Composite of 20 phase frequency responses of the Payton model for the 20 membrane taps chosen by Payton. The phase was automatically unwrapped before plotting. The log scaling effect discussed by Lyon and Mead [11] is clearly visible.

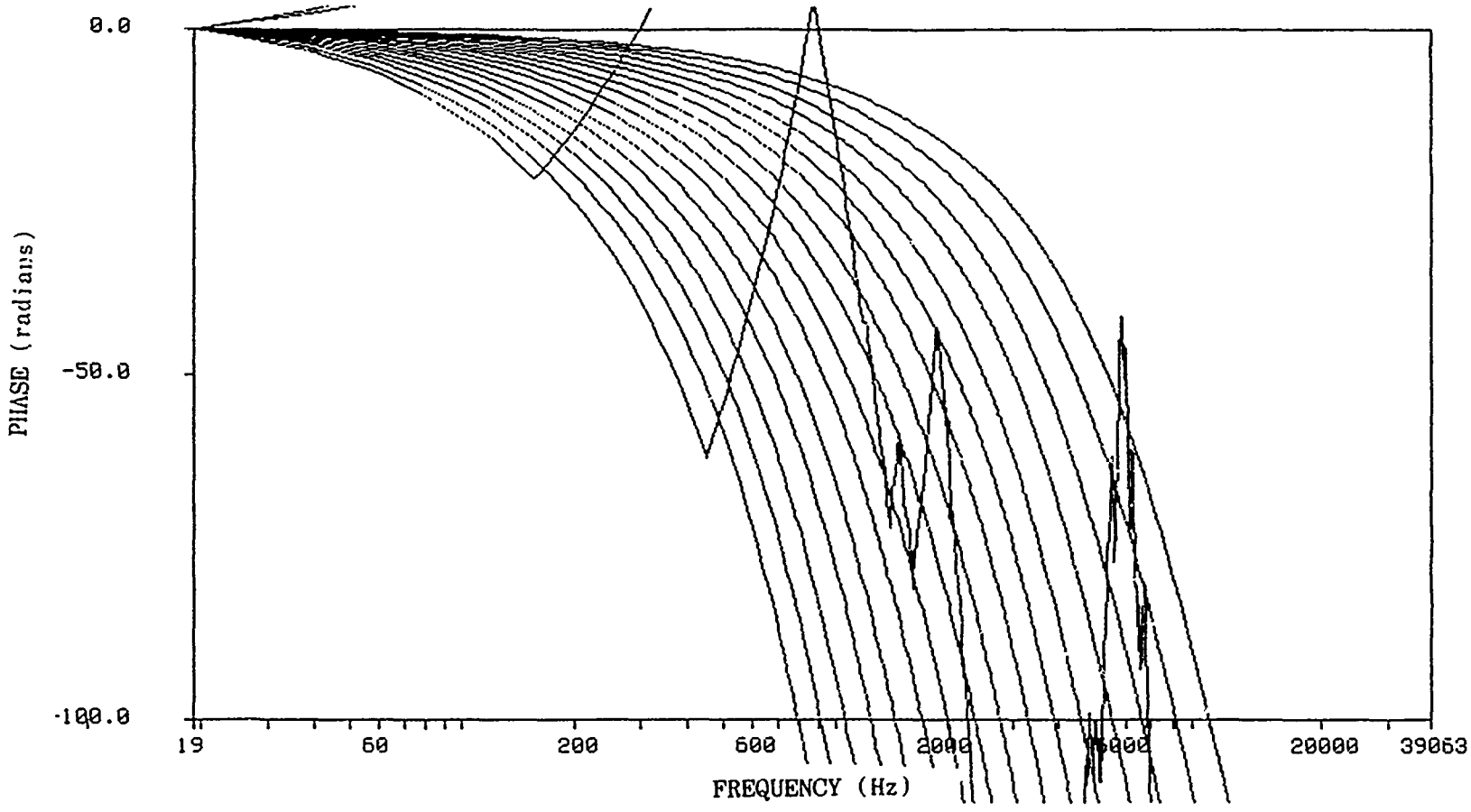


Figure 5(b). Composite of 20 phase frequency responses of the Lyon-Mead model with Q=0.8 for all sections and the 20 membrane taps chosen by Payton. There is no characteristic plateau effect (see Allen [1]) for the Lyon-Mead model as was present in the Payton model. The phase unwrapping algorithm produces erratic results for the lower frequency responses when the phase step exceeds π, but otherwise shows a log scaling effect.

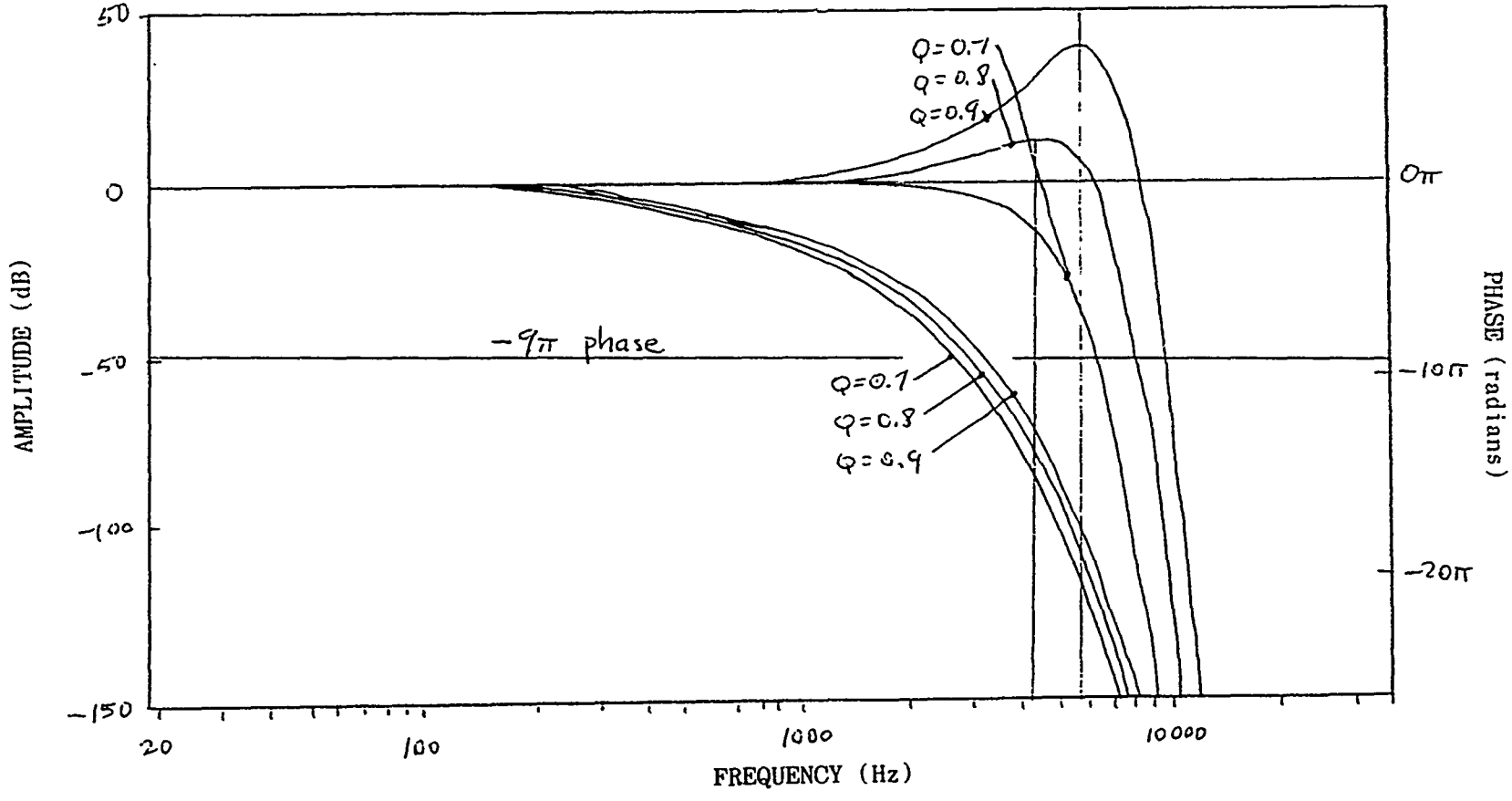


Figure 6. Composite of Lyon-Mead model frequency response for Q=0.7, 0.8, and 0.9 for tap 73 with a cutoff frequency of 7000Hz. Note that the magnitude response increases some 35dB while the phase response changes very little.

figure 3(b). Attaching the output before (after) the delays will decrease (increase) the phase at successive taps.

The Payton model incorporates static narrow tuning whereas the Lyon-Mead model relies on dynamic adaptive Q. As a result, the magnitude and phase responses of the two models produce widely varying bandwidth effects. Figure 4(a) shows a composite of 20 magnitude frequency responses for the Payton model with the membrane taps chosen to agree with the taps selected by Payton for her research. figure 4(b) shows the same collection of reponses for the Lyon-Mead model for the same pseudofrequencies as in the Payton model. The Lyon-Mead response looks broader even though the vertical axis spans twice as far. The Q value of 0.8 is slightly greater than the passive threshold of 0.7071 and yields peak gains in the +10 to +15dB range. By definition the Payton model is passive.

Refer to figures 5(a) and 5(b) to compare phase responses of the two models. Some of the lower frequency curves are erratic due to the limited performance of automatic phase unwrapping. In spite of this and the slight difference in scale, it is apparent that the Lyon-Mead model demonstrates much greater phase delay than the Payton model in particular and most of the simulations available in the literature in general.

Figure 6 shows an overlay of the magnitude and phase responses of the Lyon-Mead model with the model Q varying from 0.7 to 0.9. The phase characteristic does not change much but the effect on the magnitude is

dramatic. A line is draw in at the 9π level suggested by Allen in [1]. The Lyon-Mead model clearly exceeds this phase at pseudoresonance. Note also that the peak frequency increases with increasing Q.

Fitting the Lyon-Mead model to physiological data consists of determining an appropriate Q-adaptation scheme. Lyon shows the results of one possible adaptation strategy as figure 4 in [11] where the Q varies from 0.7 to 0.9 as a function of frequency because the level is a function of frequency. Unfortunately the method used was not given so we cannot reproduce this figure for the digital model for verification.

Computation of complex digital models with a sample rate of 160kHz consumes vast amounts of computer time during simulation runs. In order to speed up the system, the input could be lowpass filtered to 20kHz or lower before sending it into the cascade. This would permit the use of a lower sampling rate, but we would have to transform filter coefficients to accomodate the lower sample rate. Here the bilinear would be more attractive since frequency aliasing would occur if the sample rate is too low when using the impulse invariant technique.

This down-sampling process could be continued at key points along the cascade. The frequency characteristics of the membrane model include a sharp cutoff since each preceding filter attenuates frequencies above cutoff at a 6dB/octave slope. Once enough filters have processed the signal with successively lower cutoff frequencies, the signal can be decimated by a factor of two. This process can be repeated every

octave along the cascade to create lower and lower sampling rates as the filter cutoff frequencies tend toward the 20Hz lower limit. This downsampling is subject to the same aliasing or frequency distortion problems in choosing the initial sampling rate.

Table III. Multirate strategy to accomplish a computational speedup for the Lyon-Mead model and comparison with two fixed sample rate strategies.

Lyon-Mead section index	sample rate in Hertz	sections per second to compute
0-58	40,000	59x40000 = 2360000
59-106	20,000	48x20000 = 960000
107-154	10,000	48x10000 = 480000
155-202	5,000	48x5000 = 240000
203-250	2,500	48x2500 = 120000
251-298	1,250	48x1250 = 60000
299-346	625	48x625 = 30000
347-394	312.5	48x312.5 = 15000
395-442	156.25	48x156.25 = 7500
443-479	78.125	37x78.125 = 2891

TOTAL = 4.275×10^6

TOTAL AT 40kHz = 480x40000 = 19.2×10^6

TOTAL AT 160kHz = 480x160000 = 76.8×10^6

Table III indicates one possible downsampling strategy and the corresponding improvement in execution speed. As it stands, the digital Lyon-Mead model completes 8192 samples through each of 480 sections in about one hour of execution time on our Texas Instruments Explorer I lisp machine. As indicated by Table III, a factor of 18 improvement results from downsampling bringing the lisp-based system to within a factor of 4000 of real-time execution.

The digital Lyon-Mead model could be partitioned into a number of cascade segments, each of which could be executed in real time. With a processor executing five million operations per second, each section would take about one microsecond to compute. Using the total number of sections to compute given by Table III for the decimation strategy gives a total number of processors of $5*4.275x10^6/5x10^6 = 4.275$. Note that the highest frequency sections 0-58 represent 55% of the computational burden. Suitable processors exist in both fixed point and floating point formats.

VII. RECOMMENDATIONS

After considering the differences between the Payton and Lyon-Mead models, it becomes apparent that the adaptive, active undamping concept has great potential to explain the apparent narrowly tuned neural response peaks based on iso-output measurement techniques. The models cannot be directly compared as they have been constructed under disparate assumptions of the underlying mechanisms in the auditory

process. This results in significantly different signals on the basilar membrane which is the focus of attention for this study.

In this research we have made no attempt to create an adaptive Lyon-Mead model. Indeed at this time, Lyon and Mead have not reported on an adaptive implementation except for figure 4 of [11].

The next logical step in this research is to add an adaptation scheme so as to more realistically simulate the auditory system. This would necessarily require a review of the literature to obtain a body of data characterizing the adaptive behavior of the auditory system. Measurement techniques would have to be taken into account: iso-output, steady-state input, pulsed input, etc. Thus a review of the physiological literature is in order to assertain details about the extent and timing of observable adaptation mechanisms. Static and dynamic adaptation mechanisms need to be distinguished.

A smaller step for research would be to rework the model using the bilinear transform and analyze the distortions created in the frequency responses. The success of this study would permit the lower sampling rates suggested in Table III. This would accelerate the research simulation model and move us closer to a real-time implementation. Perhaps a real-time digital cochlea will offer the repeatability and flexibility required for the front end of speech processing systems.

Timothy R. Anderson of the Biodynamics/Bioacoustics Laboratory will maintain a copy of the latest software used in the models reported

herein. For the sake of brevity, computer listings will not be included in this final report. All software was written in Commonlisp and requires no additional software products to execute.

REFERENCES

[1] Allen, Jont B. (1977). "Two-dimensional cochlear fluid model: New results," J. Acoust. Soc. Am. 61, pp. 110-119.

[2] Allen, Jont B., and Man M. Sondhi (1979). "Cochlear macromechanics: Time domain solutions," J. Acoust. Soc. Am. 66, pp. 123-132.

[3] Allen, Jont B. (1980). "Cochlear micromechanics - A physical model of transduction," J. Acoust. Soc. Am. 68, pp. 1660-1670.

[4] Allen, Jont B. (1983). "Magnitude and phase frequency response to single tones in the auditory nerve," J. Acoust. Soc. Am. 73, pp. 2071-2092.

[5] Allen, Jont B. (1985). "Cochlear modeling," IEEE ASSP Magazine, pp. 3-29, Jan 1985.

[6] Brachman, M. L. (1980). "Dynamic response characteristics of single auditory-nerve fibers, Special Report IRS-S-19," Ph.D. thesis, Institute for Sensory Research, Syracuse University, Syracuse, NY.

[7] Covington, C. David and Michael K. Ellis (1988). "Auditory modeling," Final Report, AFOSR contract F49620-86-R-0004.

[8] Guinan, J. J. and W. T. Peake (1967). "A circuit model for the cat's middle ear," Res. Lab. Elect., Q. Prog. Rep. 84, pp. 320-326.

[9] Jackson, Leland B. (1989). Digital Filters and Signal Processing, Kluwer Academic Publishing, Boston.

[10] Kiang, N. Y.-S., T. Watanabe, E. C. Thomas, and L. F. Clark (1965). Discharge patterns of single fibers in the cat's auditory nerve, MIT, Cambridge, MA.

[11] Lyon, Richard F. and Carver Mead (1988a). "An analog electronic cochlea," IEEE Trans. Acoust., Speech, and Signal Processing, vol. ASSP-36, pp. 1119-1134, July 1988.

[12] Lyon, Richard F. and Carver Mead (1988b). "Cochlear hydrodynamics demystified," Caltech Computer Science Technical Report Caltech-CS-TR-88-4, Feb. 1988.

[13] Neely, S. T. (1981). "Fourth-order partition dynamics of a two-dimensional model of the cochlea," Doctoral dissertation, Washington Univ., St. Louis, MO.

[14] Neely, S. T., and D. O. Kim (1983). "An active cochlea model showing sharp tuning and high sensitivity," Hearing Research 9, pp. 123-130.

[15] Payton, Karen L. (1986). "Vowel processing by a model of the auditory periphery," Ph.D. thesis, Johns Hopkins University, Baltimore, MD.

[16] Payton, Karen L. (1988). "Vowel processing by a model of the auditory periphery: A comparison to eighth-nerve responses," J. Acoust. Soc. Am. 83, pp. 145-162.

[17] Proakis, John G. and Dimitris G.. Manolakis (1988). Introduction to Digital Signal Processing, Macmillan, New York, 1988.

[18] Seneff, Stephanie (1988). "A joint synchrony/mean-rate model of auditory speech processing," Journal of Phonetics, vol. 16, pp. 55-76.

[19] Sinex, D. G. and C. D. Geisler (1984). "Comparison of the responses of auditory-nerve fibers to consonant-vowel syllables with predictions from linear models," J. Acoust. Soc. Am. 76, pp. 116-121.

[20] Sondhi, M. M. (1978). "Method for computing motion in a two-dimensional cochlear model," J. Acoust. Soc. Am. 63, pp. 1468-1477.

[21] Zwislocki, J. J., and E. J. Kletsky (1979). "Tectorial Membrane: A possible effect on frequency analysis in the cochlea," Science 204, pp. 639-641.

Assessing the Cognitive Demands of Tracking Strategies

Research Initiation Program Final Report

Submitted by

Barry P. Goettl
Assistant Professor
Department of Psychology
Clemson University

FINAL REPORT

Title of project:	Assessing the Cognitive Demands of Tracking Strategies
Starting date:	January 1, 1989
Ending date:	December 31, 1989
Amount Funded:	$19,995
Funding agency:	Air Force Office of Scientific Research
Contract Number:	F49620-88-C-0053/SB5881-0378
Purchase Order Number:	S-210-9MG-078

Prepared by: _____

Barry P. Goettl, Ph. D.
Assistant Professor
Department of Psychology
108 Brackett Hall
Clemson, SC 29634
(803) 656-2831

Applicant Institution: Clemson University
Office of University Research
E-102 Martin Hall
Clemson, SC 29634-5107

ABSTRACT

Two experiments are reported that further research on tracking strategies. In Experiment 1 subjects perform a Sternberg memory search task concurrently with a tracking task. Central processing demands of both tasks as well as response demands of the Sternberg task are manipulated. Results suggest that strategies place equivalent demands on response related resources. Experiment 2 investigated possible demand characteristics of an optimum-maximum procedure (Navon, 1985). On-line feedback was varied between groups of subjects. In spite of small group differences no effects of feedback were obtained.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and Universal Energy Sytems for funding this project sponsoring this project. In addition, I would like thank Dr. Michael Vidulich and Gary Reid for asststance in obtaining technical support and enriching discussions. Special thanks are extended to Mark Crabtree of SRL for providing technical advice, Jim Berlin and Jane Joseph for software development.

Assessing the Cognitive Demands of Tracking Strategies

Barry P. Goettl
Clemson University

1. INTRODUCTION

1.1. Problem Statement

Throughout the course of a mission the number of concurrent activities performed by the pilot greatly varies; so too do the cognitive demands. To insure that the cognitive demands placed on the pilot do not exceed the pilot's processing capacity, it is important to understand human cognitive processing capabilities and limitations. The most important task faced by the pilot is maintaining control of the aircraft. Thus, it is important to understand the cognitive demands of this task.

Evidence suggests that the demands of maneuvering the aircraft change over the course of the mission. For example, during landings, it is vitally important to maintain the proper glide path. Precision is required to insure safety, thus, cognitive demands are high. However, during straight and level flight, it is often not crucial to maintain the command flight path. The task of controlling the aircraft is less demanding and more effort can be

devoted to other mission-oriented tasks. Human operators may deal with the changing demands of the flight task by employing different control strategies.

When controlling a system having higher-order dynamics, such as an aircraft, the operator has two different strategies available: a continuous strategy and a double-impulse strategy. The goals of the continuous strategy are to minimize deviations while maintaining a smooth ride. The continuous strategy is described by linear models of manual control. These descriptive models assume the operator perceives a continuous input display and translates it into a continuous force exerted on a control stick to produce analog stick deflections (Young and Miery, 1965). In mathematical terms, the human is assumed to generate linear output in controlling the system.

The double-impulse strategy derives its name from the large non-linear stick deflections characteristic of this strategy. The major goal of the double-impulse strategy is to minimize errors as rapidly as possible. This is accomplished by applying maximum velocity and acceleration to the system. Unlike the smooth linear control of the continuous strategy, stick manipulation with the double-impulse strategy is characterized by maximum stick deflections in one direction followed by full deflections in the opposite direction. Because of this non-linear characteristic it is often referred to as "bang-bang"

control.

The continuous strategy emphasizes a smooth ride and is typical of straight and level flight where maintaining the command flight path is not critical to accomplishing the mission. With the double-impulse strategy, deviations are reduced quickly but at the cost of comfort and increased potential for over-correction. However, because it emphasizes rapid error correction, the double-impulse strategy is characteristic of landings where the glide path must be precisely maintained.

Since these different strategies are available to the pilot, it is important to determine the cognitive demands of each. If the strategies place different cognitive demands on the pilot, then the amount of cognitive effort available for other tasks will be a function of the strategy being used. Knowing the demands of each strategy and the conditions under which each is likely to be used, the system designer can determine when task aids are needed and what type of aid will be most helpful. Knowledge of the processing demands of strategies can also indicate how the operator is overloaded and can suggest ways of reducing the load. Finally, knowledge about strategies can aid in training new operators how and when to use the strategies.

The present project continued earlier research assessing the cognitive demands associated with the different strategies of control (Goettl, 1988; Wickens &

Goettl, 1985). The goal of the project was to determine the quantitative as well as qualitative processing demands of each strategy. It is not only important to know that the two strategies differ in cognitive demands, but how they differ. Such knowledge will help determine when aids are needed and what form they should take.

1.2. Background and Rationale

Evidence suggests that the double-impulse strategy and the continuous strategies do indeed differ in cognitive demands. Such evidence is provided by laboratory observations, mathematical computation, and performance data. Smith (1962) and Wilde and Westcott (1962) have observed that non-linear control is typical in tracking tasks requiring acceleration compensation or when lead time constants exceed five seconds. These observations suggest that humans adopt the double-impulse strategy as a result of cognitive or performance limitations. This conclusion implies that the double-impulse strategy is less cognitively demanding or that it demands a different form of processing.

Young and Miery (1965) demonstrated the superiority of a pulse-type (on-off) controller over a continuous controller for certain high-order control tasks. They suggest that the superiority of pulse-type controllers for systems with higher-order dynamic, as well as the preference for the double-impulse strategy, reflects human limitations in performing the cognitive computations required of such

systems. The use of a continuous strategy in the control of a high-order system requires computation of the time integral of the control response. Such computation is avoided by a pulse-type controller or by adopting the double-impulse strategy with a continuous controller. Young and Miery illustrate that with the double-impulse strategy, the operator need only keep track of the total time a force is applied (i.e. how long a joystick is deflected). Thus, the double-impulse strategy seems to be adopted because of its relative computational ease.

Similarly, Hess (1979) showed that a "dual-loop" model of control (Hess, 1978) offers a rationale for the double-impulse strategy in higher-order systems. One of the basic assumptions of the model is that the human operator attempts to reduce the mental computations and, therefore, the cognitive load of integration. This reduction in computational load, assumed to be accomplished by generating control forces that are easily integrated, is represented in the model as a two-parameter non-linear element. The model can produce control force histories that are qualitatively similar to those produced by humans using a pulse-type controller.

This evidence demonstrates the mathematical differences between the two strategies and provides a rationale for the double-impulse strategy. The evidence also suggests that the cognitive load of the double-impulse strategy is lower

than that of the continuous strategy. However, questions concerning the psychological significance of the different strategies remain. These questions pertain to the cognitive processes associated with the different strategies. Is the double-impulse strategy less cognitively demanding than the continuous strategy? Are there qualitative differences in the processing demands, or are the differences strictly quantitative? These important questions must be answered through performance evaluation.

1.3. A Model for Assessing Resource Demands

To understand the different types of processing demands that could be placed on the human operator a model of human attention is necessary. Resource models of human attention are useful in describing such processing demands (Kahneman, 1973; Norman & Bobrow, 1975; Wickens, 1980, 1984). Modeling attention in terms of resources implies that a limited amount of processing capacity is available for performing any given task. The amount of processing demanded by the task is influenced by task difficulty, or the amount of effort required by the task. Concurrently performed tasks compete for the limited resources.

Wickens (1980) reviewed dual task literature, and developed a multiple resources model of attention that defines processing resources along three dimensions: modality, code of processing, and stage of processing. Input modality refers to source of the information. The two

primary sources of information are the visual system and the auditory system. Processing code refers to how the information is processed. Spatial information is distinguished from verbal information. Manual responses are distinguished from verbal responses. Finally, stage refers to the sequential nature of processing. Early processing stages (i.e. perception and central processing), are distinct from later stages of response execution.

Wickens' (1980) model assumes that only tasks demanding similar resources will compete for resources. The more two tasks overlap in terms of resource demands, the more they will compete. As a result of resource competition, concurrent performance will decline and resources may be traded between tasks resulting in performance "trade-offs." When tasks demand different resources, concurrent performance is good and the tasks do not trade-off. Wickens' model can be used to reveal both quantitative and qualitative differences in the processing demands of the different tracking strategies.

1.4. Dual Task Methodology

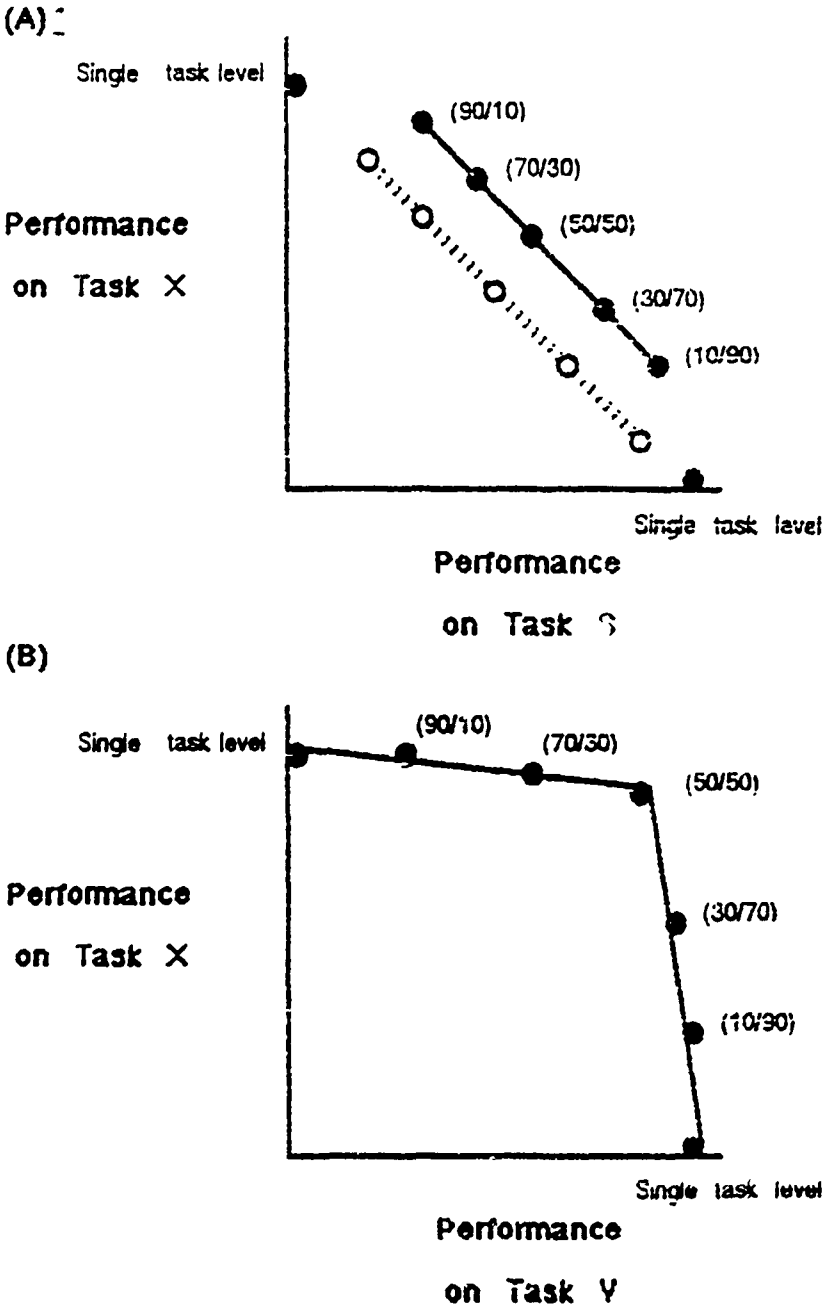
Using Wickens' (1980) multiple resource model, dual-task methodology can be used to assess the cognitive demands of various tasks. In this technique, the task to be assessed, called the primary task, is performed in dual-task configurations (i.e., performed concurrently) with one or more secondary tasks, the processing demands of which are

known. Performance measures determine the nature and extent of processing required by the primary task. The extent of dual-task interference and the specific patterns of interference and trade-offs reveal the nature of the resource demands, as the following example illustrates.

Consider the situation of attempting to determine the nature of the cognitive demands of Task X using the dual task procedure. Task X would be paired with different secondary tasks such as a spatial task (Task S), and a verbal task (Task V). Sometimes, subjects are asked to perform each dual task combination while varying their resource allocation policy. That is, subjects would be asked to devote specific amounts of processing resources to each task (e.g. 70% to Task X and 30% to Task S). Across trials the amount of processing resources devoted to Task X and Task S are varied (e.g. 90%/10%, 70%/30%, 50%/50%, 30%/70%, 10%/90%).

Performance is evaluated by plotting primary task performance (Task X) against secondary task performance (Task S and Task V). This plot is referred to as a performance operating characteristic (POC) curve. Figure 1 shows two theoretical POC curves. If the tasks demand similar resources, the pattern of results in Figure 1a is predicted. Dual-task performance of Task X declines relative to the single task level (shown on the vertical axis), performance on the two tasks trades-off as resources are shifted from one task to the other and as the

Figure 1. Hypothetical Performance Operating Characteristic (POC) curves showing results predicted for task pairs that demand (a) similar resources, and (b) different resources.



difficulty of Task S increases, and performance on S is maintained, performance on Task X decreases. The more difficult version of Task S is assumed to demand more resources and to draw them away from Task X.

Figure 1b depicts the data predicted when tasks demand different resources. Notice that none of the effects discussed above are predicted for tasks having different resource demands; (1) there is no single to dual task decrement, (2) tasks do not tradeoff as resource allocation is changed, (3) the performance on one task is not influenced by difficulty of the concurrent task.

1.5. Theoretical Criticisms

Navon (1984, 1985) has criticized dual task methodology in general and Wickens' multiple resource theory in particular. One of Navon's criticisms is that performance trade-offs resulting from resource allocation instructions may result not from re-allocation of shared resources but from demand characteristics. Navon claims that when subjects are instructed to "devote 70% of your resources to task A, and 30% to task B," they interpret this to mean "score well on task A and poorly on task B." The argument is over whether subjects actually re-allocate their resources or simply adjust their performance on both tasks to match the implied performance demands made by the experimenter.

To remove this type of demand characteristic, Navon

(1984) proposed an "Optimum-Maximum" procedure for dual task experiments. With this procedure, a specific performance level is requested for only one task, the optimized task. Subjects must perform the optimized task at the specified level (e.g. 70%). The other task, the maximized task, is to be performed as well as possible. The critical difference between this procedure and the conventional procedure is that a performance criterion is specified only for one task. The subject is instructed to devote all resources not allocated to the optimized task to the maximized task so as to "maximize" performance on it.

Navon (1985) compared the optimum-maximum procedure with the conventional resource allocation instructions to determine the extent to which demand characteristics account for performance data. Subjects performed two verbal tasks concurrently as resource allocation was manipulated using either one of the two resource allocation procedures. One task was a digit classification task, the other was a letter classification task. Both tasks demand verbal resources and would be predicted to show performance trade-offs by Wickens' model. However, performance trade-offs were obtained only when the conventional instructions (e.g. 70/30) were used, not when the optimum-maximum procedure was used. This result was interpreted as support for the argument that the conventional instructions impose strong demand characteristics that result in the observed trade-offs.

There are, however, several problems with Navon's (1985) experiment. First, single task performance was not gathered, so dual task performance cannot be thoroughly evaluated. Second, both tasks used were discrete in nature. Discrete tasks may allow, or even encourage, attention to be switched between tasks with no resulting performance decrements. Therefore, they may not be ideal for testing resource trade-offs (Wickens, Webb, & Fracker, 1987). Finally, the optimum-maximum procedure requires on-line feedback. The subjects must monitor their own performance continuously (through 2 on-line feedback displays) to insure they are performing at the proper level. This additional task imposes its own resource demands and may interfere with assessment of the demands of the primary task. Thus, although Navon's data seem to support his arguments, they can not be viewed uncritically. Navon's criticisms question dual-task methodology and the results obtained from such research. Therefore, Navon's method must be empirically investigated.

1.6. Overview of Experiments

Earlier research on tracking strategies served as the impetus for the current research. Wickens and Goettl (1985) examined the processing demands of tracking strategies using the dual-task methodology. In their experiment subjects performed a tracking task concurrently with a Sternberg memory search task. Subjects used the continuous strategy

on half the trials and the double-impulse strategy on the other half of the trials. Wickens and Goettl varied the structure of the Sternberg task in an attempt to determine processing demands via resource competition. Four variants of the Sternberg memory task were employed: visual-verbal, auditory-verbal, visual-spatial, and auditory-spatial.

Performance with the continuous strategy was not effected by the structure of the Sternberg task, but performance with the double-impulse strategy was effected by the Sternberg task input modality. When employing the double-impulse strategy, dual task decrements were small when tracking was paired with an auditory Sternberg task and larger when paired with visual Sternberg tasks. This result suggested that the double-impulse strategy demands more perceptual or central processing resources than does the continuous strategy. These results were somewhat unexpected because other research suggested that the continuous strategy demands more central processing than the double-impulse (Smith, 1962; Wilde & Wescott, 1962; Young & Miery, 1965).

Research conducted on the principle investigator's Summer Faculty Research Project (Goettl, 1988) further contributed to the inconsistency. In that experiment, subjects performed two tracking tasks concurrently. The cognitive load of one task was manipulated while subjects used one of the two strategies. The continuous strategy was

more adversely effected by increased cognitive load, suggesting that it demands more central processing load than the double-impulse strategy. Moreover, subjects using the continuous strategy did not perform as well as subjects using the double-impulse strategy. This group difference may indicate that the continuous strategy demands more resources than the double-impulse strategy at some other stage of processing.

This alternative hypothesis can be tested by manipulating response execution load. If the continuous strategy demands more response related resources than does the double-impulse strategy, then an increase in response resource demands of a concurrent task will disrupt the continuous strategy more than the double-impulse strategy. Thus, the critical test is to compare the effects of response execution load across tracking strategies. This is the objective of Experiment 1.

The other objective of the project is to further explore the criticisms of Navon (1984, 1985). The summer faculty research project used Navon's optimum-maximum procedure to investigate resource competition. The results were not extremely conclusive (Goettl, 1988). Some evidence of resource trade-offs were obtained (supporting multiple resources), however, the extent of the trade-offs were not as large as expected (favoring Navon's argument). One possible explanation of the results is that the optimum-

maximum procedure imposes its own demand characteristics. This hypothesis is addressed in Experiment 2 by examining the effects of on-line feedback on dual task performance.

2. EXPERIMENT 1: PROCESSING LOAD

2.1. Methods

2.1.1. Subjects. Six males and six females were paid $4.50 per hour to participate in eight one-hour sessions. All subjects were right-handed with normal or corrected-to-normal vision. Subjects were enrolled as undergraduate or graduate students at Clemson University and ranged in age from 17 to 30 years old.

2.1.2. Apparatus and Equipment. The tracking and Sternberg tasks were automated and run on a Zenith 258 computer. Tracking was performed with a Measurement Systems 546 joystick and responses to the Sternberg task were made with an SRL 4-key response box. An IBM enhanced color monitor was placed on a swivel table so that it could be turned toward the subject before each trial and away from the subject while the experimenter recorded data from each trial. The joystick was mounted to the right arm of the subject's seat and the keypad was placed on a desk to the left of the subject.

2.1.3. Tracking Task. Subjects were seated approximately 30 inches from the CRT screen. For the tracking task, subjects were required to keep the cursor centered on a stationary cross-hair in the center of the CRT

screen by applying left and right movements to the joystick with the right hand. A disturbance input consisting of a sum of five sine waves was used to keep the cursor in motion. The control dynamics could be adjusted for either first or second order tracking. Each trial lasted two minutes.

Subjects used two strategies for performing the tracking task: continuous strategy and the double-impulse strategy. For the continuous strategy, subjects were instructed to apply continuous movements to the joystick to minimize errors. Smooth control stick movements and linear system output are characteristic of this strategy. In contrast, for the double-impulse strategy, subjects were instructed to minimize errors as quickly as possible by applying maximum velocity or acceleration to the system. This strategy required the subjects to make large stick deflections. Subjects were instructed to compensate for overshooting the target by counteracting each stick deflection with a movement in the opposite direction.

2.1.4. Sternberg Task. A memory set of two, four or six consonants was presented to the subject at the beginning of each 2-minute trial. The memory set appeared on one line located 5 inches from the top of the CRT screen.

Subjects were given 10 seconds to memorize the letters, at which point the memory set disappeared. During the trial, positive and negative probes appeared in a box

located 1-1/2 inches above the stationary cross-hair used for the tracking task. Each probe was presented for 10 seconds or until the subject responded to the probe. Subjects responded "yes" if the probe was in the memory set (positive probe) and "no" if the probe was not in the memory set (negative probe). If the subject responded to the probe within the 10-second period, the response to the probe was recorded and the probe disappeared. If the subject did not respond within 10 seconds, an incorrect response was recorded and the probe disappeared. The time interval from one response to the presentation of the next probe varied from two to six seconds.

Subjects were required to use either a simple response to the probes or a difficult response. For the simple response, subjects pressed the left key of the response box to respond "yes" and the right key of the response box to respond "no." Key presses were made with the left index finger. For the more difficult response, subjects were required to press a sequence of three keys (left-top-bottom = yes; right-bottom-top = no). Subjects were told to use the left index finger for key presses, not to pause in between key presses and not to press keys simultaneously. An incorrect response was recorded if the subject pressed an incorrect sequence of keys or if the subject failed to complete a sequence of key presses. Before each trial, the experimenter specified the memory set size and the type of

response required. The responses and response times were recorded by the program as well as response accuracies.

2.1.5. Dual Tasks. When both tracking and the Sternberg tasks were performed, the subjects were given 10 seconds to memorize the memory set. After 10 seconds, the memory set disappeared and the cursor started moving. Subjects performed tracking with their right hand and made responses to the memory probes with their left hand. Dual task trials lasted two minutes each.

2.1.6. Procedure. Subjects participated in four one-hour practice sessions and four one-hour experimental sessions. During the first practice session, subjects were introduced to the Sternberg task and were allowed to practice with memory set sizes of two, four and six letters. Subjects practiced both the simple one-key response and the more difficult three-key response with the three memory set sizes. After practice with the Sternberg task, the experimenter described first and second order tracking and allowed the subjects to practice tracking with both levels of difficulty.

The second practice session was devoted to learning the two strategies. Subjects were given an illustrated description for each strategy. As subjects practiced each strategy, the experimenter gave subjects feedback concerning the velocity of the control stick, a crude indicator of strategy. Lower velocities are characteristic of the

continuous strategy and higher velocities are characteristic of the double-impulse strategy. The experimenter reminded subjects the purpose of each strategy if subjects seemed to stray from a particular strategy. In addition, the experimenter monitored subjects' performance to make sure they were maintaining the same level of performance with both strategies.

The third and fourth practice sessions allowed subjects to practice both the Sternberg and tracking tasks together. Subjects practiced each of the twenty-four dual task combinations at least once. These combinations were formed from the orthogonal combination of two levels of strategy, two levels of tracking difficulty, two levels of response load and three levels of central-processing load. Once subjects had practiced each combination at least once, they concentrated on improving performance on dual tasks that gave them the most difficulty. Subjects were encouraged to reach a consistent level of performance for each of the combinations of dual tasks.

For each experimental session, the subjects used either the continuous or double-impulse strategy, alternating strategies each session. At the beginning of each session, subjects were told which strategy to use and were reminded of the purpose of the strategy. Each experimental session consisted of twelve dual task trials (two levels of tracking difficulty, two levels of response load and three levels of

central-processing load). After the twelve dual task trials, subjects engaged in seven single task trials. Before each of the two-minute trials, the experimenter stated the level of tracking difficulty to be used for tracking, the number of letters to be memorized for the Sternberg task, and the type of response to be made for the Sternberg task. The experimenter then turned the CRT toward the subject, and initiated the trial. When the trial terminated, the experimenter turned the CRT away from the subject and recorded the data from that trial. This process was repeated for all nineteen trials.

2.1.7. Design. The four independent variables (i.e. strategy, tracking difficulty, response load and central processing load) were manipulated within subjects. Strategy was manipulated across sessions with half of the males and half of the females assigned to the continuous strategy for sessions 1 and 3 and the double-impulse strategy for sessions 2 and 4. The other half of the subjects received the reverse order. The twelve possible combinations of tracking difficulty, response load and central processing load were enumerated then randomly arranged into a sequence for each subject. This random sequence was assigned to sessions 1 and 4 and the reverse random sequence was assigned to sessions 2 and 3. This method of assignment assured that the random sequence and the reverse random sequence were not confounded with strategy. When subjects

performed the single tasks, they first performed the lowest level of each of the three variables then progressed to the higher levels. However, the three variables themselves were enumerated and randomly arranged for each subject. This random order was assigned to sessions 1 and 4 and the reverse random order was assigned to sessions 2 and 3.

2.2. Results

Two measures were used to assess tracking performance: Root Mean Square Error and Root Mean Square Velocity. Performance on the Sternberg task was assessed by percentage of correct responses, referred to as Accuracy, and Response Time expressed in milliseconds. For each dependent variable, a four-way strategy x Tracking difficulty x Response Load x Central Processing Load analysis of variance was performed using repeated measures analysis. The results are discussed in terms of each independent variable.

2.2.1. Cognitive Processing Load. The main effect of Cognitive Processing Load (i.e. memory set size) was significant for Accuracy ($F(2,10)=7.21$, $p<.05$) and Response time ($F(2,10)=13.54$, $p<.01$). Accuracy dropped as number of letters in the memory set increased and response time increased. With 2 letters, subjects responded correctly 95.9 percent of the time (mean RT= 720.82), with 4 letters, subjects responded correctly 94.6 percent of the time (Mean RT= 793.92), and with 6 letters, subjects responded correctly 92.3 percent of the time (Mean RT= 868.9). These

results indicate that performance on the Sternberg task declined as memory set size increased. However, cognitive load did not appear to have a main effect on tracking.

2.2.2. Response Load. The Response Load showed a reliable main effect on Accuracy ($F(1,11)=7.77$, $p<.05$), Response Time ($F(1,11)=5.02$, $p<.05$), and RMS Tracking Error ($F(1,11)= 12.2$, $p<.01$). Accuracy dropped from 96% correct to 92.5% correct and response times increased from 777 msecs to 812 msecs when the difficult response was employed. In addition, tracking performance declined when the more difficult Sternberg response was used (mean RMS error = .197 and .208 for simple and difficult responses respectively). Thus response load appeared to effect both the Sternberg and the tracking task.

2.2.3. Tracking difficulty. Tracking Difficulty had a significant effect on RMS tracking error ($F(1,11)=162.23$, $p<.01$) with subjects performing better in first-order tracking (Mean = .112) than in second-order tracking (mean = .293). Tracking Difficulty also had a significant effect on RMS tracking velocity ($F(1,11)=116.52$, $p<.01$). Subjects showed lower RMS velocity with first-order tracking tasks (Mean = .166) than with second-order tracking tasks (Mean = .433). These results are as expected since second order tracking is more difficult to perform.

Tracking Difficulty also showed a marginally reliable interaction with Cognitive Processing Load for response time

($F(2,10)=3.52$, $p<.07$) and RMS tracking error ($F(2,10)=4.96$, $p<.05$). Cognitive load showed a slightly larger effect on response time for second order tracking (mean RTs of 723 msecs and 885 msecs for memory sets of 2 and 6 respectively) than for first order tracking (mean RTs of 719 and 853 for memory sets of 2 and 6 respectively). Similarly, Cognitive load had a very smaller effect on first order tracking (.112 vs .113 for memory sets of 2 and 6 respectively), and a larger effect on second order tracking (.292 vs .298 for memory sets of 2 and 6 respectively). Surprisingly second order tracking error was lowest when a memory set of 4 was used (mean RMS error = .289).

Finally, Tracking Difficulty showed marginally reliable interactions with Response Load for RMS tracking error ($F(1,11)=4.53$, $p<.06$) and response time ($F(1,11)=5.06$, $p<.05$). For first order tracking, increased response load primarily influenced the Sternberg task. Response times increased from 763 msecs to 816 msecs compared to an increase from 790 to 808 for second order tracking. For second order tracking, increased response load was primarily observed in tracking performance. Tracking error scores for second order tracking increased from .284 to .301 as response load increased compared to an increase from .110 to .115 for first order tracking. It is not clear from the present results whether this difference reflects a qualitative difference in first versus second crder tracking

or whether it reflects a tendency for response load to effect the most difficult of two tasks. This latter argument assumes that the Sternberg task is more difficult than first order tracking but less difficult than second order tracking.

2.2.4. Strategy. Strategy showed a reliable effect on RMS velocity ($F(1,11)=216.95$, $p<.01$). Mean RMS velocities were much higher when subjects used the double-impulse strategy (Mean = .439) than the continuous strategy (Mean = .160). However, there was not a main effect of strategy for RMS tracking error. Since higher control velocities are expected with the double-impulse strategy, these results suggests that subjects were using the proper strategies and that they could perform the tracking task equally well with either strategy.

The question of whether or not the strategies differ in processing demands is addressed by an interaction between Strategy and cognitive processing load. The Strategy x Cognitive Processing Load was marginally significant for Response Time ($F(2,10)=3.8$, $p<.06$). The effect of cognitive processing load was slightly larger for the continuous strategy than for the double-impulse strategy. Response times when the continuous strategy was used varied from 712 msecs to 870 for memory sets of 2 and 6. For the double-impulse strategy these means were 730 and 868 respectively.

Additionally, reliable Strategy x Response load

($F(1,11)=12.43$, $p<.01$) and Strategy x Tracking Difficulty ($F(1,11)=49.42$, $p<.01$) interactions were obtained for RMS control velocity. For the continuous strategy control velocity increased as response load increased but for the double-impulse strategy control velocity decreased as response load increased. The interaction between strategy and tracking difficulty suggests that the effects of second order dynamics have a larger effect on control velocity when subjects use the double-impulse strategy.

2.3. Discussion Experiment 1

In this experiment two variables are used to manipulate central processing load: Sternberg memory set size, and tracking task dynamics. Both of these variables primarily effect the tasks to which they apply. Increasing memory set size results in longer response times and higher error rates but not an overall change in tracking performance. Increasing tracking dynamics results in an increase in tracking error but no change in Sternberg task performance.

These results suggest little or no resource competition between the two manipulations. On one hand, this result is expected since the Sternberg task and the tracking task utilize different codes of processing: verbal versus spatial according to Wickens' multiple resources model. On the other hand, some competition would seem to be predicted since both manipulations increase demands for central processing resources rather than perceptual or response

related resources. Evidence for some resource competition was provided by the interaction between cognitive load (i.e., memory set size) and tracking difficulty. This interaction is obtained for both response time and tracking error and for both variables it suggests that the effects of memory set are larger with second order tracking than with first order tracking. The fact that the interaction is obtained for both task measures suggests that there is some resource competition between manipulations. However, performance declines only when resource demand is high.

The other task variable, response load, effects performance on both tasks. As the response to the Sternberg task becomes more difficult, response times become longer and less accurate and tracking performance declines. This pattern of results in indicative of resource competition. As the response demands of the Sternberg task are increased the competition for response related resources increases and performance on both tasks decline. Thus, the results replicate previous work supporting Wickens' Multiple resources model.

The main variable of interest, however, is Strategy. The objective of the experiment was to further assess processing demands of strategies. The results indicated that subjects did apparently use the appropriate strategies, and were able to track equally well with either one. The only result relevant to the assessment of processing demands

is the interaction between strategy and cognitive processing load on response time. The effect of central processing load appeared to be slightly larger for the continuous strategy than for the double-impulse strategy. This interpretation is consistent with previous work on strategies (Goettl, 1988) and suggests that the continuous strategy demands more central processing resources than the double-impulse strategy.

The two strategies do not appear to differ in response related resource demands. Although increasing the difficulty of the Sternberg task disrupted tracking performance, it did not interact with tracking strategy. Thus, the effect of response difficulty was equivalent for tracking with either strategy. This experiment represents the first empirical test of the hypothesis that strategies differ in response resource demands. Since it is based on the failure to find a significant interaction, caution must be exercised in making conclusions about strategy demands. Replication of the findings is needed.

The other objective of this project was to evaluate the optimum-maximum procedure described by Navon (1985). Goettl (1988) argued that the procedure may impose demand characteristics on subjects. Performance criteria demands are imposed on subjects via on-line feedback and instructions to maintain performance at certain levels. Experiment 2 investigated this hypothesis by employing the

optimum-maximum procedure and varying the amount of on-line feedback. If performance changes as a result of amount of feedback, then it may be concluded that on-line feedback may impose demand characteristics.

3. EXPERIMENT 2: EFFECTS OF FEEDBACK

3.1. Methods

3.1.1. Subjects. Twenty-one Clemson University students enrolled in an introductory psychology class served as the subjects for the experiment. The average age of the subjects was 19.1 years, the oldest being 21 and the youngest being 18. There were twelve male and twelve female subjects, and all but two of the subjects were right-handed. All of the subjects had normal or corrected to normal vision. The subjects were paid $4.50 per hour for their participation.

3.1.2. Apparatus. All apparatus and equipment employed was the same as that used in Experiment 1 with the exception of the display monitor. A Zenith Data Systems ZCM-1490, 14-inch color monitor was used rather than an IBM enhanced color monitor. A second monitor was positioned in front of the experimenter to observe the subjects' performance.

3.1.3. Tasks. The experimental tasks were the same as those used in Experiment 1 with the exception that only the second order tracking task was used, and only a memory set size of 4 was used for the Sternberg task. In addition, subjects were not instructed to use any particular tracking

strategy. The only other difference was that the present experiment utilized on-line feedback.

On-line feedback was provided during the tasks by the color of the objects on the screen. The color of the vertical line and stationary cross provided feedback of the tracking task, and memory task performance feedback was displayed by the colors of the small box and memory probes. The feedback for each task was evaluated and changed (if necessary) every five seconds during the trial. Five colors were used to represent 5 different levels of performance: red for poor performance, yellow for below average, white for average, light blue for above average, and dark blue for excellent performance.

3.1.4. Procedure. Each session consisted of approximately sixteen trials, each trial being two minutes long. Results were only recorded during the experimental sessions. After each trial, the subjects were given general information concerning his or her performance according to the computer's summary of the trial. The subjects were required to wait twenty-four hours between sessions, and all six sessions were completed within a two week period.

During the first practice session, subjects were given instructions on the two tasks and allowed to practice each several times. If the subjects demonstrated reasonable mastery of both tasks, they were allowed to practice both tasks concurrently, otherwise they continued to practice the

tasks separately.

The second practice session consisted mainly of more practice of both tasks at the same time, after the subjects were allowed a few "warm-up" sessions at the beginning of the session. Any questions concerning the execution of the task were answered at this time. All the subjects were randomly observed to ensure that the memory task was being answered with only one finger, and they were corrected if necessary.

During the third and last practice session the subjects were introduced to the Optimum-Maximum Procedure (Navon, 1984). In this procedure, subjects are given a specified of level of performance on one task, the optimized task, and are asked to perform the other task, the maximized task, as well as possible. According to Navon, this procedure eliminates possible demand characteristics on the tasks and will lead to more accurate results.

3.1.5. Design. The subjects were randomly assigned to one of three groups of seven, each with three males and four females. Each subject attended six one-hour sessions, the first three being practice and the last three being actual experimental sessions. The amount of feedback subjects received was manipulated. Group 1 received feedback for both tasks. Group 2 received feedback only for the optimized task. Group 3 received no feedback during the practice sessions or the experimental sessions.

The optimized task was required to be performed at one of three levels during the optimum-maximum procedure, either at above average (single task level or STL), average, or below average (mean plus one standard deviation). The two levels of response difficulty were factorially combined with the three optimization levels of each task resulting in 12 different dual-task conditions which were randomly ordered during the experimental sessions. The three Single tasks were performed at the end of each session.

Performance of the tracking task, or tracking error, was measured as root mean square error (RMS). Also, both the RMS tracking velocity and RMS tracking position were recorded by the computer. Performance of the Sternberg memory task was measured in both accuracy and reaction time.

3.2. Results

Root mean square tracking error and Response times were submitted to separate 3 (Feedback: full, partial, and none) x 2 (Optimized Task: Sternberg vs tracking) x 2 (Response Difficulty) x 3 (Level of optimization: single task level, mean dual-task level, and worse than mean dual task) ANOVAs using a mixed design. Feedback was manipulated between-groups, all other variables were manipulated within-subjects.

3.2.1. Tracking performance. The analysis performed on RMS tracking error revealed main effects of Optimized task ($F(1,18)=16.18$, $p<.01$), Response load ($F(1,18)=38.65$,

p<.01), and Level of optimization (F(2,36)=39.31, p<.01). In addition, two reliable interactions were obtained: Optimized task x level of optimization (F(2,36)=30.35, p<.01), and Optimized task x response load (F(1,18)=6.36, p<.05). The main effect of Optimized task indicated that overall, subjects tracked better when the tracking task was maximized (mean RMS error = .308) than when it was optimized (mean RMS error = .337). The main effect of response load replicated findings of Experiment 1. Subjects showed better tracking when the Sternberg task required the easy response (mean RMS errors of .308 and .337 for easy and difficult response respectively). Moreover, the optimized task x response load interaction suggested that the effect of response load was larger when the Sternberg task was optimized (.289 vs .327 for easy and difficult responses respectively) than when tracking was optimized (.328 vs .347).

Table 1 shows mean RMS tracking errors and Mean response times at each optimization level for each task. The values on the left are averaged across the three groups, data on the right are separated by group. The main effect of Level of optimization showed that overall subjects performed better when they optimized either task at single task levels. However, caution must be used in interpreting this effect since Level of optimization interacted with Optimized task. Close examination of the RMS means in Table

Table 1. RMS Tracking Error and Response Time Measures for Optimized and Maximized tasks averaged across groups (top) and separated by groups (bottom).

Overall Data	RMS Error	RT (in msecs)
Tracking Optimized		
at single task level	.288	697
at mean of dual-task	.342	684
below mean of dual-task	.382	710
Sternberg Optimized		
at single task level	.312	635
at mean of dual-task	.306	699
below mean of dual-task	.306	751
FULL FEEDBACK		
Tracking Optimized		
at single task level	.307	762
at mean of dual-task	.361	742
below mean of dual-task	.399	787
Sternberg Optimized		
at single task level	.333	679
at mean of dual-task	.328	763
below mean of dual-task	.332	808
OPTIMIZED TASK FEEDBACK ONLY		
Tracking Optimized		
at single task level	.287	650
at mean of dual-task	.365	662
below mean of dual-task	.411	687
Sternberg Optimized		
at single task level	.306	614
at mean of dual-task	.297	676
below mean of dual-task	.301	728

Table 1 (continued).

NO FEEDBACK		
Tracking Optimized		
at single task level	.269	678
at mean of dual-task	.299	649
below mean of dual-task	.337	655
Sternberg Optimized		
at single task level	.297	613
at mean of dual-task	.293	659
below mean of dual-task	.286	716

1 reveals the nature of the interaction. Tracking performance significantly improved as the optimization level of the tracking task became more challenging, however, tracking performance declined slightly as optimization level of the Sternberg task became more challenging. The simple main effect of level was reliable when the tracking task was optimized ($F(2,36)=40.26$, $p<.01$), but not when the Sternberg Task was optimized ($F(2,36)=1.03$, n.s.).

There was no main effect of feedback. The only effect that feedback showed was a marginally reliable interaction with Optimized task ($F(2,18)=3.12$, $p<.07$). Although no statistically reliable group effects were found, the group means suggested that the groups did differ slightly (See Table 1). The most interesting difference was in the Level of optimization x Task optimized interaction. The no-feedback group seemed to show larger differences in optimization level when the Sternberg task was optimized. It is possible that the failure to find group differences was due to lack of power in the design. There were only 7 subjects in each group.

3.2.2. Sternberg Task Performance. Analysis of response times revealed main effects of Response load ($F(1,18)=6.12$), $p<.05$) and Level of optimization ($F(2,34)=13.46$, $p<.01$). In addition a reliable Optimized task x Level of optimization interaction was obtained ($F(2,34)=20.43$, $p<.01$). The Response load effect showed

that subjects responded more quickly when an easy response (mean RT 687 msecs) was made than when a difficult response was made (mean RT 705 msecs). This effect replicates the response load effect obtained for tracking performance and the effects obtained in Experiment 1.

The effects of Level of optimization and the Optimized task x Level of optimization interaction also replicate the results obtained for tracking performance. Overall, subjects responded more quickly when they optimized at more challenging levels. But the interaction suggests that the effect of level of optimization was larger when the Sternberg task was optimized than when the tracking task was optimized. Analysis of simple main effects confirmed this observation. The effect of Level of optimization was reliable when the Sternberg task was optimized ($F(2,34)=21.45$, $p<.01$) but not when the Tracking task was optimized ($F(2,34)=2.62$, $p>.05$).

As with tracking performance, even though the group means showed some differences (See Table 1), no reliable group effects were obtained. Again it is suggested that lack of power in the design can account for the failure to find the expected group differences.

3.3. Discussion Experiment 2

The results of Experiment 2 serve as a partial replication of Experiment 1. Response load of the Sternberg memory task effected Sternberg response latencies as well as

concurrent tracking performance. This result provides support to the conclusion that the increased response load increases competition for response related resources and supports a multiple resources model of attention.

However, the primary objective of Experiment 2 was not to support a multiple resources model per se, but to investigate the optimum-maximum procedure developed by Navon (1985). In this procedure, subjects are instructed to invest some proportion of their effort or resources to one task while performing a concurrent task as well as possible. On-line feedback is typically provided to allow subjects to monitor performance. If the tasks compete for resources then performance on the maximized task will decline as more resources are invested in the other task. Using this procedure, very little evidence of resource competition was observed. Performance on either optimized task improved as additional resources were invested in the task. However, performance on the concurrent tasks did not change in response to changes in resource investment in the optimized task. This conclusion conflicts with the conclusions from both experiments that increased response load on the Sternberg task increased the resource competition and disrupted performance on both tasks.

This conflict in the two results could be resolved by suggesting that since the tasks demand different codes of processing, according to Wickens' (1980) model, a great deal

of resource competition should not be expected. This is not an entirely satisfactory argument. Experiment 1 showed some competition for central processing resources when second order tracking was employed. A second way to resolve the conflict is to argue that the optimum-maximum procedure imposes its own demand characteristics on subjects. Subjects may adjust their performance rather than their resource investments to conform with instructions. The presence of on-line feedback would make performance adjustments easy since subjects always know what their performance is and what it should be. Subjects may artificially inflate their performance (without changing resource investments) to conform with instructions.

The argument that on-line feedback imposes demand characteristics was investigated in Experiment 2. Different groups received different amounts of feedback. One group received feedback on both tasks as in the original version of the optimum-maximum procedure. A second group received feedback only on the optimized task. The third group did not receive any on-line feedback. If on-line feedback does impose demands on subjects, then the full feedback group should show little variability on the maximized task and the no-feedback group should show larger changes in the maximized task. A decrease in maximized task performance resulting from increased investment of resources in the maximized task is consistent with a multiple resources

interpretation. Though group means hinted at this trend for both dependent measures. No reliable group effects were obtained. However, since there were only seven subjects assigned to each group the failure to find group differences may be do to a lack of power in the design.

4. GENERAL DISCUSSION AND CONCLUSIONS

There were two objectives of this project: to further investigate resource demands of tracking strategies, and evaluate the optimum-maximum procedure. The first objective was addressed in Experiment 1. Subjects performed a dual-task in which a Sternberg memory task was performed concurrently with a tracking task. Subjects used each of two different strategies for tracking. Central processing demands of both tasks were manipulated along with the response demands of the Sternberg task. Varying the response demands disrupted performance on both tasks (suggesting resource competition between tasks) but effected both strategies equally. Central processing demands showed only minor effects on both tasks, suggesting little resource competition, but effected the continuous strategy slightly more than the double-impulse strategy. To the extent that this result is reliable, it supports earlier research (Goettl, 1988).

The results concerning response load suggest that both tasks place equal demands on response related resources. Since this was the first empirical test of response demands,

the result does not conflict with previous data. However, it does counter suggestions made by Wickens and Goettl (1985). Wickens and Goettl hypothesized that the continuous strategy placed large demands on resources at some point other than perceptual resources and suggested that response related resources as the likely candidate. However, the findings of Goettl (1988) combined with the results of Experiment 1 suggest that the locus of the additional demands is central processing. To summarize the effects, it appears as though the double-impulse strategy places larger demands on perceptual processing than the continuous strategy which seems to place greater demands on central processing. Both strategies apparently draw equally on response related resources.

The second objective, to evaluate the optimum-maximum procedure was addressed by Experiment 2. It was postulated that on-line feedback used as part of the optimum-maximum procedure imposed demands characteristics on subjects. By varying the amount of feedback between groups it was hoped that the demand characteristics would be revealed. The group means suggested that there were important, though slight, group differences. Unfortunately, no reliable group differences were obtained. It was suggested that lack of statistical power may have resulted in the failure to obtain group effects.

5. References

Goettl, B. P. (1988). Cognitive Demands of Tracking Strategies as Assessed by the Optimum-Maximum Procedure. USAF-UES Summer Faculty Research Program Final Report.

Hess, R. A. (1978). Dual-loop model of the Human Controller. Journal of Guidance and Control, 1, 254-260.

Hess, R.A. (1979). A rationale for human operator pulsive control behavior. Journal of Guidance and Control, 2, 221-227.

Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, NJ:Prentice-Hall

Navon, D. (1984). Resources--A theoretical soup stone? Psychological Review, 91, 216-234.

Navon, D. (1985). Do people allocate limited resources among concurrent activities? (IPDM Report No. 26). Haifa, Isreal: University of Haifa Laboratory for Information Processing and Decision Making.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource limited processes. Cognitive Psychology, 7, 44-64.

Smith, O. J. M. (1962). Nonlinear computations in the human controller. IRE Transactions on Bio-Medical Electronics, BME-9, 125-128.

Wickens, C.D. (1980). The structure of attentional resources. In R. Nickerson & R. Pew (Eds.), Attention and

Performance VIII (pp. 239-258). Hillsdale, NJ: Erlbaum Associates.

Wickens, C.D. (1984). Processing resources in attention. In R. Parasuraman & D.R. Davies (Eds.), Varieties of Attention (pp. 63-102). New York: Academic Press.

Wickens, C. D., & Goettl, B. (1985). The effects of strategy on the resource demands of second order manual control. In R. E. Eberts & C.G. Eberts (Eds.), Trends in Ergonomics/Human Factors II (pp.235-242). Amsterdam: Elsevier Science Publishers.

Wickens, C. D., Webb, J., & Fracker, L. (1987, January). Cross modality interference: A resource, preemption, or switching phenomenon? (Technical Report EPL-87-1/NASA-87-1). Champaign, IL: University of Illinois Engineering Psychology Research Laboratory.

Wilde, R. W., & Wescott, J. H. (1962). The characteristics of the human operator engaged in a tracking task. Automatica, 1, 5-19.

Young, L.R., & Miery, J.L. (1965). Bang-bang aspects of manual control in higher-order systems. IEEE Transactions on Automatic Control, 10, 336-341.

FINAL REPORT NUMBER 105
REPORT NOT AVAILABLE AT THIS TIME
Dr. David G. Payne
210-9MG-121

FINAL REPORT NUMBER 106
EFFECT OF SYSTEM RELIABILITY ON
PROBABILISTIC INFERENCE
PENDING APPROVAL
Dr. Donald Robertson
760-7MG-094

OPTIMIZATION OF THE NONLINEAR DISCRETE PARAMETER MODEL OF THE SEATED HUMAN SPINE

by

Joseph E. Saliba

and

Mary Facciolo

Civil Engineering and Engineering

Mechanics Department

University of Dayton

to

Universal Energy Systems

4401 Dayton-Xenia Road

Dayton, Ohio 45432

December 1989

ACKNOWLEDGMENTS

We would like respectfully to thank the Air Force Systems Command, the Air Force Office of Scientific Research, and Universal Energy Systems for providing us with the opportunity to work on this project. We would like to acknowledge the Aerospace Medical Research Laboratory, and in particular, the Biomechanical Protection Branch for their collaboration and help in obtaining some of the data.

We would also like to thank Mr. James W. Brinkley for his collaboration and guidance during this project. Finally, we would like to thank Mrs. Jo Ann Riner for helping with this report.

ABSTRACT

Section I of this report reviews the use of both continuous and lumped-parameter models to describe the structural response of the human body due to the acceleration environment associated with seat ejection. The need for a nonlinear lumped-parameter model is established based on the inadequacies of linear models to reproduce results in laboratory experiments. In Section II this nonlinear lumped-parameter model is then closely examined using four different tasks. The first task is to insure that the tests conducted in the laboratory are yielding meaningful output, otherwise the model is useless. Second, a nonlinear lumped-parameter model which best predicts the behavior of the human subject is developed. Third, the model is fine-tuned through experimental studies. Fourth, the model is validated through laboratory experiments.

Section III fully describes the program SEAT. This program was written to simulate the results of the model, as well as to determine the errors generated in the model.

In Section IV a more comprehensive validation plan, encompassing considerations of sensitivity to shape, duration, and magnitude of the acceleration input, is recommended to ensure the effectiveness of this approach.

UNIVERSITY OF DAYTON
CIVIL ENGINEERING AND
ENGINEERING MECHANICS
DAYTON, OHIO 45469

I. INTRODUCTION

Investigations of aircraft escape systems, and spinal injuries sustained from seat ejections, have involved a wide spectrum of analytical and experimental studies. The analytical efforts have concentrated primarily on the development, dynamic response solution, and validation of several discrete parameter and continuum models subjected to specified acceleration-time inputs. The experimental work has entailed measurement of constitutive and inertial properties of anatomical components and determinations and comparisons of biomechanical responses of human volunteers, cadavers, animals, and dummies to controlled impacts. Injury criteria and human tolerance thresholds formulated from these studies and accident statistics provide a basis for design and analysis of escape systems.

In describing the structural response of the human body due to the acceleration environment associated with seat ejection, one can represent the physical laws that are applicable in terms of rather complex equations. Obviously, the most popular simplification is the linear discretization, which leads to simple equations for the problem under ideal conditions. This linearization is, in many instances, inadequate in describing or reproducing experimental studies performed under very controlled environments.

Theoretical models, both continuous and lumped-parameter, with varying degrees of complexity, have been formulated in efforts to predict responses to specified inputs. The lumped-parameter models usually have been single or multidegree-of-freedom systems with linear, passive elements, whereas the human body is neither a passive nor a linear system. Carmichael (1,1968) states, "the body is a nonlinear system with respect to the equations of

motion." Wittmann and Phillips (2,1969) conclude, "existing nonlinearities are sufficient to be of concern, and that appropriate measures to determine the effect of these nonlinearities should be taken." Perng (3,1970) states, "the most striking feature of the mechanical behavior behavior of biological tissues, the stress strain relation, is remarkably nonlinear." Markoff and Steidel (4,1970) state, "it is known that joint stiffness increases with deformation." Belytschko, et al. (5,1972) refers to, but neglects nonlinearities, "because of the difficulties of characterizing a nonlinear orthotropic composite material." The conclusion being that when representing nonlinear elements of the body, the assumption of linear springs and dashpots may be too great a simplification for some inputs.

II. OBJECTIVES

To remedy this inadequacy of reproducing laboratory experiments, a nonlinear discrete parameter model will be examined. To accomplish this, four different tasks need to be investigated:

-Checking of the experimental data

-Selecting a model

-Fine-tuning the model

-Validating the model

The first task is to insure that the test conducted in the laboratory on the volunteer subjects is yielding meaningful output, otherwise the reproduction, or calibration, of the model is meaningless. The second task is to find a possible model that best predicts the behavior of the human subject during seat ejection. The third step involves investigating experimental means of fine-tuning the model by proposing additional experimental studies that best determine the mechanical properties which enter the constitutive relations of the nonlinear model. The final step is to calibrate, or

validate, the model by simply attempting to reproduce as many tests on volunteer subjects as possible, making sure that the analytical output matches the experimental output.

A. CHECKING OF THE EXPERIMENTAL DATA

The first task was to become familiar with the test plan and the results of the "vertical impact of humans and anthropomorphic manikins" test conducted at the Biomechanical Protection Branch at Wright-Patterson Air Force Base. This entailed familiarization with the test procedure, fixtures used, and checking the locations of load cells, accelerometers, and other devices. Next was a careful study of the equilibrium equations, which meant checking all of the force and moment equations. This was done to insure that the tests conducted in the laboratory on the volunteer subjects yielded meaningful output. The results of these tests will be used later in the calibration of a nonlinear model which will be developed.

B. SELECTING A MODEL

There seems to be an orderly progression towards the development of realistic spinal models. Two parallel approaches are discrete lumped-parameter and continuum models.

The first spinal lumped-parameter model was proposed by Latham (1957), which consisted of a double-mass spring-coupled system subjected to a base excitation. In 1962, Payne proposed a dynamic response index (DRI) in his single degree-of-freedom model which is still in use today. Many more models have been proposed since then, extending the degrees-of-freedom to include parts of the human body other than the spine.

Different types of spinal models have been proposed with less success. These models, initiated by Hess and Lombard (1958), assume the spine to be a continuum. The spine is represented by a straight homogenous elastic rod, free at one end and subjected to a prescribed acceleration at the other end.

Non-linear models have also been attempted, again with little success.

It has been demonstrated that the lumped-parameter models represent a more viable approach because a more realistic injury model can be deduced from them. Thus, a discrete model will be used.

C. REFINEMENT OF THE MODEL

Figure 1 represents our attempt to model the spine.

The selection of the characteristics of the lumped parameter of this model can now be studied by trying first, for example, a linear spring and a linear dashpot. If such a selection is not adequate to model the test that we are trying to duplicate, then a nonlinear spring and possibly a linear or nonlinear dashpot can then be evaluated. The process can be repeated by varying the nonlinearity of both spring and dashpot until satisfied. In order to better understand this process, let us determine the governing equations for a cubic spring and dashpot.

If we call $U(t) = Z_2(t) - Z_1(t)$ (Eq. 1)

then $U'(t) = (Z'_2 - Z'_1(t))$ (Eq. 2)

where the $U' = \frac{dU}{dt}$ means the first derivative with respect to time

and $U'(t) = (Z_2"(t) - Z_1"(t))$ (Eq. 3)

From the free-body-diagram of Figure 1, Newton's second law gives the following equations of motion in the form:

$$M\,(U"(t)) + F_{(spring)} + F_{(dashpot)} = F(t) \quad \text{(Eq. 4)}$$

where $F_{spring} = -K_1\,(U(t)) + K_3\,(U(t))^3$ (Eq. 5)

$$F_{dashpot} = C_1\,(U'(t)) + C_3\,(U'(t))^3 \quad \text{(Eq. 6)}$$

so $M\,U"(t) + K_1U(t) + K_3\,[U(t)]^3 + C_1\,[U'(t)] + C_3\,[U'(t)]^3 = F(t)$ (Eq. 7)

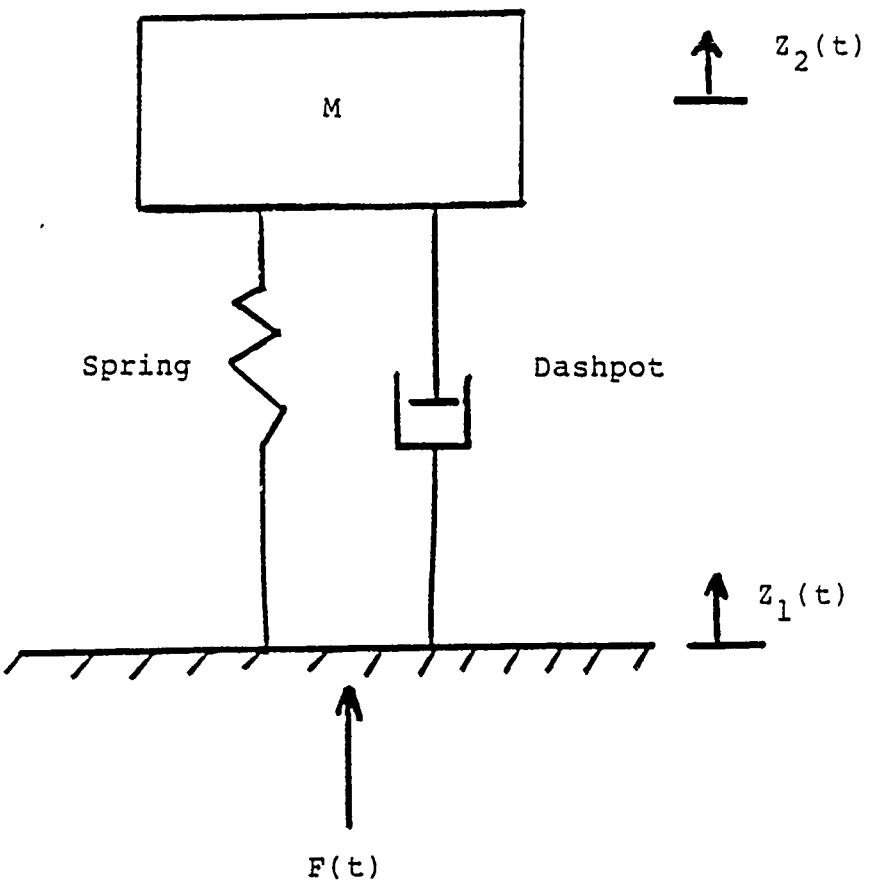


Figure 1

DISCRETE PARAMETER MODEL

Since we know $U'' = \frac{d}{dt}\frac{dU}{dt}$ we can then integrate in function of time

to get $\int_{t_1}^{t_2} U(t)''\,dt = \frac{dI(t)}{dt}\Big|_{t_1}^{t_2} = U'(t)\Big|_{t_1}^{t_2} = U'(t_2) - U'(t_1)$ (Eq. 8)

$$\text{or } U'(t_2) = U'(t_1) + \int_{t_1}^{t_2} U''(t)\,dt \qquad \text{(Eq. 9)}$$

$$\text{but } U'(t) = \frac{du}{dt} \Rightarrow U'(t)\,dt = du \qquad \text{(Eq. 10)}$$

integrating in function of time both sides, we then get

$$U(t_2) = U(t_1) + \int_{t_1}^{t_2} U'(t)\,dt \qquad \text{(Eq. 11)}$$

the integral in the above equations can be simply estimated by any numerical quadrature formula such as the Trapezoidal rule

$$\int_a^b f(x)\,dx \approx S = \frac{1}{2} * (b-a)*[f(a) + f(b)] + \text{error} \qquad \text{(Eq. 12)}$$

or by another method like Simpson's rule.

The error term for the trapezoidal rule is

$$\text{error} = -\frac{f''(\eta)\,(b-a)^3}{12} \quad \eta \in (a,b) \qquad \text{(Eq. 13)}$$

It is difficult to estimate the error term, especially since our value, for example, of the acceleration is given only in tabular form. In other words, we do not know what the actual equation for the acceleration is, hence, we cannot compute the second derivative of the acceleration to approximate the error. To remedy this, one can use the idea of the extrapolation to the limit in conjunction with the composite Trapezoidal rule. Such a method is commonly

referred to as the Romberg integration method. The best possible estimate of the integral can be found by simply looking at the table of ratios generated by this method.

Once the relative velocity and displacement are computed, the best values of the spring and dashpot constant can be found using the idea of the least squares method.

Since we do not know the values of K_1, K_3, C_1 and C_3 exactly, bor which satisfy the equality of equation (Eq. 7), an error will be introduced for every time value. The fundamental idea behind the least squares method, is to minimize the sum of the squares of these errors.

If we call

$$\phi_1(t) = U(t)$$

$$\phi_2(t) = (U(t))^3$$

$$\phi_3(t) = U'(t)$$

$$\phi_4(t) = (U''(t))^3$$

and call N the total number of point given and

$$f_m = F(t_m) - m\, U''(t_m)$$

where we can then write the square of the individual error as:

$$E(K_1,K_3,C_3) = \sum_{m=1}^{N} [f_m - G(t_m; K_1,K_3,C_1,C_3)]^2$$

where $G(t_m,K_1,K_3,C_1,C_3) = K_1\phi_1(t_m) + K_3\phi_2(t_m) + C_1\phi_3(t_m) + C_3\phi_4(t_m)$

To minimize the above error function, it is necessary thac the gradient of E vanish, i.e.;

$$\frac{\partial E}{\partial K_1} = \frac{\partial E}{\partial K_3} = \frac{\partial E}{\partial C_1} = \frac{\partial E}{\partial C_3} = 0$$

$$\frac{\partial E}{\partial K_1} = -2 \sum_{m=1}^{N} [f_m - G(t_m;K_1,K_3,C_1,C_3)]\ \phi_1(t_m) = 0$$

$$\frac{\partial E}{\partial K_3} = -2 \sum_{m=1}^{N} [f_m - G(t_m;K_1,K_3,C_1,C_3)]\ \phi_2(t_m) = 0$$

$$\frac{\partial E}{\partial C_1} = -2 \sum_{m=1}^{N} [f_m - G(t_m;K_1,K_3,C_1,C_3)]\ \phi_3(t_m) = 0$$

$$\frac{\partial E}{\partial C_3} = -2 \sum_{m=1}^{N} [f_m - G(t_m;K_1K_3,C_1,C_3)]\ \phi_4(t_m) = 0$$

the error vector

$$e = [e_1, e_2, \ldots e_N]$$

with

$$e_m = f_m - G(t_m;\ K_1,K_3,C_1,C_3)$$

are normal orthogonal to the following vector

$$\theta_i = [\phi_i(t_1),\ \phi_i(t_2) \quad \phi_i(t_N)]^T \quad i = 1,2,3,4$$

since

$$-2\ e^T\ \theta_i = 0 \quad i = 1,2,3,4$$

calling the vector

$$f = [f_1, f_2, \ldots f_N]^T$$

the normal equations can be rewritten in the form:

$$\sum_{j=1}^{4} b_j \, \theta^T \theta_i = f^T \theta_i \quad i = 1,2,3,4$$

where

$b_1 = K_1$

$b_2 = K_3$

$b_3 = C_1$

$b_4 = C_3$

Solving these normal equations will then yield the four constants K_1, K_3, C_1, C_3, which will minimize the square of the error introduced by approximating the behavior by this discrete lumped-parameter model.

D. VALIDATION OF THE MODEL

The model was validated using the VIHAM STUDY TEST: 1390 SUBJECT: B-1 weight 163.0 NOM G=8-0 CELL: X. The chest minus seat acceleration time profile is shown in Figure 2. Because of the deviation from zero, the above figure was shifted upward and then integrated to obtain the velocity, and integrated once again to obtain the displacement. Then the data was passed to the least-square program and the best possible fit was obtained. The force profile used was the sum of the three load cells of the seat. (see Figure 3).

III. THE SEAT PROGRAM

Program SEAT generally uses the method of least squares in combination with various subroutines to ultimately solve for the coefficients of the force equation (Eq. 7). The resulting force equation best describes the action on the model.

A. MASTER PROGRAM

By using the force analysis and theory developed in Section II-C it was possible to create a master program called SEAT. Program SEAT is a

Figure 2

CHEST-SEAT ACCELERATION VS. TIME
ACCELERATION (g)
6.0
4.0
2.0
0.0
-2.0
-4.0
-6.0
0
100
200
300
400
500
600
700

SEAT FORCE (Z) VS. TIME
FORCE (LB.)
2000.
1500.
1000.
500.
0.
0
100
200
300
400
500
600
700
TIME (MSEC.)

Figure 3

combination of individual subroutines. The subroutines use compatible parameters, and therefore may be used together to generate one program which can provide all of the desired output responses.

```
0001          PROGRAM SEAT
0002    ************************************************************************
0003    *                    DEFINITION OF VARIABLES
0004    *
0005    *     F()      -VECTOR CONTAINING FORCE-MASS*ACCELERATION
0006    *     ACCEL    - CHARACTER NAME OF THE ACCELERATION DATA
0007    *     FORCE    - CHARACTER NAME OF SEAT FORCE DATA FILE
0008    *     NP       - NUMBER OF POINTS
0009    *     NS       - NUMBER OF SEGMENTS ( NS=NP-1)
0010    *     NLP      - NUMBER OF LUMPED PARAMETER
0011    *     LPI      - LUMPED PARAMETER INDICATOR
0012    *                LPI=1 IF PARAMETER IS A DISPLACEMENT PARAMETER
0013    *                LPI=2 IF PARAMETER IS A VELOCITY PARAMETER
0014    *     LPP      - EXPONENT VALUE OF EITHER DISPLACEMENT OR VELOCITY
0015    *     RESPONSE- STANDARD OF DEVIATION OF THE FIT
0016    *     SUMERR   - SQUARE OF THE SUM OF COMPUTED F MINUS ACTUAL F
0017    *     TO       - INITIAL TIME VALUE
0018    *     TF       - FINAL TIME VALUE
0019    *     MASS     - MASS OF THE SUBJECT
0020    *     OUTFILE - CHARACTER NAME OF THE OUTPUT FILE
0021    *     STAT     - STATUS OF THE OUTPUT FILE
0022    *     COMFOR()- COMPUTED FORCE VECTOR OF ORDER NP
0023    *     ERROR() - COMPUTED MINUS ACTUAL FORCE VECTOR AND OF ORDER NP
0024    *     DIS()    - DISPLACEMENT VECTOR OF ORDER NP
0025    *     VEL()    - VELOCITY VECTOR OF ORDER NP
0026    *     ACC()    - ACCELERATION VECTOR OF ORDER NP
0027    *     FOR()    - FORCE VECTOR OF ORDER NP
0028    *     TIM()    - TIME VECTOR OF ORDER NP
0029    *
0030    *                K [NPL,NPL] * B (NPL) - P (NPL)
0031    *
0032    *   IS THE NORMAL EQUATION RESULTING FROM THE LEAST SQUARES METHOD
0033    *
0034    *     WHERE
0035    *
0036    *     K(,)     - MATRIX OF THE VALUES FOR THE INDEPENDENT VARIABLES
0037    *     B()      - VECTOR CONTAINING THE UNKNOWN COEFFICIENTS
0038    *     P()      - VECTOR CONTAINING THE DEPENDENT VARIABLES
0039    *
0040    *     D()      - MATRIX CONTAINING DISPLACEMENT AND VELOCITY VECTORS
0041    *                AND OF ORDER NP X NPL
0042    ************************************************************************
0043          REAL FOR(601),K(50,50),P(50),B(50),F(601),D(601,50)
0044          REAL TIM(601),ACC(601),VEL(601),DIS(601),MASS,MEANFOR
0045          REAL COMFOR(601),ERROR(601)
0046          CHARACTER*20 ACCEL,FORCE,OUTFILE,STAT
0047    ************************************************************************
0048          PRINT *,'*************************************************'
```

```
0049          PRINT *,'PLEASE INCLUDE APOSTROPHES WHEN INPUTING THE'
0050          PRINT *,'NAME OR STATUS OF ANY FILE.'
0051          PRINT *,' EXAMPLE:'
0052          PRINT *,'IF THE NAME OF YOUR OUTPUT FILE IS SEAT.OUT THEN YOU'
0053          PRINT *,'SHOULD TYPE ''SEAT.OUT'' WHEN YOU AREPROMPTEDFORTHAT'
0054          PRINT *,'FILE'
0055          PRINT *,'*************************************************'
0056          PRINT *, 'INPUT THE NAME OF YOUR ACCELERATION FILE'
0057          READ *, ACCEL
0058          PRINT *,'*************************************************'
0059          PRINT *, 'INPUT THE NAME OF YOUR FORCE FILE'
0060          READ *, FORCE
0061          PRINT *,'*************************************************'
0062          PRINT *,'THE PROGRAM WILL GENERATE AN OUTPUTFILECONSISTING OF'
0063          PRINT *,'A TABLE OF THE ACCELERATION,VELOCITY AND'
0064          PRINT *,'DISPLACEMENT VERSUS TIME.'
0065          PRINT *,'ALSO INCLUDED AT THE BOTTOM OF THE OUTPUT FILE IS A'
0066          PRINT *,'LISTING OF THE COMPUTED PARAMETERS , THE STANDARD'
0067          PRINT *,'OF DEVIATION, AND THE  COEFFICIENT OF CORRELATION.'
0068          PRINT *,'*************************************************'
0069          PRINT *,'PLEASE INDICATE WHETHER YOUROUTPUTFILE IS NEW OR OLD'
0070          READ *, STAT
0071          PRINT *,'PLEASE INPUT THE NAME OF YOUR OUTPUT FILE'
0072          READ *, OUTFILE
0073          PRINT *,'*************************************************'
0074          OPEN (10,FILE=ACCEL,STATUS='OLD')
0075          OPEN (12,FILE=FORCE,STATUS='OLD')
0076          OPEN (11,FILE=OUTFILE,STATUS=STAT,ACCESS='SEQUENTIAL')
0077    ***********************************************************************
0078          PRINT *, 'T0 = INITIAL TIME'
0079          PRINT *, 'TF = FINAL TIME'
0080          PRINT *, 'NS = NUMBER OF TIME INCREMENT'
0081          PRINT *, 'MASS = MASS OF SUBJECT'
0082          PRINT *, 'INPUT VALUES OF T0,TF,NS,MASS'
0083          READ *, T0,TF,NS,MASS
0084          NP=NS+1
0085          DO 10 I=1,NP
0086          READ(10,*) TIM(I),ACC(I)
0087          READ(12,*) TIM(I),FOR(I)
0088          F(I)=FOR(I) MASS*ACC(I)
0089     10   CONTINUE
0090    ***********************************************************************
0091          CALL TRAP(ACC,NS,T0,TF,VEL)
0092          CALL TRAP(VEL,NS,T0,TF,DIS)
0093    ***********************************************************************
0094          PRINT *,'*************************************************'
0095          PRINT *, 'INPUT NUMBER OF LUMPED PARAMETER'
0096          READ *, NLP
0097          PRINT *, 'LPI=LUMPED PARAMETER INDICATOR (1 IF DISP, 2 IF VEL)'
0098          PRINT *, 'LPP = LUMPED PARAMETER POWER'
0099          DO 15 I=1,NLP
0100          PRINT *, 'INPUT LPI,LPP OF PARAMETER ', I
0101          READ *, LPI,LPP
0102          IF (LPI.EQ.1) THEN
0103          DO 12 J=1,NP
```

```
0104          D(J,I)= (DIS(J))**LPP
0105       12 CONTINUE
0106          ELSE
0107          DO 13 J=1,NP
0108          D(J,I)= (VEL(J))**LPP
0109       13 CONTINUE
0110          END IF
0111       15 CONTINUE
0112      ************************************************************************
0113          DO 40 I=1,NLP
0114             DO 30 J=I,NLP
0115                SUM=0.0
0116                DO 20 KK=1,NP
0117                   SUM=SUM+D(KK,I)*D(KK,J)
0118       20       CONTINUE
0119                K(I,J)=SUM
0120                K(J,I)=K(I,J)
0121       30    CONTINUE
0122             SUM1=0.0
0123             DO 50 KK=1,NP
0124                SUM1=SUM1+F(KK)*D(KK,I)
0125       50    CONTINUE
0126             P(I)=SUM1
0127       40 CONTINUE
0128      ************************************************************************
0129          CALL LUD(K,P,B,NLP)
0130      ************************************************************************
0131          DO 60 I=1,NLP
0132          PRINT *,'PARAMETER NO.',I,'      VALUE =',B(I)
0133       60 CONTINUE
0134      ************************************************************************
0135          SUMERR=0.0
0136          DO 80 I=1,NP
0137          COMFOR(I)=0.0
0138             DO 70 J=1,NLP
0139                COMFOR(I)=COMFOR(I)+B(J)*D(I,J)
0140       70       CONTINUE
0141          ERROR(I)=F(I)-COMFOR(I)
0142          SUMERR=SUMERR+ERROR(I)*ERROR(I)
0143       80 CONTINUE
0144          RESPONSE=SQRT(SUMERR/NS)
0145          PRINT *,'RESPONSE =',RESPONSE
0146      ************************************************************************
0147          WRITE(11,1008)
0148          WRITE (11,1000)
0149       1000 FORMAT (T14,'TIME',T26,'DISPLACEMENT',T42,'VELOCITY',T56,
0150              X 'ACCELERATION' /)
0151       1001 FORMAT(4X,4F15.7)
0152          DO 1002 I=1,NP
0153          WRITE(11,1001) TIM(I),DIS(I),VEL(I),ACC(I)
0154       1002 CONTINUE
0155          WRITE(11,1008)
0156          WRITE(11,1003)
0157       1003 FORMAT (T5,'PARAMETER NO.', T22,'PARAMETER VALUE' /)
0158       1004 FORMAT (10X,I2,9X,E15.7)
```

```
0159          DO 1005 I=1,NLP
0160          WRITE (11,1004) I,B(I)
0161     1005 CONTINUE
0162          WRITE(11,1008)
0163          WRITE (11,1006) RESPONSE
0164     1006 FORMAT(' RESPONSE = ',F15.7)
0165          WRITE(11,1008)
0166     1008 FORMAT ('**************************************************
0167         $**************')
0168          STOP
0169          END
```

B. DETAILED DESCRIPTION

The complete program is explained in detail by proceeding through the main program first, and then returning to the subroutines in the order they are called.

DEFINITION OF VARIABLES

Variables are listed in lines 5-31 and 39-44 where they are accompanied by a brief definition. Variables are declared real and matrices are dimensioned in lines 46-49. All other variables will be integers or reals by default.

INDICATING INPUT AND OUTPUT FILES

In lines 51-79 many PRINT statements are included to assist the user in inputting the necessary data files. The user is instructed to enter the file name of the file containing the measured forces. The user then indicates the name of the file to which the output is to be sent. These files are all opened so that data may be read from them and output generated.

ERROR IN MEASURED FORCE

The value of F(I), which is calculated in lines 81-92, represents the error between the calculated force and the measured force from the model.

INTEGRATION

Lines 94 and 95 send the acceleration values to the subroutine TRAP. There they are integrated to obtain values of velocity and deflection.

LUMPED PARAMETERS

This program is extremely flexible in that it is able to use any number of lumped parameters desired. Many times a greater number of lumped parameters will more accurately simulate the given model. In lines 97-114 the number of lumped parameters and powers are given and the method of least squares is begun.

MATRICES

Lines 116-130 form matrices which contain the independent and dependent variables (the K matrix and the P matrix respectively). Each of these two matrices is the product of two matrices, and these two pairs of matrices are functions of the velocity, displacement, force error, and number of lumped parameters.

SOLVING FOR COEFFICIENTS

Line 132 calls a subroutine which uses the method of LU Decomposition to to solve for the matrix of unknown coefficients to complete the force equation. A description of this method is given later with the review of the subroutines. The coefficients are printed out in lines 134-136.

ERROR OF LINEAR REGRESSION

Lines 138-156 determine how well the least squares fit has done by calculating the standard of deviation and the coefficient of determination. A perfect fit would mean that the force equation is a precise indication of the forces in the model. This would be true if the standard of deviation were 0 and the coefficient of determination were 1. Deviations from 0 and 1 are indications of error which may have numerous sources.

OUTPUT FILES

In lines 158-179 pertinent information such as the time, displacement, velocity, acceleration, parameters, standard of deviation, and coefficient of

correlation is sent to the output file. The output is arranged neatly in tabular form of analysis.

SUBROUTINE TRAP

```
0001    ************************************************************************
0002    ************************************************************************
0003          SUBROUTINE TRAP(FF,NS,XO,XF,INTEGR)
0004    ************************************************
0005    *           DEFINITION OF VARIABLES            *
0006    *                                              *
0007    *     NS       =NUMBER OF SEGMENTS             *
0008    *     FF()     =DEPENDENT VARIABLE             *
0009    *     XO,XF    =INTEGRATION LIMITS             *
0010    *     INTEGR()=INTEGRAL                        *
0011    ************************************************
0012          REAL FF(601),INTEGR(601)
0013          INTEGR(1)=0.0
0014          H=(XF-XO)/NS
0015          DO 10 I=2,NS+1
0016          INTEGR(I)=INTEGR(I-1)+(FF(I)+FF(I-1))/2.*H
0017       10 CONTINUE
0018          RETURN
0019          END
```

This subroutine uses the trapezoidal rule as a method of integration. Values of acceleration are sent to TRAP and TRAP returns values of velocity to the master program. The velocities are in turn sent to trap and displacements are returned.

SUBROUTINE LUD

```
0001    ************************************************************
0002          SUBROUTINE LUD(A,C,X,N)
0003    ********************************************
0004    *           DEFINITION OF VARIABLES        *
0005    *                                          *
0006    *      N    = NUMBER OF EQUATIONS         *
0007    *      A[]  = MATRIX OF COEFFICIENTS      *
0008    *      C()  = RIGHT-HAND-SIDE VECTOR      *
0009    *      X()  = UNKNOWNS                    *
0010    *      O()  = ORDER VECTOR                *
0011    *      S()  = SCALE VECTOR                *
0012    ********************************************
0013          DIMENSION A(50,50),C(50),X(50),S(50)
0014          INTEGER O(50)
0015          CALL ORDER(S,A,O,N)
0016          CALL DECOMP(A,O,S,N)
0017          CALL SOLVE(A,C,X,O,N)
0018          RETURN
0019          END
```

This subroutine begins solving for the coefficients of the force equation by using the method of LU decomposition. At this point the force equation is in the form [K](B)=(P). The B matrix, which contains the coefficients of the force equation, is the one which will be solved for.

SUBROUTINE ORDER

```
0001    ********************************************
0002          SUBROUTINE ORDER (S,A,O,N)
0003          DIMENSION A(50,50),S(50)
0004          INTEGER O(50)
0005          DO 10 I=1,N
0006             O(I)=I
0007             S(I)=ABS(A(I,1))
0008             DO 20 J=2,N
0009                IF (ABS(A(I,J)).GT.S(I)) THEN
0010                   S(I)=ABS(A(I,J))
0011                ENDIF
0012       20    CONTINUE
0013       10  CONTINUE
0014          RETURN
0015          END
```

This subroutine is used in conjunction with subroutine pivot to reorder the set of equations to be solved.

SUBROUTINE DECOMP

```
0001    ********************************************
0002          SUBROUTINE DECOMP(A,O,S,N)
0003          DIMENSION A(50,50),S(50)
0004          INTEGER O(50)
0005          INTEGER*2 I,J,K
0006          J=1
0007          CALL PIVOT(A,S,O,N,J)
0008          DO 10 J=2,N
0009             A(O(1),J) = A(O(1),J)/A(O(1),1)
0010       10 CONTINUE
0011          DO 20 J=2,N-1
0012             DO 30 I=J,N
0013                SUM = 0.0
0014                DO 40 K=1,J-1
0015                   SUM=SUM+A(O(I),K)*A(O(K),J)
0016       40          CONTINUE
0017                A(O(I),J)=A(O(I),J) - SUM
0018       30       CONTINUE
0019             CALL PIVOT(A,S,O,N,J)
0020             DO 50 K=J+1,N
0021                SUM = 0.0
0022                DO 60 I=1,J-1
0023                   SUM=SUM+A(O(J),I)*A(O(I),K)
```

```
0024        60      CONTINUE
0025                A(O(J),K)=
0026          *           (A(O(J),K)-SUM)/A(O(J),J)
0027        50      CONTINUE
0028        20  CONTINUE
0029            SUM = 0.0
0030            DO 70 K=1,N-1
0031               SUM=SUM+A(O(N),K)*A(O(K),N)
0032        70  CONTINUE
0033            A(O(N),N)=A(O(N),N)-SUM
0034            RETURN
0035            END
```

This subroutine involves decomposing the set of simultaneous equations into a product of lower and upper matrices.

SUBROUTINE PIVOT

```
0001    ******************************************
0002            SUBROUTINE PIVOT(A,S,O,N,J)
0003            DIMENSION A(50,50),S(50)
0004            INTEGER O(50)
0005            INTEGER*2 J,PIVIT,II,IDUM
0006            PIVIT = J
0007            BIG=ABS(A(O(J),J)/S(O(J)))
0008            DO 10 II=J+1,N
0009               DUMMY=ABS(A(O(II),J)/S(O(II)))
0010               IF (DUMMY.GT.BIG) THEN
0011                  BIG=DUMMY
0012                  PIVIT=II
0013               ENDIF
0014        10  CONTINUE
0015            IDUM=O(PIVIT)
0016            O(PIVIT)=O(J)
0017            O(J)=IDUM
0018            RETURN
0019            END
```

This subroutine performs a partial pivoting to avoid division by zero and to minimize round-off-errors.

SUBROUTINE SOLVE

```
0001    ******************************************
0002            SUBROUTINE SOLVE(A,C,X,O,N)
0003            DIMENSION A(50,50),C(50),X(50)
0004            INTEGER O(50)
0005            X(1)=C(O(1))/A(O(1),1)
0006            DO 10 I=2,N
0007               SUM = 0.0
0008               DO 20 J=1,I-1
0009                  SUM=SUM+A(O(I),J)*X(J)
```

```
0010       20     CONTINUE
0011              X(I) = (C(O(I))-SUM)/A(O(I),I)
0012       10  CONTINUE
0013           DO 30 I=N-1,1,-1
0014              SUM = 0.0
0015              DO 40 J=I+1,N
0016                 SUM=SUM+A(O(I),J)*X(J)
0017       40     CONTINUE
0018              X(I)=X(I)-SUM
0019       30  CONTINUE
0020           RETURN
0021           END
```

Simply, this subroutine solves the set of equations after they have been decomposed by a trivial combination of forward and back substitutions.

IV. OUTPUT

A sample output file is shown below:

```
**************************************************************************
            TIME         DISPLACEMENT     VELOCITY
ACCELERATION

          .0000000        .0000000        .0000000        .0000000
         1.0000000        .0000000        .0000000        .0000000
         2.0000000        .0000000        .0000000        .0000000
             .                .               .               .
             .                .               .               .
             .                .               .               .
             .                .               .               .
             .                .               .               .
             .                .               .               .
             .                .               .               .
             .                .               .               .
             .                .               .               .
        75.0000000        .0000127        .0011588        .1287600
        76.0000000        .0000139        .0012409        .0354100
        77.0000000        .0000151        .0012763        .0354100
        78.0000000        .0000164        .0012940        .0000000
        79.0000000        .0000177        .0012940        .0000000
        80.0000000        .0000190        .0012473       -.0933501
        81.0000000        .0000202        .0012006        .0000000
             .                .               .               .
             .                .               .               .
             .                .               .               .
             .                .               .               .
             .                .               .               .

       595.0000000        .0193880        .0806900       -.1770500
       596.0000000        .0194686        .0804485       -.3058300
       597.0000000        .0195489        .0801604       -.2704100
       598.0000000        .0196289        .0799898       -.0708300
```

```
      599.0000000        .0197089        .0800365        .1641500
      600.0000000        .0197891        .0802360        .2349700
**************************************************************************
   PARAMETER NO.     PARAMETER VALUE

          1            -.3132372E+04
          2             .3041510E+04
**************************************************************************
 RESPONSE =     642.6951000
**************************************************************************
```

The output of program SEAT reveals the experimentation errors encountered in the laboratory. To obtain useful data these experimentation errors must be identified, and methods must be found to compensate for them. The following paragraphs will discuss observations of experimental error, the source of the error, and the methods used to compensate for the error.

A. INITIAL OUTPUT

The original output from the Data Acquisition System is plotted in Figure 4. From an understanding of the fundamental principles of motion, the errors in this output are obvious. The acceleration must begin at zero, since this would indicate conditions prior to acceleration. The output, however, begins at a point other than zero. Figure 5 represents the velocities found in the program SEAT. These velocities were found using the accelerations in Figure 4. From the basic principles of motion it can be expected that the velocity will be zero prior to the test, and will return to zero at the end of the test. The plot clearly shows the velocity trailing off without returning to zero. The resulting displacements, plotted in Figure 6 exhibit the same erroneous characteristics. The displacement should begin and end at zero, yet instead it trails off to come other value.

The velocities and displacements are produced by the program SEAT, and are dependent upon that program's input. The input for program SEAT is the accelerations recorded from the accelerometer. These accelerations then are the basic source of error in the program SEAT. The error in the

Figure 4

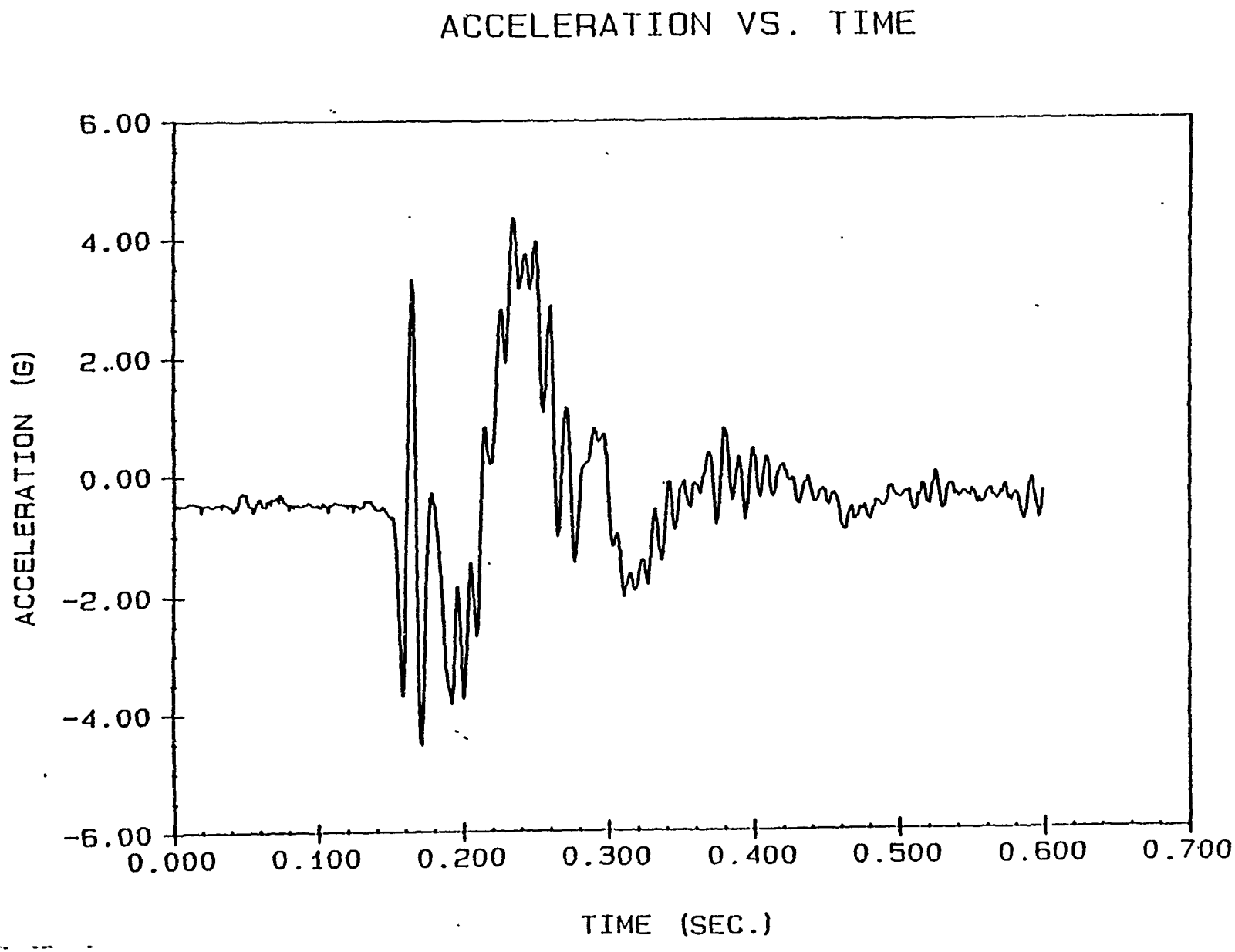
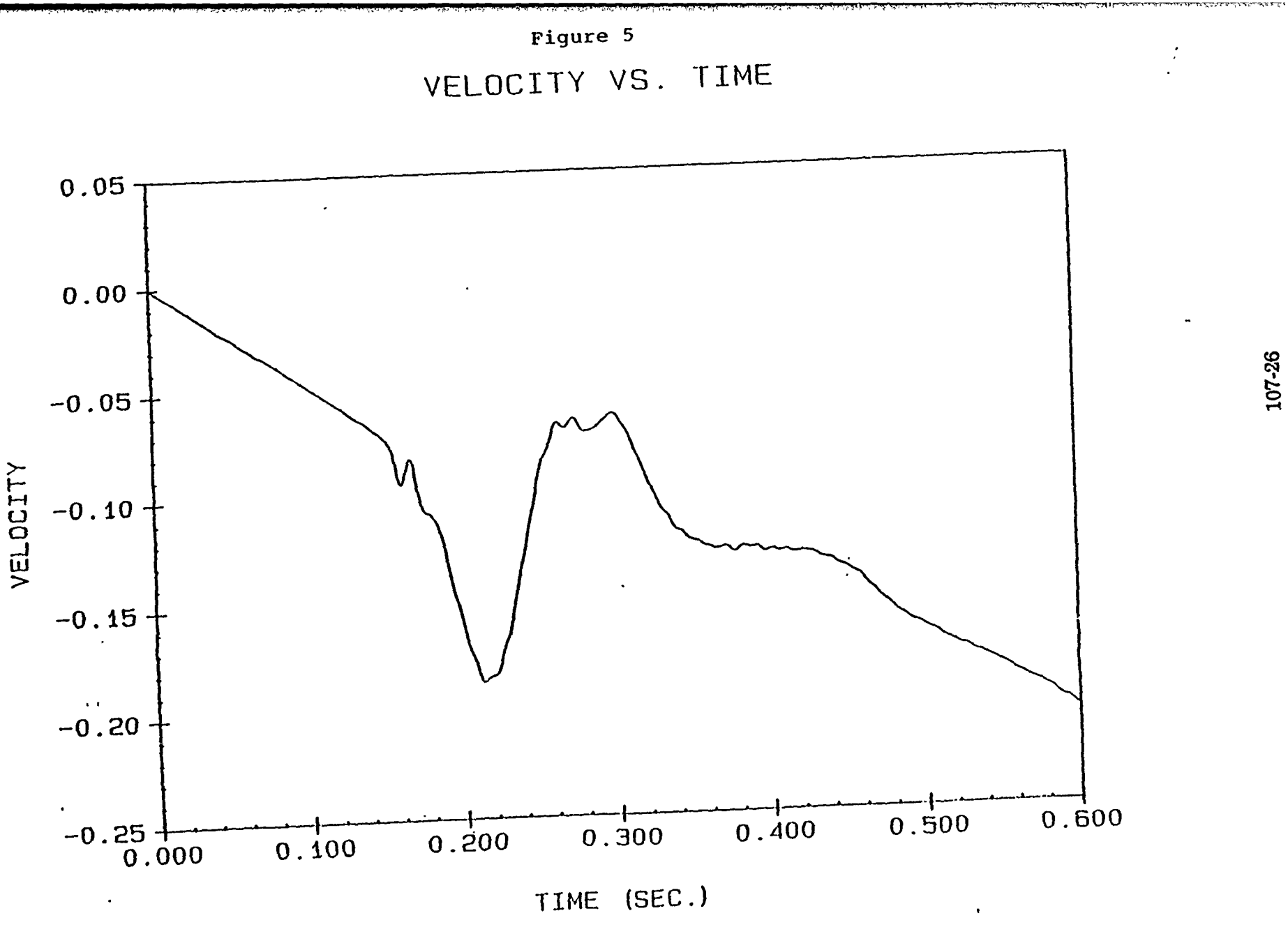
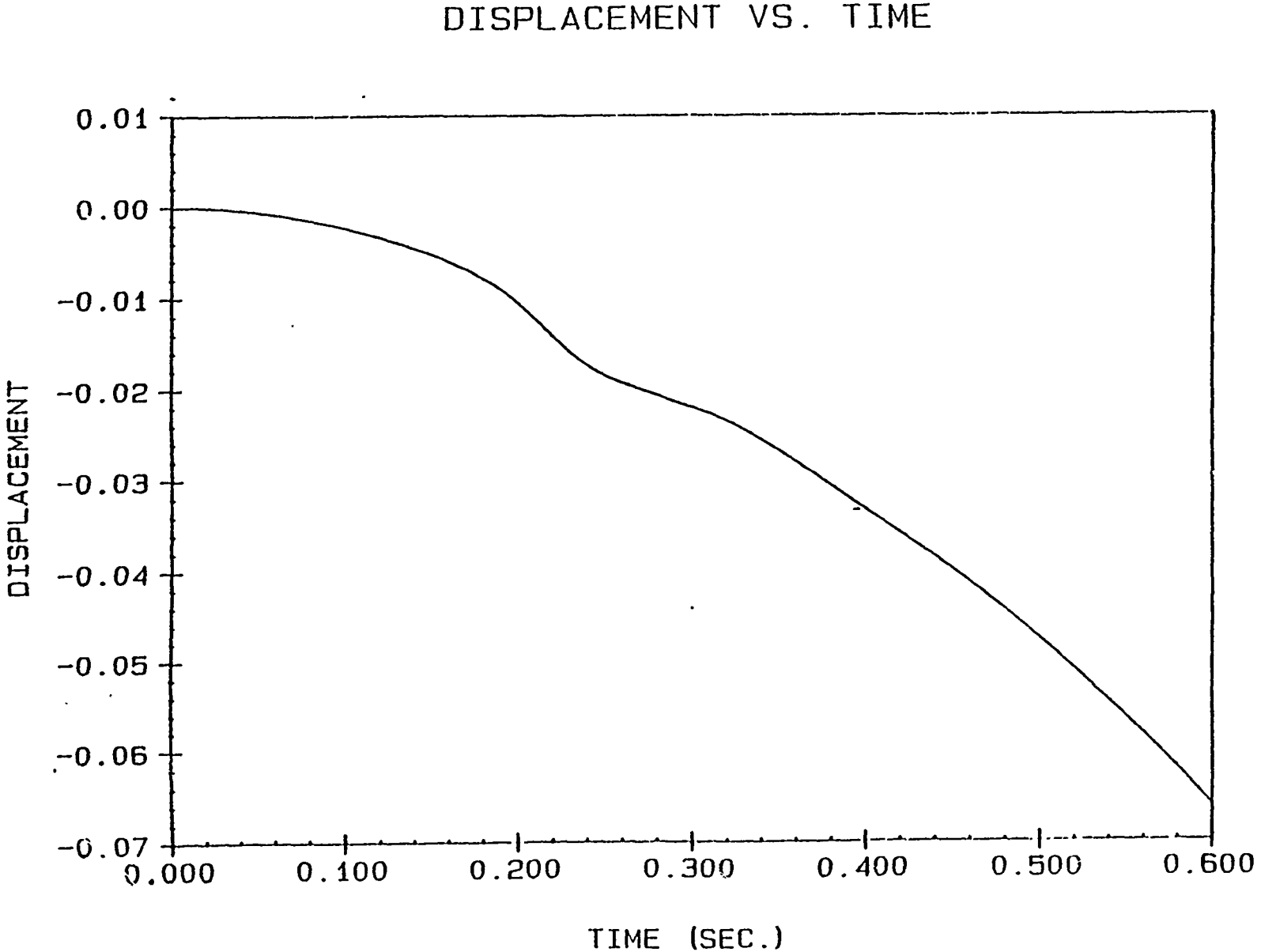

ACCELERATION VS. TIME
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
ACCELERATION (G)
0.000
0.100
0.200
0.300
0.400
0.500
0.600
0.700
TIME (SEC.)

Figure 5



107-26

Figure 6


DISPLACEMENT VS. TIME
0.01
0.00
-0.01
-0.02
-0.03
-0.04
-0.05
-0.06
-0.07
DISPLACEMENT
0.000
0.100
0.200
0.300
0.400
0.500
0.600
TIME (SEC.)

accelerations, themselves, must be traced to the model and accelerometer. The accelerometer, not being rigidly attached to the chest of the model, was allowed to jiggle during testing. This jiggling motion unbalanced the accelerometer, forcing its original position away from zero. A perfectly rigid attachment of the accelerometer is not feasible in this situation, so the data must now be manipulated to simulate a rigid attachment. The most simplistic manipulation is to shift the acceleration data so that the initial value is forced to zero.

B. SHIFTED OUTPUT

Figure 7 is a plot of the same acceleration data as Figure 4, however, the curve has been shifted up to bring the initial acceleration value to zero. The subsequent data has also been shifted, resulting in completely new acceleration values to be used as input into program SEAT. Figures 8 and 9 are plots of the velocities and displacements found by program SEAT. Again, the values trail off, and do not return to zero. The shifting has been beneficial in that the values of acceleration are now reasonable, however, problems still remain in that the velocities and displacements fail to return to zero.

C. SMOOTHED OUTPUT

The final manipulation performed on the output was a data smoothing. Smoothing techniques are purely graphical methods used for interpreting scattered points. Instead of plotting a curve which passes through all points in a collection of output, the smoothing chooses a path which most closely represents the trend of the data. Smoothing, in this case, was used to simulate a steadying of the accelerometer. This compensated for the scattering of points which was caused by the jiggling of the non-rigidly attached accelerometer. The method used was a Fast Fourier Transform which

Figure 7

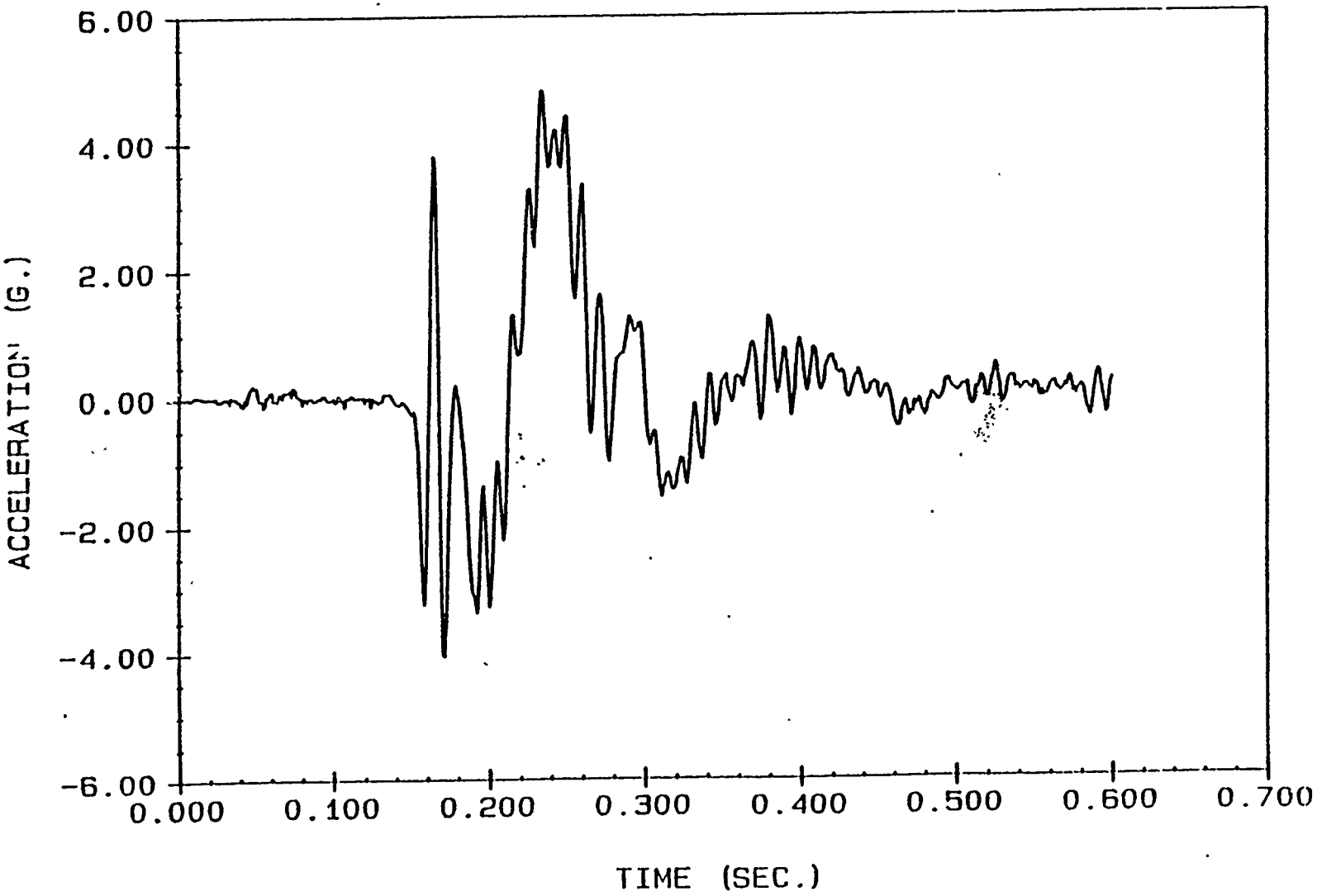

ACCELERATION VS. TIME
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
ACCELERATION (G.)
0.000
0.100
0.200
0.300
0.400
0.500
0.600
0.700
TIME (SEC.)

Figure 8

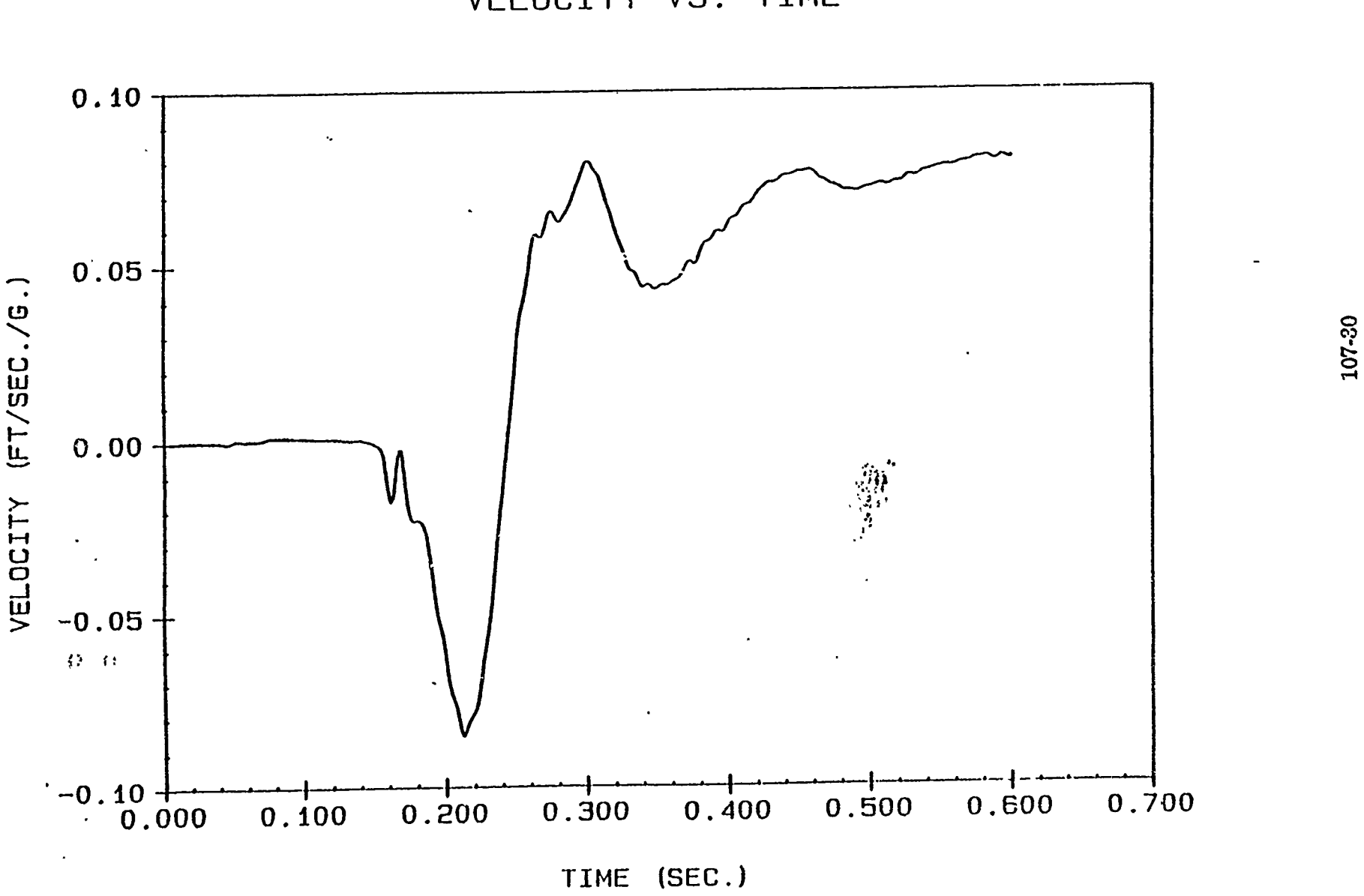
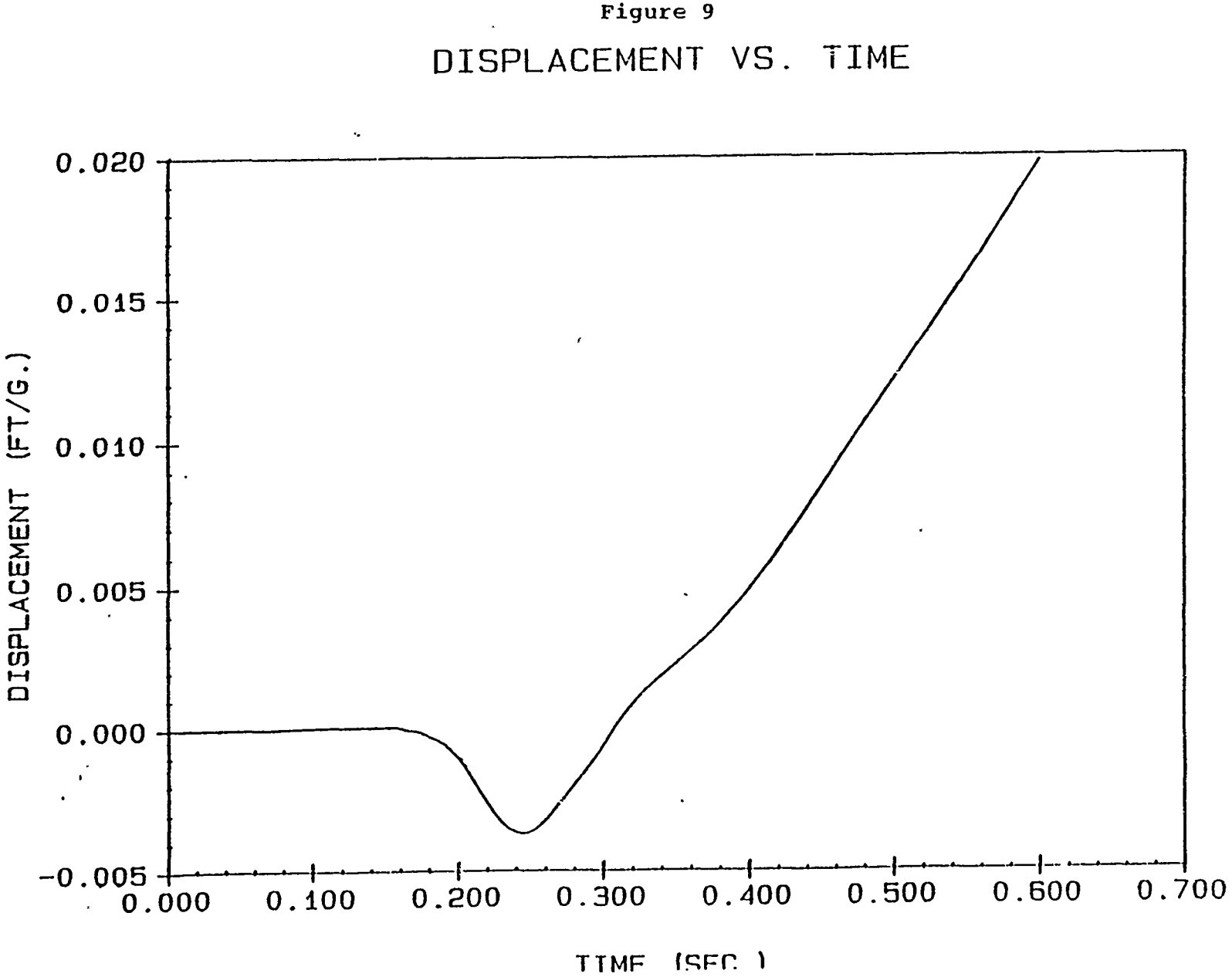

VELOCITY VS. TIME
0.10
0.05
0.00
-0.05
-0.10
VELOCITY (FT/SEC./G.)
0.000
0.100
0.200
0.300
0.400
0.500
0.600
0.700
TIME (SEC.)

Figure 9


DISPLACEMENT VS. TIME
DISPLACEMENT (FT/G.)
0.020
0.015
0.010
0.005
0.000
-0.005
0.000
0.100
0.200
0.300
0.400
0.500
0.600
0.700
TIME (SEC.)

acted as a filter to eliminate the constant bouncing of the data points. To see the results of the smoothing a comparison between Figures 4 and 10 must be made. The plot in Figure 10 contains the filtered, or smoothed, data. The subsequent velocities and displacements are plotted in Figures 11 and 12 respectively. The velocities now come close to returning to zero. The displacements, however, seem unaffected by the smoothing.

It must be kept in mind that data smoothing is a graphical technique, and is not justified for use as a numerical approximation method. It is used here to simulate a reduction of movement in the accelerometer.

Figure 10

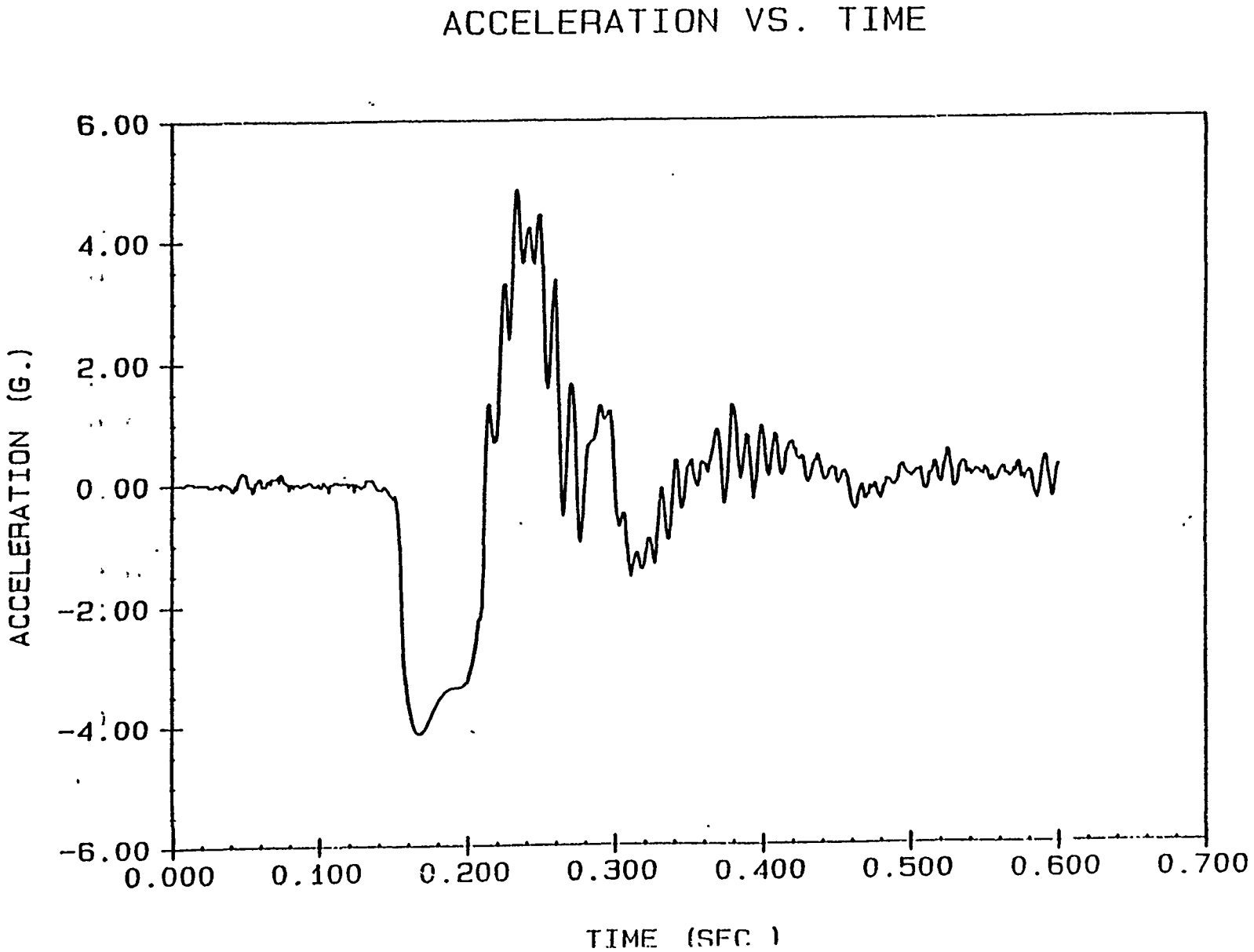

ACCELERATION VS. TIME
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
ACCELERATION (G.)
0.000
0.100
0.200
0.300
0.400
0.500
0.600
0.700
TIME (SEC.)

Figure 11

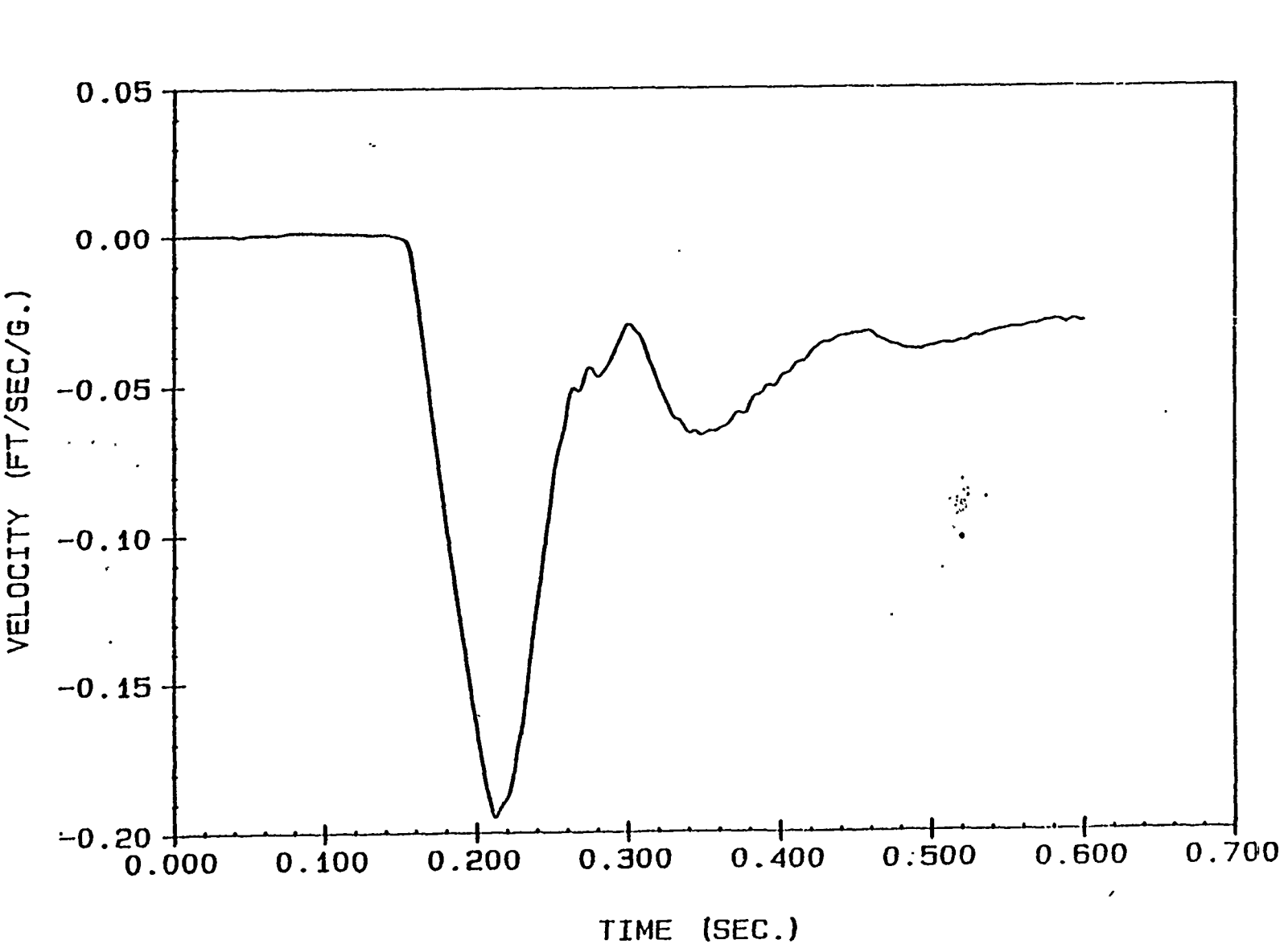
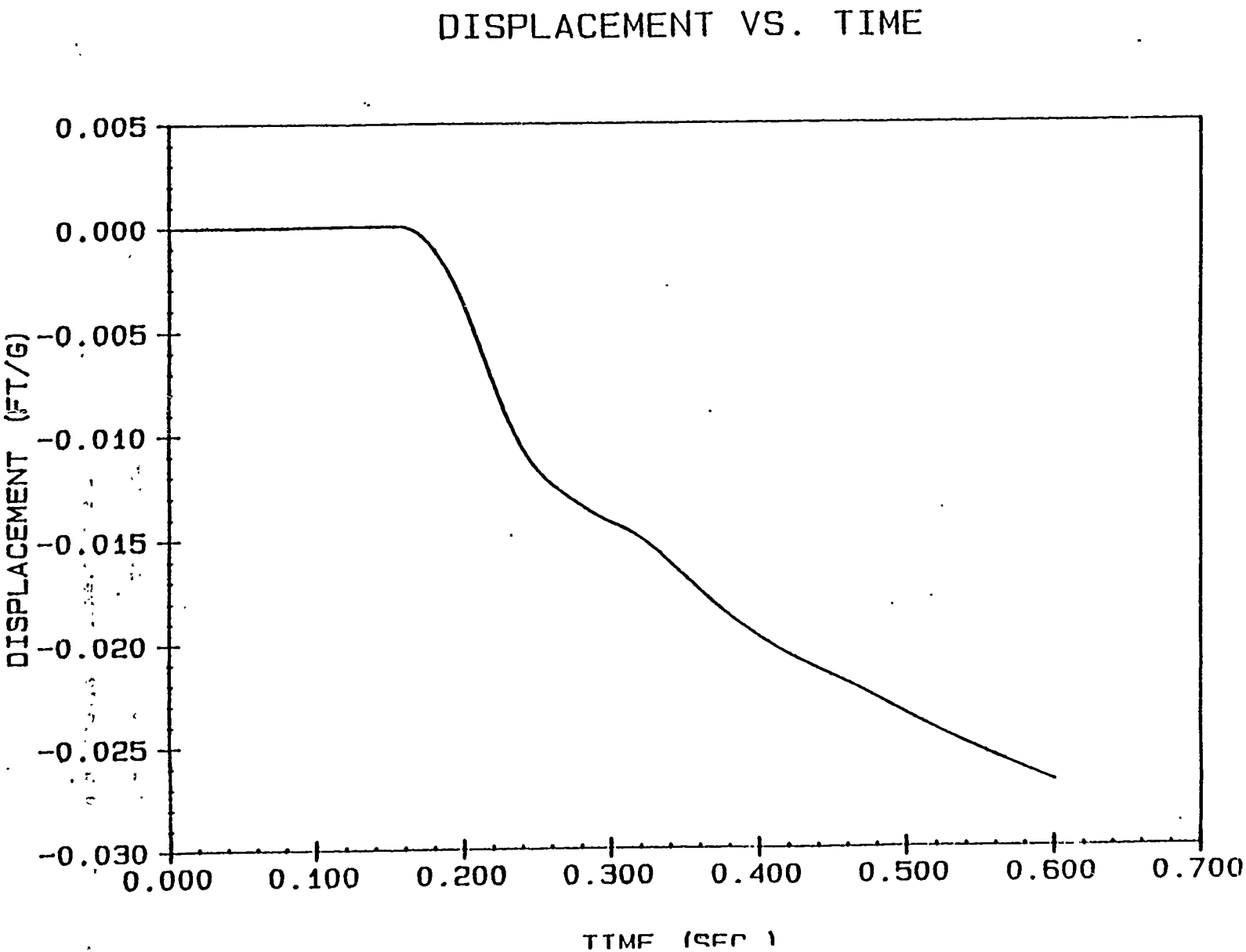

VELOCITY VS. TIME
0.05
0.00
-0.05
-0.10
-0.15
-0.20
VELOCITY (FT/SEC/G.)
0.000
0.100
0.200
0.300
0.400
0.500
0.600
0.700
TIME (SEC.)

Figure 12


DISPLACEMENT VS. TIME
0.005
0.000
-0.005
-0.010
-0.015
-0.020
-0.025
-0.030
DISPLACEMENT (FT/G)
0.000
0.100
0.200
0.300
0.400
0.500
0.600
0.700
TIME (SEC)

V. RECOMMENDATIONS

Due to the nature of the project and its complexity, validation of this nonlinear model was accomplished on one single test alone. A more realistic and comprehensive validation plan should be proposed and implemented to ensure the effectiveness of this nonlinear lumped-parameter model. This plan should include a wide spectrum of excitation magnitudes (4g to 10g) and also different excitation durations and shapes.

Although this model was developed for the A direction, the methodology is the same for a similar model in the other two directions. Similar validation plans can be used to determine the effectiveness in the X and Y directions.

The same method can also be used on future spiral models of impact which will also include the abdominal musculature and interaction of ribs with the thoratic vertebrae. This will mean that more lumped masses should be included, as well as more spring and viscous damping forces.

REFERENCES

1. Carmichael, J.B., Jr. (1968) Asst. Chief, Vibrations and Impact Branch, Biodynamics and Bionics Division, Wright-Patterson AFB, Ohio. Correspondence.

2. Wittman, T.J. and Phillips, N.S. (1969) "Human body non-linearity and mechanical impedance analyses". J. Biomechanics 2, pp. 281-288.

3. Perng, F.G. (1970) "Strain energy function for biological tissue". ASME Paper No. 70-BHF-11.

4. Markoff, K.L. and Steidel, R. F., Jr. (1970) "The dynamic characteristics of the human interverteral joint". ASME Paper No. 70-WA/BHF-11.

5. Belytschko, T., Schultz, A.B., Kulak, R. F. and Galante, J. O. (1972) "Numerical stress analysis of intervertebral disk." ASME Paper No. 72-WA/BHF-12.

6. Latham, F. (1957) "A study in body ballistics: seat ejection". Proc. R. Soc. (B) 147. pp.121-139.

7. Payne, P. R. (1962) "The dynamics of human restraint systems, in impact acceleration stress". Publ. No. 977, Nat. Acad. Sci.--National Research Council. pp.195-257. Washington DC.

8. Hess, J. L. and Lombard, C. F. (1958) "Theoretical investigations of dynamic response of man to high vertical accelerations". Aviation Medicine 29, pp. 66-75.

1989 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Universal Energy Systems, Inc.

FINAL REPORT

IN VITRO MODELING OF PERFLUORO-N-DECANOATE EFFECTS ON ENZYMES OF FATTY ACID METABOLISM

Prepared by: SANFORD S. SINGER PH.D.

Academic Rank: PROFESSOR

Department: CHEMISTRY DEPARTMENT

University: UNIVERSITY OF DAYTON

Research Location: Rooms 412, 416-418 Wohlleben Hall
University of Dayton

Date Submitted 12 December 1969

Grant/Contract No.: S-49620-88-C-0053

IN-VITRO MODELING OF PERFLUORO-N-DECANOATE EFFECTS ON ENZYMES OF FATTY ACID METABOLISM

by

Sanford S. Singer

ABSTRACT: Rat liver carnitine acyltransferase, acyl-SCoA oxidase, acyl-SCoA synthetase, & acyl-SCoA hydrolase were isolated by literature methods we modified as neeeded. The enzymes were examined, with palmitate & laurate or palmitoyl-SCoA & lauroyl-SCoA substrates(where appropriate) to model, in vitro, the basis for effects of perfluoro-n-decanoate (PFDA) on fatty acid metabolism. We also attempted organic synthesis of PFDA-SCoA for use in parallel enzyme studies. Methodologic difficulties complicated many aspects of the work. Good progress was made, considering funds available. Our present examination of PFDA effects on the test enzymes produced useful data & provides a beginning for understanding in-vivo effects of PFDA on lipid metabolism. As to:

(1) Carnitine Acyltransferase(CAT) Enzyme was prepared as reported in the literature, with one modification of the isolation method. Six CAT preparations were isolated, frozen at -25° C, & used over time periods up to six months. They were quite stable(e.g., one preparation, reexamined after 6 months, retained 75-85% of initial activity). Assay was optimized for palmitoyl transfer to L-carnitine. Comparison showed lauroyl-SCoA to be a better substrate than palmitoyl-SCOA. PFDA, at concentrations from 55 to 440 µM, activated acyl transfer from both substrates. It was a more effective activator of palmitoyl transfer than of lauroyl transfer. Double reciprocal plots gave typical data in the absence of PFDA. However, we were unable to obtain K_ms(or activation constants) in the presence of PFDA because PFDA turned CAT on to a maximum rate--regardless of the acyl-SCoA level used as substrate-- & this rate was maintained until substrate gave out. Later reexamination of CAT preparations showed that this exciting PFDA effect of begins to disappear after 3 to 4 months of frozen storage. It is likely that the effect is due to a conformation of CAT that eventually disappears upon frozen storage.

(2) Acyl-SCoA Oxidase(AO) Enzyme was purified by combination of aspects of two reported literature methods, required for success. Six AO preparations were isolated, stored at -25° C, & used over time periods of up to four months. They were stable frozen(e.g., after 4 months one preparation retained 65-75% of its initial activity). Thawed AO was very unstable & use required addition of bovine serum albumin for assays. AO assay was optimized--within its limitations--for palmitoyl-SCoA oxidation. Comparison of 80 μM lauroyl- and palmitoyl-SCoAs showed lauroyl-SCoA to be a better AO substrate than the C-18 acyl-SCOA(100% vs 69.0±6.5%). PFDA inhibited AO action on both acyl-SCoAs, as follows. First, 440 μM PFDA inhibited oxidation of 80 μM palmitoyl-SCoA or lauroyl-SCoA 22.3±2.3% or 53.2±4.3%, respectively. Stronger inhibition of 80 μM lauroyl-SCoA oxidation was supported by observation that 36 μM PFDA inhibited it, but had no effect with 80 μM palmitoyl-SCoA. PFDA was a competitive inhibitor of palmitoyl-SCoA oxidation, K_I = 318±58 μM. We could not obtain K_I values with lauroyl-SCoA. However, trends here supported PFDA inhibition as being competive, with a low K_I(possibly 1/3 that for palmitoyl-SCoA). Problems encountered were high lauroyl-SCoA lability during the reaction & the instability of the enzyme. This made kinetic study with lauroyl-SCoA quite difficult.

(3) Acyl-SCoA Synthetase(AS) Five complete AS purifications were attempted. The last two were unsuccessful because the final chromatography step produced inactivate AS. Our efforts with the successful preparations used palmitate & laurate. They were comparable substrates at 143 μM concentrations. In contrast, AS did not convert appreciable amounts of PFDA to PFDA-SCoA with 110 or 440 μM PFDA substrate. However, 110 μM PFDA inhibited conversion of both fatty acids to acyl-SCoA. Inhibitions of palmitoyl- and lauroyl-SCoA production--from 143 μM substrate--were 70±16% and 88±20%, respectively.

(4) Acyl-SCoA Hydrolase(AH) We were unable to demonstrate AH activity in homogenates or partly purified enzyme preparations, using the spectrophotometric assay methodology proposed. It may be possible to study AH in the future, with an alternate radioisotope method. Time & budgetary constraints precluded further effort here.

(5) PFDA-SCoA Synthesis We attempted to prepare PFDA-SCoA for use in enzyme inhibition studies to parallel study of the PFDA effect. The method utilized was recommended to us as the best procedure available. Organic synthesis was attempted 4 times, with absolutely no success. Efforts to this end were dropped because: a. We noted that lauroyl-SCoA was much less stable than palmitoyl-SCoA & this was taken to indicate that PFDA-SCoA was just too unstable for isolation. b. We found that rat liver acyl-SCoA synthetase did not make an appreciable amount of PFDA-SCoA, so planned enzymologic studies did not appear to be essential. c. Conference with another researcher who had attempted the organic synthesis indicated a similar lack of success.

ACKNOWLEDGEMENT

I wish to thank the Air Force Systems Command & the Air Force Office for Scientific Research for sponsoring our effort in the Research Initiation Minigrant Program for 1988-1989. My research effort was a continuation of a summer 1988 involvement at AAMRL/TH & I am grateful to Dr. Melvin Andersen for giving me this research opportunity. My thanks also to Marilyn George for invaluable aid in procurement, & logistics aspects. Universal Energy Systems is to be commended for administrative effort in the program.

I also wish to thank my research assistants Cathy Knapke, Kate McGee, & Susan Payson. These three excellent young ladies spent an average of 600 hours each on the project. Some this effort was a ten week effort remunerated with a modest stipend from grant funds. However, they each also gave generously of their time when they were remunerated only by their interest in the project & its use in their senior research thesis efforts.

I. INTRODUCTION

A. RELATIONSHIP TO EFFORTS AT AAMRL/TH Ten-carbon, perfluorodecanoic acid (PFDA) was recently found to be toxic. In recent years, a number of related fluorinated chemicals, used in places including foam fire extinguishers & vascular fluid replacements, have also been shown to be toxic(see IB). Air Force personnel may be exposed to such chemicals. Consequently, evaluation of the extent of & basis for their toxicity is important.

The Biochemistry Branch of AAMRL/TH is actively engaged in evaluating the toxicology of PFDA & related chemicals & in toxicologic modeling of the processes. Whole animal studies have been conducted with rats & other species in which the pathology, variation in species mortality & alteration of energy & lipid metabolism have been determined. These studies suggest a central role for liver in the toxicity. It is presently unclear what the basis for the response to PFDA is. However, it is clear that the responses must involve enzymes of lipid metabolism.

Understanding of the enzymatic basis for changes of lipid metabolism is likely to be obtained from examination of the enzymes involved in the process. Also, appropriate utilization of in-vitro enzymology should add a new dimension to screening chemicals deemed likely to operate like PFDA. The advantages of such study over whole animal studies are simplicity & amenability to experimental manipulation, lack of need for large numbers of animals on a routine basis, cost advantages & diminished researcher risk due to the smaller quantities of toxicant/reagent utilized for evaluation processes. The use of in-vitro enzymology is complicated by limitation of the exactitude of interpretation of the relationship to in-vivo situations, where the information base relating to in-vivo concentrations of metabolic products is not complete & needs for specialized equipment & personnel with training in enzyme purification/enzymologic technique.

My enzymologic research interests & long experience in purification, characterization & manipulation of enzymes suits me to carry out the effort. This fits with the AAMRL/TH move to use of in-vitro technology, in addition to whole animal studies, for toxicologic estimations.

B. TECHNICAL BACKGROUND: Perfluorocarboxylic acids & related compounds have many industrial uses(1-3), including corrosion inhibitors, hydraulic fluids, wetting agents, foam fire extinguishers, coatings that impart water & oil resistance to paper & fabrics. Such uses have been predicated upon their supposed inertness & lack of toxicity. Recent study(4-6) has shown that perfluorinated carboxylic acids of chain length exceeding 8 carbons are toxic. Perfluoro-n-decanoic acid(PFDA), $CF_3(CF_2)_8COOH$, has been shown to be one of the more toxic of these compounds in rats & other species(5,7).

Most toxicologic research on PFDA has been carried out with Fisher 344 rats. There, a single i.p. dose of 41 mg PFDA kills 50% of injected rats. Toxic manifestations of PFDA(5,6) ressemble those of 2,3,7, 8-tetrachloro-p-dioxin. These include acute anorexia, extensive weight loss, testis & thymus atrophy, bone marrow depression, excessive liver size & disrupted hepatic architecture. Both toxic agents also alter cellular lipid levels in liver, relating particularly to early alterations of fatty acid levels (8,9). Although the mechanism of the alterations is unclear, it appears likely to include alterations of lipid metabolism.

To this end, it is already clear(10) that PFDA leads to peroxisome proliferation & large elevations of apparent acyl-SCoA oxidase activity due to those organelles. Such enzyme elevation could alter production & disposition of cellular fatty acids, as the enzyme will not use short chain acyl-SCoA substrates(11). Disruption of hepatic fatty acid metabolism could disturb cell function, leading to the toxicity observed, by altering the composition & properties of cellular membranes &/or altering the availability of lipids for energy production, hormone production & other cell processes.

PFDA itself could have disruptive effects on lipid metabolism, by interacting with various enzymes involved. These interactions could cause different activation or inactivation of various enzymes. Certainly, it has been shown that PFDA is retained at high concentrations in hepatocytes for 30 days after initial insult with a 50 mg/kg dose(9). Also, PFDA appears to be converted to a polar metabolite(9). This molecule may also contribute to PFDA action. Or, alternatively, it may be responsible for "late effects" of PFDA, exemplified by a bimodal response of acyl-SCoA oxidase described in a recent report(10).

More information is required to explain the basis for lipotransformations resultant from PFDA administration. It would seem possible to model the basis for the effects of PFDA & its polar metabolite, using purified enzymes. Such modeling could enable us to ascertain why observed alteration of fatty acid metabolism occurred. It would also provide insight into metabolites to which PFDA could be converted.

C. INITIAL EFFORTS BY PRINCIPAL INVESTIGATOR

Commercially available, nonrat acyl-SCoA synthetase, acyl-SCoA oxidase & carnitine acetyltransferase were studied in our efforts(during the 1988 SFRP)to model, in vitro, the basis for effects of perfluoro-n-decanoate (PFDA) on fatty acid metabolism. We found that:

1) Pseudomonas Acyl-SCoA synthetase used palmitate, oleate, laurate & decanoate as substrates to similar extents. However, it converted PFDA to PFDA-SCoA very slowly. PFDA inhibited acyl-SCoA formation from the fatty acids. The inhibitions appeared to be competitive. Palmitoyl-SCoA formation was inhibited most & decanoyl-SCoA formation was inhibited least. Palmitcyl -SCoA formation was inhibited up to 30% when the [PFDA]/[palmitate] was 4.

2) Candida Acyl-SCoA oxidase used palmitoyl-, lauroyl- & decanoyl-SCoA as substrates. It preferred the smaller acyl-SCoAs. Inhibition of oxidation of the C-10 & C-12 acyl-SCoAS by PFDA was more extensive than that of palmitoyl-SCoA. PFDA inhibition of decanoyl-SCoA & palmitoyl-SCoA oxidation was examined & found to be competitive, with K_Is of 593 ± 150 µM & 76 ± 6.0 µM.

3) Pigeon Liver Carnitine acetyltransferase used acetyl-SCoA as its best substrate. Butyryl-, hexanoyl-, & octanoyl-SCoA were less effective substrates than acetyl-SCoA. Transfer of all acyl groups to carnitine was inhibited to a similar extent by PFDA. The K_Is were 111 ± 15 µM & 76.0 ± 28 µM with the C-2 & C-8 acyl-SCoAs. Inhibition was competitive with acetyl-SCoA & noncompetitive with octanoyl-SCoA.

Our examination of the PFDA effects on the enzymes gave data that may provide connections between isofunctional enzymes of rat liver & in-vivo effects of PFDA on lipid metabolism in rats. Study of effects of PFDA-SCoA was tabled, as it was not available at AAMRL/TH during the SFRP.

II. SUMMARY OF PROPOSED RESEARCH GOALS:

1. Examine the enzymology of & purify three or more of the following rat liver enzymes: acyl(palmitoyl)-SCoA synthetase, acyl(palmitoyl)-SCoA oxidase, carnitine acyl(palmityol)transferase, & palmitcyl-SCoA dehydrogenase by methods reported in the literature.

2. Test the purified enzymes to identify the effects of PFDA.

3. Examine the ability of rat liver acyl(palmitoyl)-SCoA synthetase to make PFDA-SCoA in vitro.

4. Prepare PFDA-SCoA by organic synthesis & examine its effects as as done for PFDA in goal 2, if justified.

5. Plan to examine, other crucial enzymes of fatty acid metabolism in rat liver(e.g., acyl desaturase) & how to use the group of 7 or more enzymes to assess the toxicity of potentially hazardous chemicals(e.g., CTFE) in <u>in vivo</u> studies.

III. STUDY OF CARNITINE ACYL(PALMITOYL)TRANSFERASE

A. MATERIALS AND METHODS

<u>Supplies(all from Sigma)</u> CATs are involved in transport of acyl -SCoAs across mitochondrial membranes. We studied the enzyme form that transports palmitoyl groups(& others down to lauroyl groups). Its study models whether PFDA affects this transport in ways that help to explain observed changes of fatty acid metabolism & energy production after PFDA. Supplies needed are 5,5'-dithiobis-2-nitrobenzoic acid(DTNB), L-carnitine, palmitoyl -SCoA, lauroyl-SCoA, tetrasodium EDTA, KH_2PO_4, propylene glycol, & PFDA. Assay, as shown below, uses reaction of HSCoA(freed by transfer of acyl groups to carnitine) with DTNB to yield a yellow product, followed spectrophotometrically at 412 nm.

$$\text{Acyl-SCoA + Carnitine} \xrightarrow{\text{Acetyltransferase}} \text{HS-CoA + Acyl-Carnitine}$$

$$\text{HS-CoA + NBA-SS-NBA} \xrightarrow{\text{Nonenzymatic}} \text{NBA-SS-NBA + NBA-SH(yellow)}$$

Solutions used:

Assay Buffer(.75 M KH_2PO_4). Neutralize to pH 7.25 in .8 vol water, dilute to desired vol & reneutralize. Store at 4 C. Use for up to 2 weeks.

Dilution Buffer Dilute assay buffer 1:10 & check pH. Use for making all other solutions. Prepare as needed, store at 4 C, use for up to 2 weeks.

2.50 mM Acyl-SCoA Prepare 5 mL of each acyl-SCoA by adding appropriate amounts of solid to 4.0 mL dilution buffer. Quickly neutralize to pH 7.25, bring to 5 mL & store frozen in 1 mL aliquots. Use for up to 2 weeks (palmitoyl-SCoA) or 1 week(laurcyl-SCoA). Refreeze no more than twice.

10.0 mM DTNB Dissolve DTNB in 25 dilution buffer. The solution is used for 2 weeks, or until it turns noticably yellow.

15.0 mM L-Carnitine Dissolve carnitine in .8 vol dilution buffer & neutralize to pH 7.25. Bring to desired volume. Store at 4 C. Use for 2 weeks.

Enzyme Diluant Mix dilution buffer & glycerol 1:1 (v/v) & vortex. Diluant is used for a month, stored in the freezer.

PFDA diluant This diluant, propylene glycol:dilution buffer:¼% Triton X-100(2:1:1 by volume), allows use of PFDA in inhibitor study. It is added to all reaction mixtures. It is stable for 1 month, frozen.

4.40 mMPFDA This solution is prepared in PFDA diluant. Be sure to vortex & allow the foamy solution to settle, to be sure that all PFDA dissolves. Then, neutralize to pH 7.25. Store & use for up to 1 month month in the freezer. Dilute as needed for study. Store diluted samples similarly. **Be cautious this solution is not only toxic, but it causes skin burns.**

The Enzyme Assay For each 4 reaction mixtures mix: .40 mL acyl-SCoA, .40 mL DTNB, .80 mL assay buffer, 2.4 mL carnitine & preincubate at 25 C for 5 minutes. Then, transfer .80 mL aliquots to cuvets. Compare cuvets to a water blank at 412 nm in a spectrophotometer at 25 C. Next, add 0.10 mL PFDA diluant or PFDA & between 0 & .10 mM enzyme(diluted as needed) or enzyme diluant, to bring to a total volume of 1.00 mL Enzyme additions are made last, 10 seconds apart, with immediate mixing. Then, readings are taken at 2 minute intervals, for up to 14 minutes. The reaction rate is constant for 8-12 minutes, depending on enzyme content.

Basic Enzyme Purification The 2.5-day reported method(12) uses 60 g liver from male rats(200 g BW), fasted overnight. Chemicals are from Sigma. Solution preparation is not described, for the sake of brevity.

Livers are homogenized in 10 vol 0.25 M sucrose-5 mM HEPES buffer, pH 7.4(2 strokes with loose pestle), followed by centrifugation (1,500 x g, 10 min). Supernatant is then centrifuged(12,000 x g, 10 min) & mitochondrial pellet is washed twice by the same centrifugation method. It is resuspended in the same buffer at 40-50 mg protein/mL. Digitonin(60 mg/mL in 5 mM Tris, pH 7.4)is added to a concentration of 6 mg/mL(.12 mg/mg protein),followed by bovine serum albumin to .5 mg/mL. Ten min after addition of digitonin 3 vol .25 M sucrose-5 mM Tris(pH 7.4),.5 mg/mL bovine serum albumin is added The mixture is centrifuged(10,000 x g, 20 min). The pellet is resuspended in 5 mL .1% Tween 80, .4 M KCl, .125 M tris, pH 8.0(Buffer A) & sonicated, two 50 sec bursts, followed by centrifugation(48,000 x g, 10 min). The supernatant is saved. The pellet is suspended in 5 mL Buffer A, resonicated & recentrifuged. The supernatants are combined & mixed with 10 mL saturated AMS & stirred gently for 15 min, followed by centrifugation(48,000 x g,10 min). The pellet is redissolved in a minimum amount of Buffer A & applied to a 1.5 x 60 cm column of Sepharose 6B equilibrated with Buffer A. Elution follows with Buffer A. CAT-containing fractions are mixed with 1 vol saturated AMS & centrifuged(48,000 x g,10 min). Pellet, dissolved in 5 mL Buffer A is usable for 3 weeks.

B. RESULTS

1. Examination of the Enzyme Assay The procedure was that of Bergstrom's group(12), modified as needed. Routine assay mixtures were saturated with carnitine & acyl-SCoA. The great expense of the coenzymes limited our study to palmitoyl & lauroyl-SCoA, chosen as long & short chain acyl-SCoAs for comparison with our nonrat enzyme studies(Section IC). We obtained reproducible data, adequate for all studies planned. Double reciprocal plots gave useful data.

2. Enzyme Purification Efforts The 2.5 day preparation was repeated six times, as just outlined, except that the digitonin step was omitted. This was because a preliminary effort, carried out with digitonin, was not

successful. As Bergstrom's group(12) has indicated that CAT on the inner & outer mitochondrial membranes is kinetically identical, & we obtained purifications within the reported range, this was deemed appropriate. Similar purifications & yields were obtained with fasted or unfasted rats & with 200-400 g rats from AAMRL/TH.

3. PFDA Effect on Acyl Transfer from Palmitoyl-SCoA to Carnitine. This was first examined with 200 μM palmitoyl-SCoA & 110 μM PFDA. Ten experiments showed that the enzyme was activated by 75%(70.2±21% & 73.3±19% for 6 & 8 min time points, where v_o was constant for 8 min). Variation of the [PFDA] between 13.8 & 440 μM showed that effects were similar with all concentrations from 110 μM up(e.g., Figure 1) & that a small effect was obtained with 13.8 μM PFDA(Fig. 2).

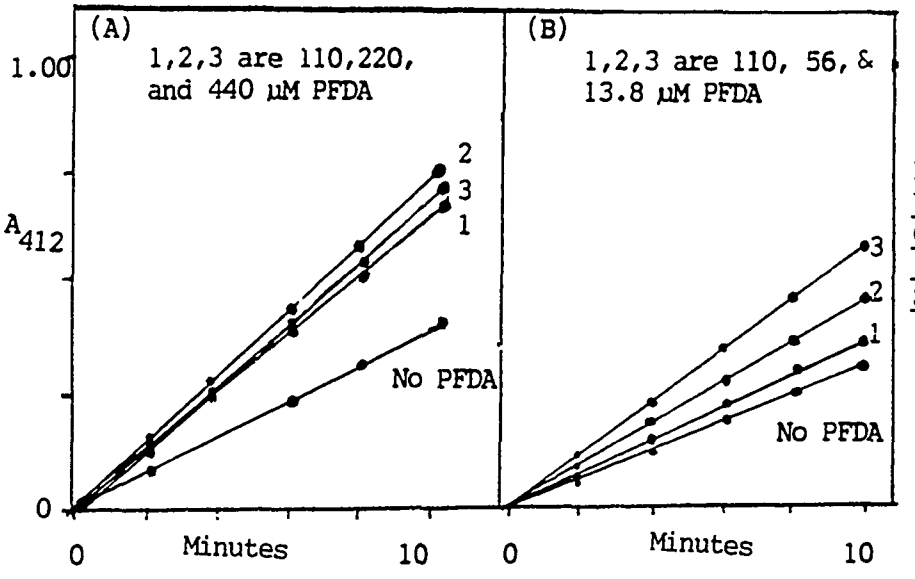


Fig 1. Effect of Varying [PFDA] on CAT Activity With 200 μM Palmitoyl-SCoA. (A) With [PFDA] 110-440 μM (B) With [PFDA] 13.8-110 μM.

4. PFDA Effect on Palmitoyl Transfer with Varied [Palmitoyl-SCoA]. This study used 110 uM PFDA & varied [palmitoyl-SCoA]. Numerous experiments were carried out because the data appeared to be somewhat inconsistant at first. The inconsistency was as follows. In most cases CAT appeared to be turned on to nearly maximum levels by PFDA--regardless of palmitoyl-SCoA levels--until substrate gave out(Figure 2A). CAT samples assayed simultaneously, without PFDA(Figure 2B) gave data that could be used to calculate K_ms. This sugested an effect analagous to the control effect of cholera toxin on translocase. However, a number of experiments with one CAT preparion gave data more like those without PFDA(compare Figures 2A & 2C). The

inconsistancy now appears to be partly resolved, due to experiments with all CAT preparations that have been stored for 3-4 months. This "old" CAT gives data more like Figure 2C. It appears likely that storage modifies CAT, so that it behaves more typically. However, data with PFDA still do not allow estimation of K_ms or K_Is. This effect of PFDA seems likely to be important & it should be reinvestigated with great care, if funds are available.

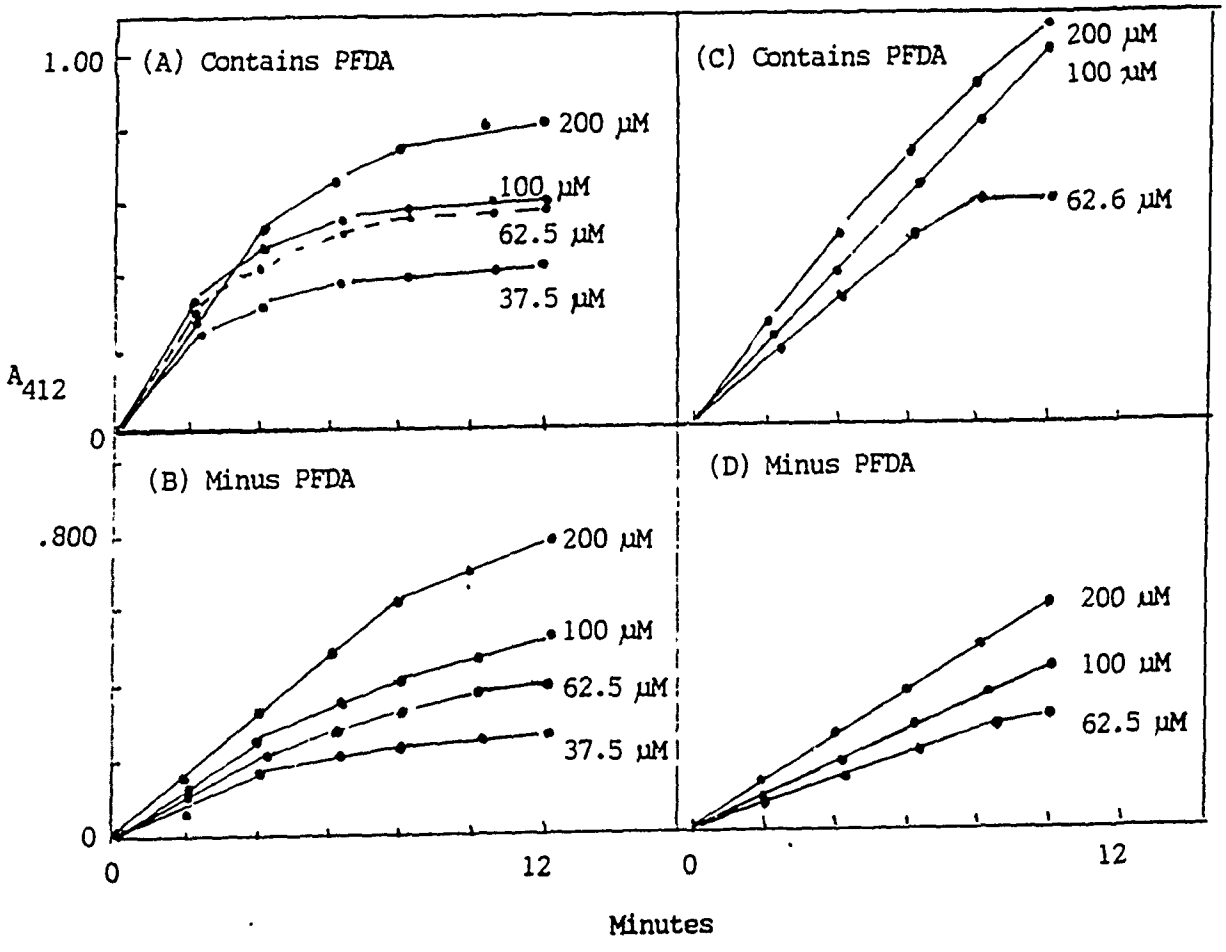


Figure 2. Effect of Varying [Palmitoyl-SCoA] Activation of CAT by 110 µM PFDA. A & B represent most fresh preparations. C & D represent the inconsistant fresh prepararion or 3- to 4-month-old preparations.

3. Comparison of Acyl Transfer from 200 µM Lauroyl- & Palmitoyl -SCoAs to Carnitine. Here, we compared the rate of transfer of lauroyl & palmitoyl groups from the two acyl-SCoAs. Five experiments showed that 200 µM lauroyl SCoA .as a better substrate than 200 µM palmitoyl-SCoA, as the

ratio of lauroyl to palmitoyl transfer was 1.99±22. Figure 3 shows representative data.

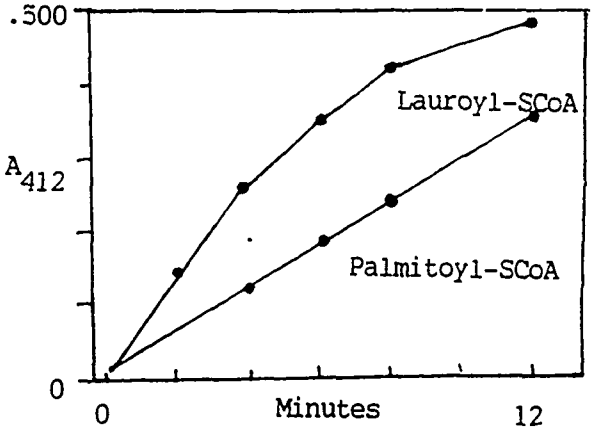


Figure 3. Comparison of Lauroyl & Palmitoyl transfer by CAT.

4. Comparison of the effect of 110 µM PFDA on Lauroyl & Palmitoyl Transfer from 200 µM Acyl-SCoAs via CAT There was a major difference in the effect of PFDA on lauroyl & palmitoyl transfer. Fifteen experiments, all CAT preparations with several different batches of each acyl-SCoA(e.g., Figure 4) indicated that PFDA activated lauroyl transfer less effectively (33±12 %) than palmitoyl transfer(79±16%). However, one batch of lauroyl -SCoA(4 experiments) behaved anomalously, showing inhibition by PFDA(42±9%). It is not clear what the problem here was because the ratio of CAT activity with the two acyl-SCoAs in the absence of PFDA was as described in the preceding section. When we called Sigma about this batch of lauroyl-SCoA, they indicated that it was an very old batch of the acyl-SCoA that "should not have been sold", & replaced it at no cost to us. This enigmatic occurrence led to the assumption that the data may have been due to some breakdown product of the "old" acyl-SCoA, & we disqualified the data. In the future, the acyl-SCoAs from several vendors should be examined to allow absolute identification of the extent of the difference of the PFDA effect on lauroyl & palmitoyl transfer. Figure 4 shows data from a representative experiment.

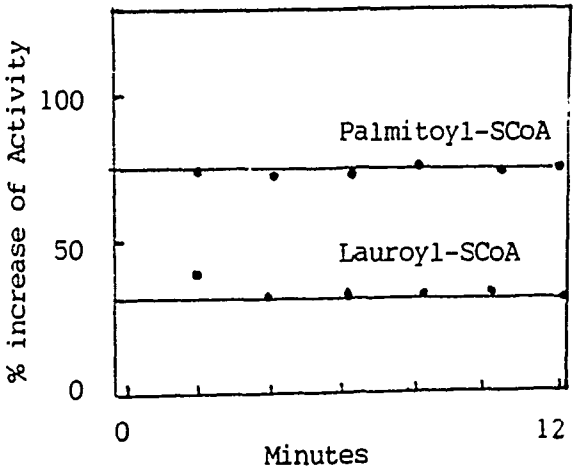


Figure 4. Comparison of the Effect of PFDA on Lauroyl & Palmitoyl Transfer.

5. The PFDA Effect on Lauroyl Transfer with Varied [Lauroyl-SCoA]. This study used 110 μM PFDA & varied [laurcyl-SCoA] between 63 & 200 μM. Seven experiments were carried out. The data(not shown) were similar to those with palmitoyl-SCoA. CAT appeared to be turned on to nearly maximum levels by PFDA --regardless of lauroyl-SCoA levels-- until substrate gave out. Again, CAT samples assayed simultaneously, without PFDA, gave data that could be used to calculate K_ms, while data taken with PFDA were not usable in this way. Experiments with "old" enzyme were not carried out, & should be done in the future.

IV. STUDY OF ACYL-SCoA OXIDASE

A. MATERIALS AND METHODS

Supplies(all from Sigma) The enzyme was used to model whether PFDA affects acyl-SCoA oxidation in a way that could help to explain observed alterations of fatty acid metabolism after PFDA. Palmitoyl-SCoA & lauroyl -SCoA were tested. The assay method is that described by Shimazu et al(13). Supplies needed are palmitoyl-SCoA, lauroyl-SCoA, 4-aminoantipyrine, phenol, horseradish peroxidase, KH_2PO_4, K_2HPO_4. Assay is based on production of a pink quinoneimine dye, as shown next, quantitated by increased 500 nm absorbance.

$$\text{Acyl-SCoA} + \text{E-FAD} \xrightarrow{\text{Acyl-SCoA Oxidase}} \Delta^2\text{-Enoyl-SCoA} + \text{E-FADH}_2$$

$$\text{E-FADH}_2 + O_2 \xrightarrow{\text{Acyl-SCoA Oxidase}} \text{E-FAD} + H_2O_2$$

$$\text{Phenol} + \text{4-Aminoantipyine} \xrightarrow[2H_2O_2 \quad 4H_2O]{\text{Peroxidase}} \text{Pink Dye}$$

Solutions used:

Buffer 1 (50 mM Potassium Phosphate, pH 7.5) Prepare 50 mM KH_2PO_4 & 50 mM K_2HPO_4 in water. Mix aliquots of the 2 solutions as needed to prepare 1 L buffer at the desired pH. Store at 4 C for 3 weeks & discard.

Buffer 2 (50 mM Potassium Phosphate, pH 8.0) Use solutions described for Buffer 1 preparation to make Buffer 2. Handle like Buffer 1.

Color Reagent [0.79 mM 4-aminoantipyrine,11.0 mM phenol, Peroxidase 5 units/mL)] Add 4-aminoantipyrine, phenol & 500 U peroxidase to 100 mL cold buffer 2 & dissolve. Use for up to 2 weeks, stored at 4 C.

Enzyme Diluant, Acyl-SCoAs, PFDA, & Fatty Acid Diluant Solutions are prepared as described in III.

The Enzyme Assay Mix 2.50 mL color reagent with 0.15 mL of acyl-SCoA & .35 mL Buffer 1. Incubate at 25 C for 5 min. Then, transfer .8 mL samples of the mixture to cuvets & add 0.10 mL PFDA(or fatty acid diluant)as appropriate. Then add up to .1 mL enzyme &/or enough enzyme diluant to bring the volume in each cuvet to 1.00 mL & mix. Take 500 nm readings every 2 minutes. The assay proceeds at a constant rate for about 4 minutes.

Basic Enzyme Purification In brief, the 2.5 day reported method(14) used 30 to 60 g liver from male rats(200 g BW), fasted overnigt. It was modified to yield the following method by adding the last step from(15) because the similar last step in(14)did not work in our hands. Chemicals came from Sigma. Solution preparation is not given, for the sake of brevity.

Liver is homogenized in 5 volumes of 50 mM Tris, pH 7.8(10 passes) & centrifuged(17,000 x g,15 min). FAD is added to the supernatant to 25 µM & the solution is immersed in a 60 C bath, with gentle mixing, until 55 C is attained. Stirring is stopped & the solution is kept immersed for 2 more min. Then, it is cooled to 4 C on ice & denatured protein is removed by centrifugation. $(NH_4)_2SO_4$-- AMS --is added to 30% saturation, slowly with mixing. The resultant precipitate(after 5 min mixing) is removed by centrifugation(15,000 x g) & the supernatant is brought to 45% saturation with AMS(7 min) & stirred for 5 min. The precipitate, collected by centrifugation, is dissolved in 5 mL 10 mM K phosphate, pH 7.4(buffer I). This solution is frozen overnight & thawed gently the next morning. It is then applied to a 2 x 50 cm Sepharose 6B column equilibrated with buffer I-FAD & eluted with Buffer I-25 uM FAD. Four-mL fractions are eluted & those containing enzyme are precipitated by adding AMS to 45% saturation. The precipitate is collected by centrifugation & dissolved in 3 mL Buffer I. It is stable at -20 C for 3 months.

B. RESULTS

1. Examination of the Enzyme Assay The procedure was essentially that of Shimazu(13), modified as needed. The routine assay mixtures were saturated with 80 μM acyl-SCoA. The great expense of the coenzymes again limited our study to palmitoyl & lauroyl-SCoA, chosen as long & short chain fatty acyl-SCoAs for comparison with nonrat enzyme studies(see IB). We obtained reproducible data, adequate for all studies planned & double reprocal plots gave appropriate data. The most highly purified AO was diluted with 10 mg/mL bovine serum albumin in Buffer I, before use, to stabilize it. If this was not done the assays gave very erratic results.

2. Enzyme Purification Efforts The 2-day preparation was carried out six times, as indicated in the preceding summary. The second column step replaced the planned cellulose-phosphate column step(from 14) which took a week to carry out in our hands & left us with inactive enzyme. The overall preparation was expected to give a similar purification. We were unable to assay the AO content of the first extract, due to color interference, & could not quantify the total purification. However, initial & final protein content of pools was as expected & the individual chromatography steps gave expected purifications. Therefore, we assumed that our efforts produced AO of expected purity. Future efforts should probably utilize a radioisotope assay. We did not have funds for that purpose.

3. Examination of the AO Activity with Palmitoyl-SCoA or Lauroyl-SCoA The assay method described in IVB1 was very reproducible. Duplicate reaction mixtures differed by only a few %. AO was saturated with 80 μM acyl-SCoA & exhibited zero order kinetics with the enzyme levels tested. Reactions proceeded at a constant rate for 4 minutes, but the enzyme was very unstable & had to be used immediately after storage, even with added bovine serum albumin. Relative activities for 80 μM lauroyl- or palmitoyl-SCoA were 100% or $69.0 \pm 6.5\%$(6 experiments with 3 different AO preparations).

4. Inhibition of 80 uM Palmitoyl-SCoA or Lauroyl-SCoA Oxidation by PFDA Initially we tested 440 μM PFDA. In separate experiments PFDA inhibition of oxidation(6 experiments) was $22.3 \pm 2.3\%$ with palmitoyl-SCoA &

53.2±4.3% with lauroyl-SCoA. So, PFDA appeared to be a better inhibitor of oxidation of the short chain acyl-SCoA. This difference was supported by examination of the ability of varying amounts of PFDA to inhibit oxidation of the acyl-SCoAs. With lauroyl-SCoA, significant inhibition occurred with 36 μM PFDA. This [PFDA] did not inhibit palmitoyl-SCoA oxidation significantly(5 experiments.). Figures 5A & 5B show the effect of varying the [PFDA] on oxidation of palmitoyl- & lauroyl-SCoA, respectively. Fig 5C shows the effect of 36 μM PFDA on the oxidation of both acyl-SCoAs.

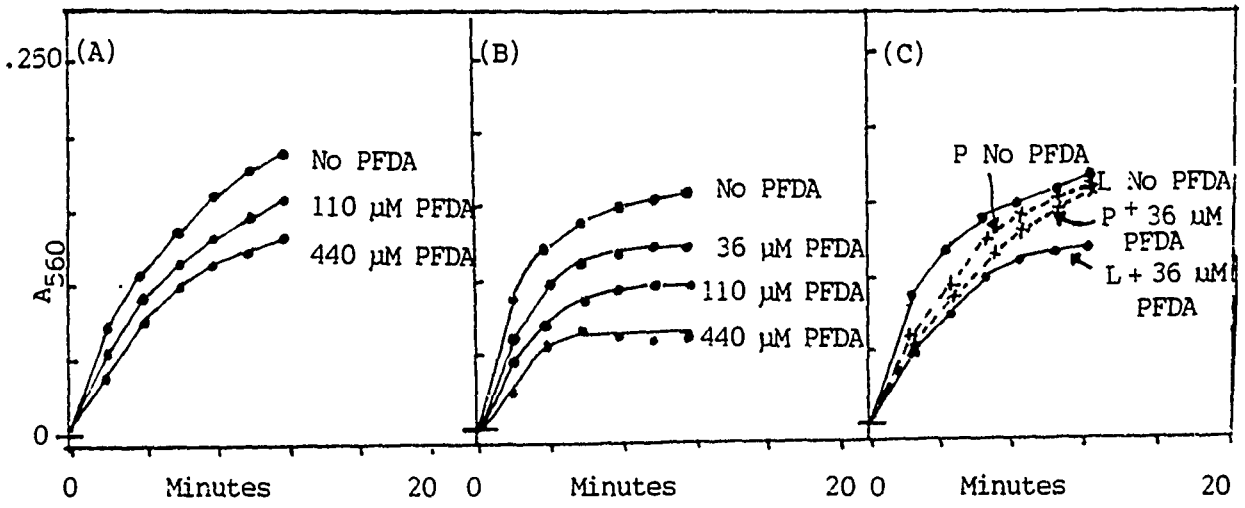


Fig. 5. PFDA Effect on Oxidation of 80 μM Palmitoyl-SCoA & Lauroyl-SCoA by AO. In (A) PFDA content is varied, using 80 μM palmitoyl-SCoA. In (B) the same experiment is carried out with 80 μM lauroyl-SCoA. In (C) the two acyl-SCoA are compared, using 36 μM PFDA.

5. The Nature of PFDA Inhibition of Oxidation of Palmitoyl- & Lauroyl -SCoA We used 110 uM & 55 μM PFDA uM PFDA concentrations with palmitoyl- & lauroyl-SCoA, respectively. These PFDA levels were expected to give 25-35% inhibition with the highest acyl-SCoA levels tested. Double reciprocal plots of palmitoyl-SCoA inhibition studies(e.g., Fig. 6) showed that PFDA was a competitive inhibitor. The K_I obtained was 318±58 μM(six experiments). We were unable to obtain usable data with studies(12 experiments) carried out with lauroyl-SCoA. However the trends observed in supported the effect of PFDA on the short-chain acyl-SCoA as being competive, with a K_I about 1/3 the value for palmitoyl-SCoA. We believe that the problem here due to

the fact that lauroyl-SCoA is more labile than palmitoyl-SCoA. We have found that it breaks down nonenzymatically during the reaction & that the rate of breakdown varies with lauroyl-SCoA concentration. Due to the constraints of the instability of the enzyme(allowing a 4 minute usable kinetic period) it is not possible to carry out the complex experiment required to generate adequate controls & still obtain appropriate kinetics. In the future, it would be valuable to use a radioisotopic assay &/or to modify enzyme-handling methods enough to find a way to prolongze it so the usable kinetic period.

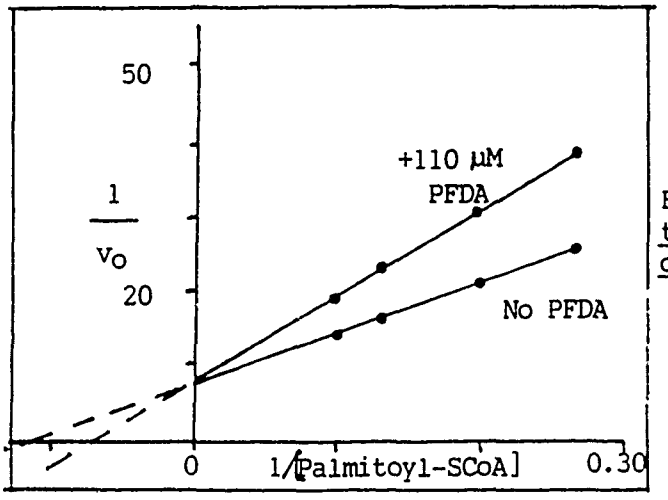


Figure 6. Double Reciprocal Plot of the Kinetics of PFDA Inhibition of of Palmitoyl-SCoA Oxidation.

V. STUDY OF ACYL-SCoA SYNTHETASE

A. MATERIALS AND METHODS:

Supplies(all from Sigma Chemical Company) The enzyme was used to model whether PFDA affects conversion of fatty acids to coenzyme A derivatives & whether PFDA is converted to PFDA-SCoA. Palmitic & lauric acids were tested. The assay is basically that described by Shimizu et al(16). Supplies needed are pyruvate kinase-LDH, myokinase, coenzyme A, NADH, disodium EDTA, Triton X-100, palmitate, laurate, phosphoenolpyruvate, disodium ATP, trizma base, $MgCl_2$. Assay is based upon reading 340 nm decreases due to NADH oxidation as a consequence of:

Fatty Acid + ATP + HSCoA	Synthetase →	Acyl-SCoA + AMP +PPi
ATP + AMP	Myokinase →	2ADP
2ADP + 2PEP	Pyruvate Kinase →	2 ATP + 2 Pyruvate
2 Pyruvate + 2 NADH + 2H+	Lactate Dehydrog. →	2 Lactate + 2 NAD+

Solutions used:

Assay Buffer(200 mM Tris, 20.0 mM $MgCl_2.6H_2O$, 2.00 mM Disodium EDTA, .25% Triton X-100, pH 8.1) Add to .8 vol water: trizma base, $MgCl_2$ disodium EDTA, Triton X-100. Bring to pH 8.1. Dilute to one vol & reneutralize. Store at 4 C and use for one month.

Reagent Buffer(100 mM Tris, pH 7.5) Add Trizma base to .8 vol water & bring to pH 7.5. Dilute to 1 L. Reneutralize. Store at 4 C and use for 1 month.

0.25% Triton X-100 Mix Triton x-100 with water. Store at 4 C and use for one month.

14.5 mM ATP(in 100 mM tris) Weigh 3.00-mg samples of ATP into test tubes. Cover with parafilm & store in freezer. When needed, mix each sample with 1.0 mL reagent buffer. Use for 3 days, stored on ice.

42.8 mM Phosphoenolpyruvate(in 100 mM tris) Weigh 10.0-mg samples of PEP into test tubes & handle like ATP.

Myokinase(100 U/mL in 100 mM tris,0.70 M EDTA) This is a suspension in 3.2 M EDTA. Store as 50 µL samples in parafilm-capped tubes and dilute to 100 U/mL as needed. Use for 4 days, stored on ice.

Pyruvate Kinase/Lactate Dehydrogenase Treat like myokinase, but store as 200 µL aliquots. Dilute each with 800 uL reagent buffer.

5.00 mM Coenzyme A,15.0 mM Dithiothreotol(in 100 mM Tris) Weigh 7.70 mg coenzyme A + 4.35 mg dithiothreotol into test tubes. Handle like ATP, except that 2.00 mL reagent buffer is added to each sample

3.77 mM NADH Store 5.00 mg NADH samples like ATP. Dilute with 2.00 mL water.

1.10 mM Palmitic Acid(in 0.125% Triton X-100-6.57 M propylene glycol Store 3.00 mg samples of palmitic acid in 10-mL tubes at 4 C. As needed, mix a sample with 4.5 mL propylene glycol & vortex well. Then, add 1 vol

0.25% Triton X-100 & mix well. Use for 2 weeks, stored on ice.

The Enzyme Assay: Mix 4.00 mL assay buffer, 0.40 mL NADH, 0.20 mL each ATP, PEP, PK/LDH, myokinase, coenzyme A/dithiothreotol. Incubate at 25 C for 5 min. Then, transfer 1.20 mL samples of mixture to cuvets. Compare cuvets to a water blank at 340 nm in a spectrophotometer at 25 C. Next, add 0.1 mL enzyme diluant to one cuvet & 0.10 mL of enzyme to the other cuvet. After mixing, equilibrate until the initial A 340 nm drop stops(6 min) and add 0.10 mL palmitic acid & 0.10 mL fatty acid diluant each cuvet. Take A 340 nm readings at 2 minute intervals. The reaction rate is constant for 12-18 minutes.

Basic Enzyme Purification In brief, the 4 day reported method(17) used 30 to 40 g liver from male rats(200 g BW), fasted overnigt. All hemicals came from Sigma. Description of solution preparation is not given, for the sake of brevity.

Livers are removed and homogenized in 3 vol .25 M, sucrose pH 7.4(3 strokes of loose-fitting pestle). Homogenate is centrifuged(600 x g, 10 min).The supernatant is recentrifuged(13,000 x g, 20 min). The 13,000 x g supernatant is centrifuged(230,000 x g,1 hr) and the microsomal pellet is saved. It is washed with a ¼ vol .25 M sucrose by resuspension and recentrifugation. Microsomes are suspended in 84 mL 5 mM Triton X-100(TRI), 50 mM K phosphate, pH 7.4, 5 mM mercaptoethanol(MERSH), 1 mM EDTA. After 1 hr the mixture is centrifuged(230,000 x g, 1 h). The supernatant is diluted with 400 mL 2 mM TRI, 5 mM MERSH and applied to a 1 x 18 cm column of Blue Sepharose equilibrated with 20 mM K phospate, pH 7.4, 2 mM TRI, 5 mM MERSH (Buffer A). The column is washed with 5 column vol Buffer A,followed by 5 column vol Buffer A, 10 mM ATP. The column is eluted with Buffer A, 10 mM ATP, .8 M NaCl(5 mL/h). Three-mL fractions are collected. Enzyme elutes as a 25 mL peak. This is mixed with .7 vol 50% polyethylene glycol and stirred for 30 min. The resultant precipitate is collected by centrifugation(23,000 x g, 30 min) and homogenized in 5 mL 20 mM K phosphate, pH 7.4, 5 mM TRI, 5 mM MERSH. It is mixed,30 min, and centrifuged as before. The purified enzyme may be stored for 2 months at -30 C.

B. RESULTS

1. Acyl-SCoA Synthetase Assay With Palmitate as Substrate The method modified the procedure of Shimazu et al(15). Its reproducibility was good, as duplicates differed by only 5%. The assay exhibited zero order kinetics with 146 μM palmitate & 83.5 μM coenzyme A. The reaction rate was constant for 12-18 minutes.

2. Enzyme Purification Efforts The 4-day preparation was carried out six times. Preparations 2-4 provided usable enzyme. The Blue Agarose step yielded inactive enzyme in preparations 5 & 6. We did not carry out a seventh preparation because we saw no reason why a third attempt would succeed; because of budget constraints; & because it was unlikely that additional obtainable data(within the grant period) would add significantly to the information we had obtained. Purification appeared to be as expected. It was not quantitated because the spectrophotometric assay did not work until after microsomes were treated with detergent. Data with enzyme from that point on were in the expected range. Future efforts using an alternate radioisotopic assay should be carried out as the budget allows.

3. PFDA as Substrate for Acyl-SCoA Synthetase Here, 110 μM PFDA, 440 μM PFDA & 146 μM palmitate were compared. The data from four experiments with both PFDA concentrations were similar(e.g., Figure 7, with 440 μM PFDA). The probable maximum amount of PFDA-SCoA that could be expected to be produced was ½ % as much as the amount of palmitoyl-SCoA made with 146 μM palmitate. The very low activity of the enzyme with PFDA was not due to inhibition of the coupled enzymes used to visualize the catalyzed reaction, as shown by our earlier studies(IC).

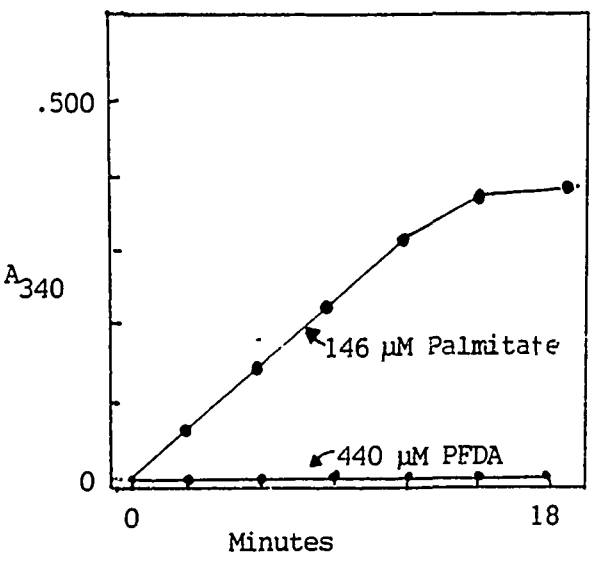


Fig.7 Palmitate & PFDA as Acyl-SCoA Acyl-SCoA Synthetase Substrates

4. PFDA Effects on Formation of Palmitoyl- & Lauroyl-SCoAs. These experiments were not as extensive as planned because the enzyme isolation stopped working after we had carried out only 3 successful preparations. In our effort here, we first compared the ability of 143 μM palmitate and 143 μM laurate to support acyl-SCoA formation. We found that their relative activities(3 experiments) were 100% and 95.1$\pm$5.1%. Next, we tested the ability of 110 μM PFDA to inhibit conversion of each fatty acid to acyl-SCoA. It apperared that the production of palmitoyl- and lauroyl-SCoAs were inhibited 70$\pm$16%(4 experiments) and 88$\pm$20% with laurate(3 experiments). It is regrettable that we were not able to obtain more data, because the rat enzyme appears to be more sensitive to PFDA than the nonmammalian enzyme we studied earlier. It may thus be valuable to continue these studies in the future. Figures 8A and 8B give representative data. The dashed lines shown represent the extent of the inhibition we would have expected with the nonrat enzyme. Also, note that these experiments were careied out at different times. They were chosen for comparison here because similar amounts of lauroyl- or palmitoyl-SCoA were produced in the absence of PFDA and so they are deemed the best comparison we have, presently.

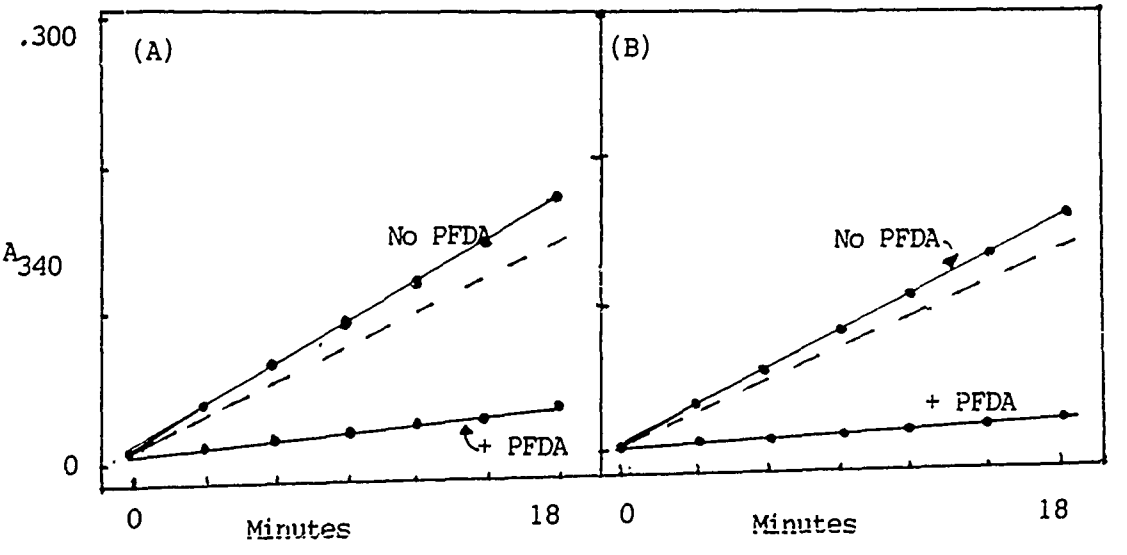


Fig. 8 The Effect of 110 μM PFDA on Formation of Acyl-SCoAs from 146 μM Palmitate(A) or Laurate(B).

5. Kinetic Examination of Inhibition of Palmitoyl-SCoA We examined the effect of 18 μM PFDA(this gave 25-30% inhibition with 146 μM palmitate) on palmitoyl-SCoA formation with differing amounts of palmitate. Six experiments were attempted. The amount of enzyme activity seen with [palmitate] that allowed us to see changes in the inhibition was so low that the data were inconclusive. Not only could we not obtain K_ms here, but at usable palmitate concentrations the enzyme inactivated. If the fifth and sixth enzyme preparations had worked, we would have pursued the study. Such studies will have to await future efforts.

VI. STUDY OF PALMITOYL-SCoA HYDROLASE

A. MATERIALS AND METHODS

Supplies & Assay(all from Sigma) Modeling this enzyme, which releases palmitate and other fatty acids from acyl-SCoAs may help to explain fatty acid buildup due to PFDA. The assay is similar in principle to that of carnitine acyltransferase(**IIIA**). It uses reaction of HSCoA, released by hydrolase action, with DTNB & the yellow product is quantitated at 412 nm. The 1.5 mL assay mixtures(**18**)contain 45 μmol HEPES(pH 7.5) 1.5 μmol EDTA,.45 μmol DTNB, 60 nmol palmitoyl-SCoA, 0.37 mg bovine serum albumin. Reaction mixtures minus enzyme(1.3 mL)are incubated at 25 C & compared to a water blank. Then, enzyme or diluant is added and 412 nm absorbance is read. The procedure--except for reagent concentrations--is very similar to the carnitine palmitoyltransferase assay & will not be discussed further.

Basic Enzyme Purification This method(**18**) uses 5 male rats, fasted 12 hr. Livers are homogenized in 10 vol .25 M sucrose,5 mM HEPES, pH 7.4 (2 strokes with loose fitting pestle).Homogenate is centrifuged(1,500 x g, 10 min)and supernatant is recentrifuged(12,000 x g, 10 min). The mitochondrial pellet is washed with ¼ vol buffer and resuspended in enough buffer to give 40-50 mg protein/mL. Digitonin(60 mg/mL in .25 M sucrose, pH 7.4 is added to give 120 μg digitonin/mg protein, followed by bovine serum albumin to .5 mg/mL. Ten min later 3 vol .25 M sucrose, 5 mM Tris(pH 7.4) .5 mg/mL bovine serum albumin is added. The mixture is centrifuged(10,000 x g, 20 min). The pellet is resuspended in 10 mM HEPES, pH 7.4, sonicated

(20 μM, 30 sec, MSE 150 W Sonifier) and centrifuged(100,000 x g,1 hr). Supernatant is used immediately or stored for up to 30 days frozen before use for additional purification.

B. RESULTS

Examination of the Enzyme Assay and Enzyme Purification The procedure was attempted under a wide variety of conditions that will not be described here. Rats from both our colony and from AAMRL/TH were used. Enzyme assays were carried out with samples from all stages in the preparation. We did not find enzyme activity in any of them. It may be that better results can be obtained in the future with a radioisotopic assay.

VII. EFFORTS AT ORGANIC SYNTHESIS OF PFDA-SCoA

A. MATERIALS AND METHODS

Supplies(from Sigma/Aldrich) Perfluorodecanoic acid(PFDA), n-hydroxy-succinimide, dry ethyl acetate, dicyclohexylcyclocarbodimide, ethanol, thio-glycolic acid, coenzyme A, dry tetrahydrofuran, perchloric acid, acetone,

Synthetic Method This method, reported for synthesis of hexadecanoyl -SCoA (19, 20) and recommended to us for PFDA-SCoA synthesis, was carried out in two steps.

Step 1: Synthesis of the N-hydroxysuccinimide(NHS) ester of PFDA by the Exact Method used for NHS-Hexadexanoate PFDA(3 mmol) is added to 3 mmol NHS in 15 mL dry ethyl acetate. Then, dicyclohexylcarbodiimide(3 mmol) in 5 mL dry ethyl acetate is added & reaction is carried out at 25° C, overnight. Dicyclohexylurea formed is removed by centrifugation & supernatant is concentrated under reduced pressure, to yield white crystals which are recrystallized from ethanol. NHS ester(yield, 91% with hexadecanoate)is identified by thin layer chromatography in chloroform.

Step 2: a. Coenzyme A(50 mg) in 3 mL water is mixed with .5 mmol thioglycolic acid & 2 mmol $NaHCO_3$. To this solution is added 2 mmol of NHS ester in 6 mL tetrahydrofuran(just distilled from $LiALH_4$). The solution is mixed for 4 hrs & 5% perchloroacetic acid(12 mL) is added to precipitate acyl-SCoA and side products. The mixture is then concentrated by flash evaporation,

precipitate is collected by centrifugation, & side products are removed by washing with 25 mL .8% perchloric acid.

b. This is followed by sequential extraction with four 15-mL portions of acetone & three 10-mL portions of ethyl ether.

c. The white residue is then extracted with three 4-mL portions of water, reprecipitated with 6 mL 5% perchloric acid, washed with 10 mL .8% perchloric acid, & then with three 5-mL portions of acetone. The final yield is 50 mg(of hexadecanoyl-SCoA).

B. RESULTS

Preparation of PFDA-SCoA The reaction sequence was attempted four times. Each time Step 1 yielded about ¼ the suggested yield. At the end of Step 2b all of the perchlorate precipitate had redissolved & thus there was no evidence that any PFDA-SCoA had been produced.

VIII. DISCUSSION

A. RESULTS DURING SFRP In rats given 50 mg/kg PFDA, hepatic toxicant levels quickly at 500 μM. Hepatic fatty acid levels(particularly, oleate & palmitate) are elevated quickly. Cell membranes soon become more fluid, less permeable and less fragile. These occurrences are deemed important to PFDA toxicity. As changed liver lipid enzymology could be important to the toxic response mechanism, our first effort(IC) examined PFDA interaction with central, fatty acid-metabolizing enzymes available in purified form, Pseudomonas acyl-SCoA synthetase(**pAS**), Candida acyl-SCoA oxidase(**cAO**) & pigeon muscle carnitine acetyltransferse(**pCT**). Questions asked were: 1) Is PFDA-SCoA made via **pAS**? 2) What is the effect of PFDA on activation of long-chain fatty acids(oleate & palmitate) & shorter-chain fatty acids(laurate & decanoate)? 3) How does PFDA affect action on palmitoyl-, lauroyl- & decanoyl -SCoA of **cAO** & **pCT**? Effects observed provided inferences that made it seem worthwhile to purify the rat enzymes for study. These were:

Acyl-SCoA synthetase(**pAS**) made little PFDA-SCoA slowly, compared to palmitoyl- or decanoyl-SCoA, suggesting that PFDA-SCoA might only be made

slowly by rat **AS**. However, PFDA inhibited palmitoyl-, oleoyl-, lauroyl- & decanoyl-SCoA formation. C-18 acyl-SCoA formation was affected most. The data, extrapolated to rats, suggest that inhibition rat **AS** could lead to elevated palmitate & oleate levels as well as to less long-chain acyl-SCoA available for energy formation &/or membrane lipid synthesis. PFDA is probably a competitive inhibitor of **pAS**, so [PFDA]/[fatty acid] are important (e,g.,[PFDA]/[palmitate]= 4 gave 30% inhibition of palmitoyl-SCoA formation). With 500 uM PFDA in liver, significant inhibition could be expected with up to 300 uM palmitate. The total of the 6 major fatty acid levels in normal liver averages 25 mM**(9)**. Because [PFDA]/[major fatty acids] is thus .02, it would seem that PFDA inhibition of rat **AS** is potential, not actual, if it is like **pAS**. However, **pAS** & rat **AS** may not be similar & rat **AS** should be tested. Examination of rat **AS** will also clarify whether PFD-SCoA, unavailable during my SFRP, is made in rats & the rate at which it is made?

Acyl-SCoA oxidase(**cAO**) mediated oxidation of palmitoyl-, lauroyl- & decanoyl-SCoA was inhibited by PFDA. Inhibition with palmitoyl-SCoA was much weaker than with lauroyl or decanoyl-SCoA. The effect with palmitoyl- & decanoyl-SCoA was competitive inhibition, K_Is $59\overset{+}{-}15$ and $7\overset{+}{-}6.0$ μM. Again, [PFDA]/[substrate] determines the extent of inhibition. Consequently, to acertain the relevance of the PFDA effect *in vivo*, it is necessary to know that acyl-SCoA levels in rat liver are**(21)** 125 uM. As maximum early PFDA levels in liver are about 500 uM, [PFDA]/[long chain acyl-SCoA] is near 4. This could be expected(from our data) to give 40% inhibition of palmitoyl -SCoA oxidation & almost complete inhibition of oxidation of shorter acyl -SCoAs. If rat **AO** is similar, oxidation of palmitoyl-SCoA(& perhaps other long chain acyl-SCoAs)could be less extensive than usual & might diminish energy production from fatty acid oxidation. In addition, oxidation of long -chain fatty acids would cease almost entirely once C-10 or C-12 acyl-SCoAs were produced. Not only would this diminish energy production, but shorter acyl-SCoAs could accumulate & be incorporated into complex lipids more extensively, contributing to observed changes of properties of membranes and other complex lipids.

Carnitine acetyltransferase(**pCT**) mediated transfer of acetyl butyryl, hexanoyl & octanoyl groups from acyl-SCoAs to carnitine was inhibited by

PFDA. A [PFDA]/[acyl-SCoA] of .40, diminished hexanoyl & acetyl transfer by 49.2±5.5% & 43.4±.70%. Inhibition of acetyl transfer was competitive, K_I 111±15 µM, so the basis for acetyl-transfer inhibition by PFDA is [PFDA]/[substrate]. If **pCT** models long-chain carnitine acyltransferases appropriately, then 500 uM PFDA in livers of treated rats would cause significant acyltransferase inhibition even if 2 mM acyl- or acetyl-SCoA were present. Thus, it seems possible that <u>in vivo</u>, PFDA effects in rat liver might act to minimize acetyl & other acyl transfer across mitochondrial membranes & diminish energy production/other important aspects of fatty acid metabolism/utilization. Even, if the noncompetitive inhibition of octanoyl transfer we observed is the rule for longer acyl-SCoAs, its 76.0±28 µM K_I would indicate that PFDA was a potent inhibitor of utilization of acyl-SCoA & have similar consequences.

B. RESULTS DURING THIS GRANT PERIOD The overall goals of the project were to 1) Examine the enzymology of & purify, by methods reported in the literature, three or more rat liver enzymes viewed as important to action of PFDA therein. 2) Test the purified enzymes to identify the effects of PFDA. 3) Examine the ability of rat liver acyl(palmitoyl)-SCoA synthetase to make PFDA-SCoA. 4) Prepare PFDA-SCoA by organic synthesis & examine its effects as done for PFDA, if justified. 5) Plan to examine, other crucial enzymes of fatty acid metabolism in rat liver(e.g., acyl desaturase) & how to use the purified enzymes to assess toxicity of potentially hazardous chemicals.

Our results related to goals 1 - 3 are considered together. There, we achieved success with three enzymes. The data were interesting and allowed us to extend our understanding of the enzymes to rat liver, as follows:

<u>**Carnitine Acyltransferase(CAT)**</u> was purified--essentially as reported in the literature(**IIIA, IIIB2**)--& preparations were examined kinetically. Methodology was designed for optimum use in study PFDA effects on palmitoyl and lauroyl transfer from the acyl-SCoAs to carnitine. The effect of PFDA here was very different from the competitive inhibition of acyl transfer from short chain acyl-SCoAs observed with pigeon muscle carnitine acetyltransferase(**IC3, VIIIA**). With **CAT**, 55 to 440 uMPFDA activated acyl transfer from

either substrate. PFDA was a more effective activator of palmitoyl transfer than of lauroyl transfer. We could not obtain K_ms or activation constants because PFDA turned **CAT** on maximally--regardless of the acyl-SCoA level used as substrate-- & this rate was maintained until substrate gave out(e.g., Figure 2). This effect is reminiscant of the action of cholera toxin on adenyl cyclase(**22**). It is interesting that later reexamination of **CAT** preparations showed that the PFDA effect began to disappear after 3 to 4 months of frozen storage. It is likely due to a conformation of **CAT** that disappears on storage.

The PFDA effect on **CAT** would be expected to occur in rats given toxic doses of the perfluoroacid & may be related to abnormal lipid metabolism seen in those animals if expected differences of the effect on acyl-SCoAs of varying chain length occur. The overall effect here might be unequal transport of acyl groups across the mitochondrial membrane that could uncouple normal energy production and other aspects of lipid metabolism. The effect should be studied in greater depth in the future.

Acyl-SCoA Oxidase(AO) was purified by combination of aspects of two reported literature methods required for success(**IVA, IVB2**). **AO** assay was optimized--withir its limitations--for palmitoyl-SCoA oxidation. **AO** was very similar to the equivalent enzyme from Candida(**IC3, VIIIA**) in substrate preference & the effect of PFDA. Lauroyl-SCoA was a better substrate than palmitoyl-SCOA(100% vs $69.0 \pm 6.5\%$). PFDA inhibited **AO** action on both acyl -SCoAs, as follows. First, 440 µM PFDA inhibited oxidation of 80 µM palmitoyl- & lauroyl-SCoA $22.3 \pm 2.3\%$ & $53.2 \pm 4.3\%$, respectively. Stronger effects of PFDA on lauroyl-SCoA oxidation were supported by observation (Figure 5C) that 36 µM PFDA inhibited oxidation of 80 µM lauroyl-SCoA but not that of 80 µM palmitoyl-SCoA. PFDA was a competitive inhibitor of palmitoyl-SCoA oxidation, $K_I = 318 \pm 58$ µM. We could not obtain K_I values with laurcyl-SCoA. However, trends supported PFDA inhibition as competive, with a K_I about 1/3 that for palmitoyl-SCoA.

Thus, it appears that the [PFDA]/[substrate] determines the extent of inhibition. Because hepatic long-chain acyl-SCoA in rat liver is reportedly 125 µM(**21**) and maximum early PFDA levels in liver are about 500 µM the [PFDA]/[total long chain acyl-SCoA is near 4. This would be expected to

give 40-50% inhibition of palmitoyl-SCoA oxidation and completely inhibit oxidation of shorter-chain acyl-SCoAs(e.g., lauroyl-SCoA). Such effects could lead to diminished oxidation of palmitoyl-SCoA(& perhaps other long chain acyl-SCoAs) and diminished energy production from fatty acids. In addition, oxidation of fatty acids would be expected to almost cease once C-12 acyl-SCoAs were made. Not only would this diminish energy production, but shorter acyl-SCoAs might accumulate and be incorporated into complex lipids more extensively. In turn, such incorporation could contribute to observed changes of properties of membranes and other complex lipids.

Acyl-SCoA Synthetase(AS) was purified by a reported literature method (**VA, VB2**) and its assay was optimized for palmitoyl-SCoA synthesis. The enzyme was similar to that from Pseudomonas(**IC**), except that it was more sensitive to PFDA. Palmitate & laurate were comparable substrates. In contrast, **AS** did converted even less PFDA to PFDA-SCoA with 110 or 440 μM PFDA substrate(Figure 7) than did the bacterial enzyme. However, 110 μM PFDA inhibited palmitoyl- and lauroyl-SCoA production from 143 μM substrate by 70$\pm$16% and 88$\pm$20%, respectively. This inhibition(Figure 8) was much stronger than that observed with the pseudomonas enzyme(**IC, VIIA**). It is unfortunate that technical difficulties prevented us from proceeding with additional planned studies because the greater sensitivity of the rat enzyme to PFDA supported its possible importance in the PFDA effect. This enzyme should be studied more extensively in the future.

Our efforts concerning goals 4 and 5 will be considered next. We attempted to prepare PFDA-SCoA for use in enzyme inhibition studies to paralliel study of the PFDA effect. The method utilized(**19,20**) was recommended to us as the best procedure available. However, we had absolutely no success with it. Our efforts to this end were dropped for several reasons. First, we noted that lauroyl-SCoA was much less stable than palmitoyl-SCoA & this was taken to indicate that PFDA-SCoA was just too unstable to isolate. Second, we found that rat liver acyl-SCoA synthetase did not make an appreciable amount of PFDA-SCoA, as just stated, so planned enzymologic studies with this chemical did not appear to be essential. Finally, conference with another researcher who had attempted the organic synthesis indicated a similar lack of success.

As to goal 5, "planning to examine, other crucial enzymes of fatty acid metabolism in rat liver(e.g., acyl desaturase) & how to use the group of seven or more enzymes to assess toxicity of potentially hazardous chemicals", it is just too soon to do this. However, the studies just carried out support the eventual importance of this goal.

IX. RECOMMENDATIONS FOR ADDITIONAL STUDY -This assumes additional funding by AFOSR.

1. Complete study of rat liver acyl-SCoA synthetase to allow exact identification of its importance.

2. Purify and study rat liver palmitoyl-SCoA dehydrogenase and palmitoyl-SCoA hydrolase, as originally planned.

3. Identify the inhibitory effects of PFDA-SCoA on all of the five enzymes if it is shown to be justified and if PFDA SCoA can be made.

4. Plan to examine, other crucial enzymes of fatty acid metabolism in rat liver(e.g., acyl desaturase) and how to use the group of 7 or more enzymes to assess the toxicity of potentially hazardous chemicals(e.g., CTFE) minimizing in vivo studies.

LITERATURE REFERENCES

(1) Clark, L.C., Jr., Becattini, F., Kaplan, S., Obrock, V., Cohen, D. & Becker, C.(1973) Science 181, 680-682.

(2) Guenther, R.A. & Victor, M.L.(1962) Ind. Eng. Chem. Prod. Res. Dev. 1, 165-169.

(3) Shinoda, K. & Nomura, T.(1980) J. Phys. Chem. 84, 365-369.

(4) Griffith, F.D. & Long, J.E.(1980) Am. Ind. Hyg. Assoc. J. 41, 576-583.

(5) Anderson, M.E., Baskin, G. & Rogers, A.(1981) Toxicologist 1, 16.

(6) Olson, C.T. & Anderson, M.E.(1983) Toxicol. Appl. Pharmacol. 70, 362-372.

(7) Van Rafelghem, M.J., Mattie, D.R., Bruner, R.H. & Anderson, M.E.(1987) Fund. Appl. Toxicol. 9, 522-540.

(8) Olson, C.T., George, M.E. & Anderson, M.E.(1983) Toxicologist 3, 99.

(9) George, M.E. & Anderson, M.E.(1986) Toxicol. Appl. Pharmacol. 85, 169-180.

(10) Harrison, E.H., Lane, J.S., Luking, Van Rafelghem, M.J. & Anderson, M.E. (1988) Lipids 23, 115-119.

(11) Zubay, G.(1988) Biochemistry, 2nd Edition, p. 611, MacMillan, New York.

(12) Bergstrom, J.D., & Reitz,R.C.(1980) Arch. Biochem. Biophys. 204, 71-79.

(13) Shimizu, S., Yasui, K., Tani, Y. & Yamada, H.(1979) Biochem. Biophys. Res. Commun. 91, 108-115.

(14) Osumi, T., Hashimoto, T., & Ui, N. J. Biochem.(1960) 87, 1735-1746.

(15) Inestrosa, N.C., Bronfman, M., & Leighton, F.(1980) Biochim. Biophys. Res. Commun. 95, 7-12.

(16) Shimazu, S.(1979) Anal. Biochem. 98, 341-345.

(17) Tanaka, T., Hosaka, K., & Numa, S.(1981) Meth. Enzymol. 71, 334-341.

(18) Berge, R.K. & Farstad, M.(1981) Meth. Enzymol. 71, 234-242.

(19) Blecher, M.(1981) Meth. Enzymol. 72, 404-408.

(20) Lapidot, Y., Rappaport, S., & Wolman,Y.(1967) J. Lipid Res. 8,142-145.

(21) Berge, R.K. & Aarsland, A.(1985) Biochim. Biophys. Acta 837, 141-151.

(22) Schulz, G.E., Elzinga, M. Marx, F, & Schirmer, R.H.(1974) Nature 250, 120-123.

FINAL REPORT NUMBER 109
REPORT NOT ACCEPTABLE AT THIS TIME
Dr. Ronna E. Dillon
760-7MG-100

1988-1989 USAF-UES Faculty Research Initiation Program

Sponsored by the

Air Force OFfice of Scientific Research

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Prepared by: J. Kevin Ford, PhD. and Douglas J. Sego

Academic Rank: Associate Professor

Department and Psychology Department

University: Michigan State University

USAF Researcher: Mark Teachout

Date: 15, September, 1989

Contract No: UES Project S-210-9MG-066

An Investigation of Training Content Validity and Training Efficiency in the Air Force Airmen Basic-In-Residence Training Course

by

J. Kevin Ford, PhD.

and

Douglas Sego

ABSTRACT

Based on the conceptual work of Ford (1988), this research focused on developing methodologies for examining the training content validity and training efficiency of the Aerospace Ground Equipment (AGE) Airmen Basic-in-Residence training program. Three objectives were met in this research project. First, a methodology was developed to identify the training content domain. Results indicated that the domain could be defined in terms of 105 task statements. Second, a methodology was developed and information collected from AGE training instructors on how much emphasis is actually placed on each task trained. Third, the Matching Technique (Ford & Wroten, 1984) was used to identify tasks that might be over- or undertrained. Based on this work, recommendations are given on incorporating performance information to more fully address training efficiency issues.

Acknowledgements

We are grateful to the combined efforts of Universal Energy Systems and the Air Force Office of Scientific Research for providing the opportunity to conduct this research. Various members of the Human Resources Laboratory helped make this research project an enjoyable and productive experience including Captain Marty Pellum and Mark Teachout. Special thanks to Mr. John Predmore and SMGSTG Kevin Watt of the Training Division Branch for the Aerospace Ground Equipment (AGE) ABR training course for giving up their valuable time to provide information for this research project. Thanks also to the instructors in the AGE ABR school who completed surveys regarding the course content. A special thanks to Judy Solecki for data entry and assistance in data analysis.

I. INTRODUCTION

The design of Air Force technical skills training courses is based on a five step Instructional Systems Development (ISD) model. This ISD process which guides the development and revision of technical training programs includes the following steps: (a) analyze system requirements; (b) define education and training requirements; (c) develop objectives and tests; (d) plan, develop, and validate instruction; and (e) conduct and evaluate instruction (AF Manual 50-2, 1979).

The final step in the ISD process specifies that a thorough evaluation of the training program must be conducted. Although evaluation is a critical step in the ISD model for developing a continuously evolving and refining training process, it often receives the least attention.

There are at least five distinct purposes for conducting training evaluation, each corresponding to specific questions that may be asked of a training system (Ford, 1988): (1) is the training course content job relevant?; (2) are certain tasks over- or undertrained?; (3) did the trainees learn the material or meet the training objectives?; (4) are graduates able to perform their jobs?; and (5) does performance during training predict on-the-job performance?. Decision makers may be interested in answering one or more of these evaluation questions. Evaluation systems must be developed that collect and analyze the information needed to address the questions that decision makers are interested in answering. The quality of the information gathered to address each question can have important implications for improving the quality of the training system.

II. OBJECTIVES

The specific objective of this research was to examine content validity (is the training course content relevant?) and training efficiency (are certain tasks over- or undertrained?) for the Air Force Airman Basic-in-Residence (ABR) technical training courses. The Aerospace Ground Equipment (AGE) specialty was used to develop and apply a methodology that addresses these evaluation issues. The AGE career field was chosen for study as it represents a field in which a large number of personnel are trained per year in the ABR course and the course has remained fairly constant over the last five years. This allows for a more detailed examination of the content validity and training efficiency issues.

The conceptual models of Ford (1988) indicate that there are three key requirements that need to be met to address issues of content validity and training efficiency. First, the job and training content domains (e.g., what tasks are performed on the job and what tasks are trained in the technical skills programs) must be defined using the same type of information (e.g., either task statements or knowledge and skill statements). Second, the emphasis that should be placed on a task during training and the emphasis actually placed on each task during training must be identified. Third, the information collected must be of high quality so that it is an accurate reflection of the job and training domains.

An analysis of the Air Force training evaluation system for the ABR courses reveals that the content validity question is answered at the general task level (i.e., the Specialty Training Standards task level) rather than the more detailed task level (i.e., the Occupational Survey Report (OSR) task statement level). The analysis also revealed that there is no systematic collection of training emphasis information in terms of OSR task levels which would allow for the examination of the training efficiency issues.

Based on the conceptual models developed (Ford, 1988) and the analysis of

the Air Force ABR training system (Sego, 1988), a three step research plan was designed. The first step was to develop a methodology to identify OSR task statements that are trained in the AGE ABR course. Based on the information collected in the first step, the second step was to develop a methodology to collect information on how much emphasis is placed on each task trained in the AGE ABR course. Using the information collected from the first two steps, the third step was to link OSR emphasis information with the training emphasis information obtained in this project to determine tasks which may be over- or undertrained. The results for each research step are reported separately. This includes the methodology developed to collect and analyze the information required as well as the results of those analyses.

III. STEP 1: IDENTIFICATION OF TRAINING CONTENT DOMAIN

The first step in the research plan was to identify OSR task statements that are actually trained in the AGE ABR training course. The output from this step is the foundation for the next two steps in the research plan. Therefore, much effort was placed on developing a systematic and comprehensive methodology for accurately identifying the training content domain.

The last OSR for AGE was completed in 1983. The OSR indicated that there were 615 tasks that describe the AGE career field. Not all 615 tasks are trained in the AGE ABR course. Therefore, the first phase of Step 1 was to separate tasks into two categories or lists; those that are likely to be trained in the current AGE ABR course and those that are not trained.

Two sources of information from the Training Extract were used to separate the tasks into these two lists. The first source was Report Number 2 that matches OSR survey data with the AGE Specialty Training Statements. All OSR tasks linked to STS statements, which have a skill level 3 course requirement, were placed into the preliminary "training content" list. The second source was Report Number 6

that reports the percentage of first enlistment airmen who perform each task. Tasks reported to have been performed by more than 30% of the Airmen during their first enlistment and not identified in the linkage of the STS and OSR were added to the preliminary "training content" list. The resulting list of 192 tasks was put into a questionnaire labeled the "Training Content Domain" (see Appendix A). The remaining 423 tasks in the AGE career field were included in a second questionnaire labeled "Additional Job Content Domain" Tasks (see Appendix B).

These two questionnaires were sent to the AGE ABR Training Development Branch (TDB) at Chanute AFB. Two TDB SMEs were chosen because of their experience and knowledge of the entire AGE ABR course. The two SMEs were asked to independently review the tasks from both lists and to determine if each task is currently taught in the AGE ABR course. If the SME considered the task to be taught, he/she also identified in which block(s) the task is taught.

In the AGE ABR course there are 11 instructional blocks. These 11 blocks are placed into 5 general areas of instruction. Instructors are assigned to an area of instruction and train one group of trainees for the entire area. Block information was important for linking OSR task information to the Plan of Instruction as the POI is divided into blocks.

The questionnaire responses were analyzed for inconsistencies between the two SMEs. The researchers went to Chanute AFB to discuss and resolve the inconsistencies with the SMEs. The two SMEs were presented those tasks in which there was disagreement. There were two possible areas of disagreement: (a) was the ta taught or not taught in the ABR course; and (b) if taught, which block(s) of instruction is the task taught.

There were 16 tasks in which the SMEs failed to agree as to whether the task was or was not taught as part of the ABR course. More inconsistency was found in regards to which block the tasks were trained. Nevertheless, most of the

inconsistencies were eliminated when considering the Area of Instruction level rather than the more specific block level of analysis. Given, that instructors teach an entire area to a group of trainees, the appropriate level of analysis is the Area level.

The results from the questionnaire and subsequent consensus judgment meeting with the SMEs revealed that the SMEs deleted 101 tasks from the "Training Content Domain" list. Eight tasks from the "Additional Job Content Domain" list were added to the Training Content Domain list which resulted in a list of 99 tasks that are most likely taught in the AGE ABR course. Disagreements on which block of instruction a task was taught were resolved through an examination of course documents. This process resulted in two sets of information. First, the training content domain was tentatively identified at the OSR task level of specificity in terms of 99 task statements. Second, the list of 99 tasks were separated into the five areas of instruction. This provided a task matrix for each area in the AGE ABR training course.

Instructional Area 1 was found to have the least amount of OSR tasks being trained (N =12), while Area 3 had the greatest number of tasks taught (N =53). Area 2 had 30 tasks assigned, Area 4 had 27 tasks and Area 5 had 19 tasks. Some tasks are taught across multiple areas so that the number of tasks add up to more than 99.

Each AGE instructor received a survey that was appropriate to their area of instruction. In addition to emphasis judgments to be described below, the instructors were asked to independently check off those tasks that they teach in their area of instruction. The instructors were also given the opportunity to write in additional tasks that they teach that were not on the list.

The results of the survey indicated that all tasks were taught in at least one area of instruction by at least one instructor. Six tasks were added based on

instructors input resulting in 105 tasks. While no tasks were deleted from the list, some tasks were found not to be taught in specific areas of instruction as indicated on the survey. For example, in Area 3, OSR task 241 "adjust diesel engine fuel racks" was not taught by any of the instructors who teach Area 3. Thus, while tasks were eliminated from certain Areas of Instruction, they were taught in at least one other area. The survey of instructors validates the judgments of the SMEs from the TDB. A list of the 105 OSR task statements that define the training content domain is presented in Table 1.

IV. STEP 2: IDENTIFICATION OF TRAINING EMPHASIS

The first step of the research revealed that 99 of the 615 OSR tasks considered relevant to the AGE career field are trained in the AGE ABR course. The Area(s) of Instruction that each task is taught was also identified.

The second step of the research project was to develop a methodology to collect information on how much emphasis is actually placed on each task trained in the AGE ABR course. This information can then be compared with the OSR emphasis ratings (how much should each task be emphasized in training) using the Matching Technique (described in Step 3).

Training emphasis information was collected from the AGE ABR training instructors through survey methodology. Five surveys were developed (see Appendix C for copies of the five surveys), one for each Area of Instruction. Of the 80 instructors who received the survey appropriate to their training area, 53 completed useable surveys (66% response rate). The instructors were asked to estimate the total amount of instructor time devoted to each task trained in their area.

The data was coded onto the computer for analysis. To eliminate the effects of differences in perceptions of time between instructors, the data for each task was transformed into a relative time scale for each instructor to insure

a common metric. The relative time estimates for each task were combined across instructors and instructional areas. These relative time estimates were translated into z-scores (Ghiselli, Campbell, & Zedeck, 1981) in order to facilitate comparison with the OSR emphasis judgments as described in Step 3.

Table 1 presents the z-score for the 105 tasks that comprise the training content domain. An examination of the table reveals a large variance in the relative time spent across tasks. For example, Task 108 (Maintain AFTO Forms 244 and AFTO Forms 245) received a high training emphasis while Task 387 (Remove or Install Refrigerant or Equipment Cooler Lines or Fittings) received a relatively low training emphasis.

V. STEP 3: APPLICATION OF THE MATCHING TECHNIQUE

The third phase of the research matched OSR emphasis judgments with the training emphasis time estimates calculated in Step 2. The Matching Technique (Ford & Wroten, 1984) was used to link OSR emphasis judgments with training emphasis information. The matching of OSR emphasis and training emphasis is based on 99 tasks rather than the 105 tasks identified as part of the training content domain. This is due to the fact that the 6 tasks added to the training content domain by instructors (see discussion in Step 2) were not part of the 1983 OSR. Therefore, the additional statements provided by instructors could not be matched with any OSR task emphasis data.

The most recent OSR for AGE (1983) surveyed 75 senior Noncommissioned Officers (NCOs) from different Major Commands (MAJCOM) for ratings of training "emphasis". Training emphasis is defined as the extent to which a task should be taught to first enlistment airmen in the ABR technical training course. The emphasis scale ranges from 0 (no training recommended) to 9 (extremely high training emphasis). The NCO emphasis judgments for the 99 OSR tasks that are part of the training content domain were standardized through z-score transformations.

The z-scores for these emphasis judgments are presented in Table 1.

To use the Matching Technique (Ford & Wroten, 1984) to identify tasks that may be over- or undertrained, the NCO emphasis judgments were directly compared to the instructor training emphasis data. Table 1 presents both types of emphasis information in terms of z-scores.

The comparison of the two emphasis judgments identifies training "hits" and "misses". Training hits refer to those tasks where emphasis received in training approximately reflects training needs. Training misses can involve areas of deficiencies (undertraining) and excesses (overtraining). Training deficiencies are content areas whose high training needs (i.e., should be trained) are not matched by a high degree of emphasis in the training program. Training excesses are tasks that are receiving an excessive amount of emphasis in training relative to their need to be trained.

The Matching Technique is operationalized by plotting the two z-score emphasis judgments for each task. The plot can be easily translated into a correlation coefficient that provides an overall summary of the statistical relationship between the two emphasis judgments. The correlation between the two judgments was found to be significantly greater than zero ($r = .30$; $N = 99$; $p < .01$). While the correlation coefficient indicates some degree of agreement, it also suggests that there may be a number of tasks in which the two emphasis judgments are not in agreement.

To examine these results more closely, the Matching Technique was completed for each of the 99 tasks in the training content domain. Through this process, we identified 53 tasks that may be training "misses" (53%) and 46 tasks (47%) that are training "hits" or matches. Of the training misses, 24 tasks (24%) may be overtrained while 29 tasks (29%) might be undertrained. Some examples of how the Matching Technique is operationalized at the task level are provided in Figures 1

and 2.

Figure 1a, OSR Task 441 ("Adjust pneumatic system clutches or components") presents a miss in which instructor emphasis is much higher than the OSR emphasis judgement. This indicates a possible training excess or overtraining situation. Figure 1b presents a possible undertraining situation as Task 246 ("Adjust Gas Turbine Engine Governors") has a high emphasis score from NCOs but is given a low training emphasis by instructors.

Figures 2a and 2b present two examples of training "hits" or matches. For Task 264 ("Isolate Engine Motor or Generator Mechanical Malfunction"), there is high emphasis scores from both the NCO ratings and the instructor ratings (see Figure 2a). Figure 2b presents a task in which there is low emphasis scores from the NCO and instructor judgments (Task 225; "Remove or Install Cannon Plug Parts").

VI. RECOMMENDATIONS

The three objectives specified for this research project have been met. These include the identification of the AGE training content domain in terms of OSR tasks, the development of a methodology for determining instructor training emphasis and the use of the Matching Technique to identify tasks that may be over- or undertrained. To more fully address the issue of training efficiency (are tasks over- or undertrained?), the next step is to examine how well recent graduates are actually performing the tasks on the job.

Job performance information is critical to determine whether the emphasis devoted to training at the technical school should be changed. As an example, Task 246 is possibly undertrained. If it is found that the task is not being performed well by most graduates on the job, then the need for more training emphasis during the technical training program is indicated. If a task is being performed well by most graduates, this indicates that training deficiency is not

an important issue.

The Matching Technique also identified some tasks as being possibly overtrained. An examination of job performance may indicate that some of these tasks are not performed well on the job. This would indicate a serious problem that is best addressed through other means such as on-the-job training.

Currently, job performance information is not used to examine the over- and undertraining issue. A new methodology called the Job Performance Measurement System (JPMS) developed by the Air Force Human Resources Laboratory (AFHRL) (e.g., see Hedge & Teachout, 1986) provides the type and quality of job performance data needed to examine this issue. The next research step, therefore, is to integrate the JPMS data into the training evaluation system to more fully address the issue of over- and undertraining.

References

Air Force Manual 50-2. (1979, May 25). Instructional System Development. Washington, DC: Department of the Air Force.

Ford, J.K., & Wroten, S.P. (1984). Introducing new methods for conducting training evaluation and for linking training evaluation to program redesign. Personnel Psychology, 37, 651-665.

Ghiselli, E.E., Campbell, J.P., & Zedeck, S. (1981). Measurement Theory for the Behavioral Sciences. San Francisco: W.H. Freeman & Company.

Hedge, J.W., & Teachout, M.S. (1986). Job performance measurement: A systematic program of research and development. Air Force Human Resources Laboratory, Brooks AFB, Texas.

Sego, D.J. (1988). Air force training evaluation system: A case study. Final Report, Air Force Office of Scientific Research, Air Force Human Resources Laboratory, Brooks AFB, Texas.

TABLE 1 TRAINING CONTENT DOMAIN TASK STATEMENTS & EMPHASIS JUDGEMENTS IN Z-SCORES

		INSTRUCTOR EMPHASIS	OSR EMPHASIS
108	FILL OUT AFTO FORMS 244	2.91	.31
141	MAKE ENTRIES ON AFTO FORMS 22	1.39	-1.73
142	MAKE ENTRIES ON AFTO FORMS 349	.38	1.20
143	MAKE ENTRIES ON AFTO FORMS 350	.70	1.11
152	PERFORM AIR COMPRESSOR INSPECTIONS	.17	1.39
153	AIR-CONDITIONER VISUAL INSPECTIONS	-.54	1.16
154	PERFORM GENERATOR INSPECTIONS	.20	1.36
155	PERFORM LOAD BANK INSPECTIONS	-.11	.63
157	PERFORM BOMB LIFT INSPECTIONS	-.62	1.30
160	PERFORM GAS TURBINE COMPRESSOR INSPECTIONS	.16	1.59
161	PERFORM HEATER INSPECTIONS	-.21	1.22
162	HYDRAULIC TEST STAND SERVICE INSPECTIONS	-.51	1.56
171	AIR COMPRESSOR PERIODIC INSPECTIONS	1.03	1.35
172	AIR-CONDITIONER PERIODIC INSPECTIONS	-.08	1.13
173	PERFORM GENERATOR PERIODIC INSPECTIONS	.59	1.35
174	PERFORM LOAD BANK PERIODIC INSPECTIONS	.32	.50
176	PERFORM BOMB LIFT PERIODIC INSPECTIONS	-.05	1.15
179	GAS TURBINE COMPRESSOR PERIODIC INSPECTIONS	1.12	1.51
180	PERFORM HEATER PERIODIC INSPECTIONS	.34	1.25
181	HYDRAULIC TEST STAND PERIODIC INSPECTIONS	-.11	1.29
190	ADJUST CONTRACTOR POINTS	-.60	-.87
197	CLEAN CONTACTOR POINTS	-1.21	-.84
203	ISOLATE MALF. IN ELECTRICAL CIRCUITS	3.78	.78
209	MEASURE RESISTANCE IN ELECTRICAL CIRCUITS	1.41	.14
212	MEASURE VOLTAGES IN ELECTRICAL CIRCUITS	1.31	.60
215	PERFORM ELECTRICAL SYSTEM OPERATIONAL CHECKS	.56	.31
225	REMOVE CANNON PLUG PARTS	-1.23	-.98
226	REMOVE OR INSTALL CANNON PLUGS	-.92	-1.13
236	RESEARCH TO'S, FOR ELEC. MAINT. INSTRUCTIONS	-.04	.58
237	SOLDER ELECTRICAL WIRING	.77	.73
238	SPLICE ELECTRICAL WIRING	-.11	.33
241	ADJUST DIESEL ENGINE FUEL RACKS	.72	-.49
242	ADJUST DIESEL ENGINE GOVERNORS	.20	.78
244	ADJUST ENGINE VALVES	.54	-.51
246	ADJUST GAS TURBINE ENGINE GOVERNORS	-1.14	.69
247	ADJUST MAGNETO OR DISTRIBUTOR POINTS	-.66	.96
248	ADJUST ENGINE FUEL SYSTEM COMPONENTS	.46	.10
249	ADJUST STARTER CLUTCHES	.08	-1.36
250	ADJUST TURBINE ENGINE BLEED AIR SYSTEMS	-.81	.80
251	ADJUST TURBINE ENGINE FUEL SYSTEMS	-.80	.78
252	ADJUST TURBINE ENGINE TEMPERATURE CONTROLS	-1.07	.35
257	CLEAN AND ADJUST SPARK PLUGS	-1.13	-.54
259	CLEAN MAGNETO OR DISTRIBUTOR POINTS	-1.12	.03
263	FABRICATE ENGINE FUEL LINES	-.68	-1.82
264	ISOLATE ENGINE MECHANICAL MALF.	1.28	1.13

268	LOAD TEST GENERATOR SETS	1.55	.94
269	PERFORM COMPRESSION TESTS	1.33	-.21
270	PERFORM ENGINE OPERATIONAL CHECKS	.53	.57
273	PREPARE ENGINES FOR STORAGE	-.92	-1.68
274	REMOVE ENGINES FROM STORAGE	-.93	-1.77
276	REMOVE DIESEL ENGINE FUEL LINES	1.80	-.59
287	INSTALL ENGINE HYDRAULIC GOVERNORS	-1.02	-.85
289	REMOVE OR INSTALL ENGINE INTAKE MANIFOLDS	-1.16	-.34
304	REMOVE MOTOR OR GENERATOR ARMATURES	.31	-2.50
315	INSTALL TURBINE ENGINE ATOMIZER COMPONENTS	.15	-.32
316	REMOVE OR INSTALL TURBINE ENGINE ATOMIZERS	-.05	-.14
317	INSTALL TURBINE ENGINE COMBUSTOR CANS	2.02	.12
322	RESEARCH TO'S ON ENGINES, MOTORS, OR GENERATORS	1.29	.63
328	TIME DIESEL ENGINE INJECTORS	.32	-.30
332	ISOLATE HEATING SYSTEM MALFUNCTIONS	1.36	.45
333	PERFORM CARBON DIOXIDE (CO-2) TESTS	1.21	-.34
334	PERFORM HEATING SYSTEM OPERATIONAL CHECKS	1.16	.21
356	USE TO'S TO PERFORM MAINTENANCE ON HEATERS	-.40	.18
359	CHARGE REFRIGERANT SYSTEMS	.07	.45
360	EVACUATE REFRIGERANT SYSTEMS	-1.19	-.20
361	ISOLATE AIR CONDITIONER MALFUNCTIONS	-.12	.17
362	PUMP DOWN REFRIGERANT SYSTEMS	-.94	-.96
363	PERFORM AIR CONDITIONER LEAKAGE TESTS	-.92	-.42
364	PERFORM AIR CONDITIONER OPERATIONAL CHECKS	.42	-.10
367	PURGE REFRIGERANT SYSTEMS	-1.19	-.98
381	REPLACE REFRIGERANT DRIER STRAINERS	-1.25	-1.68
387	REPLACE REFRIGERANT COOLER LINES OR FITTINGS	-1.26	-2.03
395	RESEARCH TO'S, FOR REFRIG. SYSTEMS MAINT. INSTRUC.	-.07	.11
398	ADJUST HYDRAULIC FILL AND BLEED SYSTEMS	-1.26	.26
399	ADJUST HYDRAULIC HIGH PRESSURE SYSTEM COMPONENTS	-.11	.18
400	ADJUST HYDRAULIC SYSTEM VALVE ASSEMBLIES	-.29	-.25
406	ISOLATE HYDRAULIC SYSTEM MALFUNCTIONS	.15	.97
407	PERFORM HYDRAULIC SYSTEM OPERATIONAL CHECKS	.46	1.09
436	REPLACE SEALS IN HYDRAULIC SYSTEM COMPONENTS	-.71	.01
437	USE TO'S TO PERFORM MAINT. ON HYDRAULIC SYSTEMS	-.38	.92
439	ADJUST BLEED AIR LOAD CONTROL VALVES	-1.13	.36
441	ADJUST AIR COMPRESSOR SYSTEM OR COMPONENTS	.84	-1.20
446	ISOLATE AIR COMPRESSOR SYSTEM MALFUNCTIONS	1.60	-.05
447	PERFORM AGE PNEUMATIC SYSTEM OPERATIONAL CHECKS	1.28	.42
457	INSTALL AIR COMPRESSOR FILTERING SYSTEM COMPONENTS	.62	-1.43
472	USE TO'S TO FIND AGE PNEUMATIC SYS. MAINT. INSTRUC.	-.77	.59
473	ADJUST BRAKE SYSTEMS	-.53	.10
475	ISOLATE BRAKE SYSTEM MALFUNCTIONS	-.76	-.21
477	PACK WHEEL BEARINGS	.76	-.21
478	PAINT, STENCIL, OR MARK AGE	-.86	-1.04
479	PERFORM BRAKE SYSTEM OPERATIONAL CHECKS	-.87	-.91
482	PREPARE AGE FOR PAINTING	-.98	-1.99
483	PURGE FUEL TANKS	-1.30	-1.36
484	REFLECTORIZE AGE	-.99	-1.71
485	REMOVE OR INSTALL AGE BRAKE ASSEMBLIES	-1.19	-.89
486	REMOVE OR INSTALL AGE BRAKE ASSEMBLY COMPONENTS	-.58	-1.03

488	INSTALL AGE TIRE, TUBE, OR WHEEL ASSEMBLIES	-.41	-.25
503	RESEARCH TO'S, FOR AGE ENCLOSURES	-1.02	.39
568	PERFORM HOUSEKEEPING	-.37	-2.29

Note: The task number refers to the number assigned by the 1983 AGE OSR.

TASK 441 ADJUST PNEUMATIC SYSTEM CLUTCHES

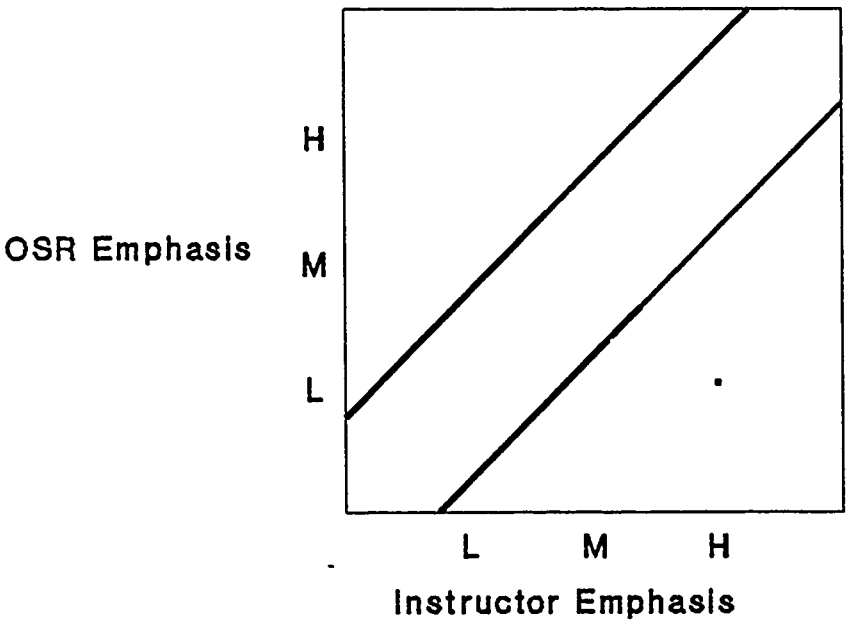


Figure 1a: Possible overtrained task

TASK 246 ADJUST GAS TURBINE ENGINE GOVERNORS

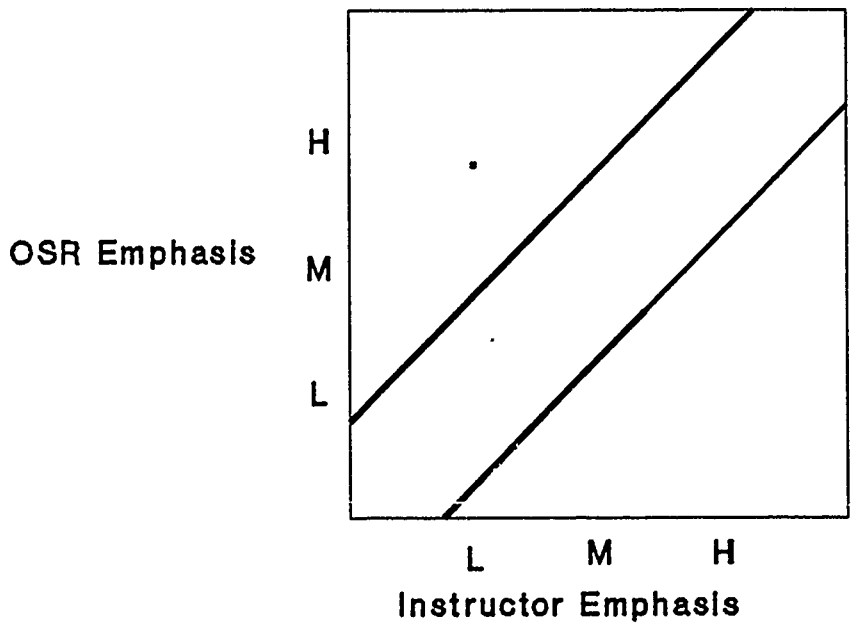


Figure 1b: Possible undertrained task

TASK 264: ISOLATE ENGINE OR MOTOR MECH. MALF.

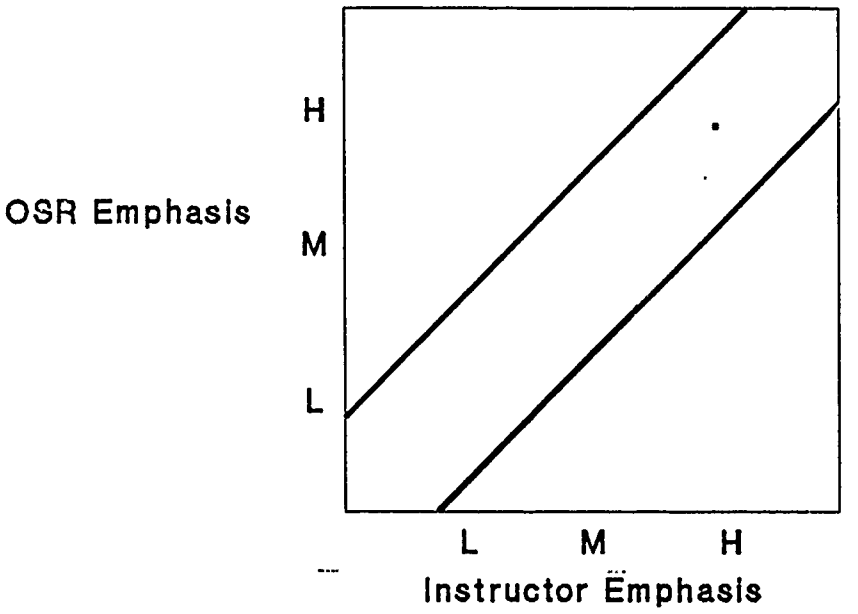


Figure 2a: Training hit, high emphasis

TASK 225: REMOVE OR INSTALL CANNON PLUG PARTS

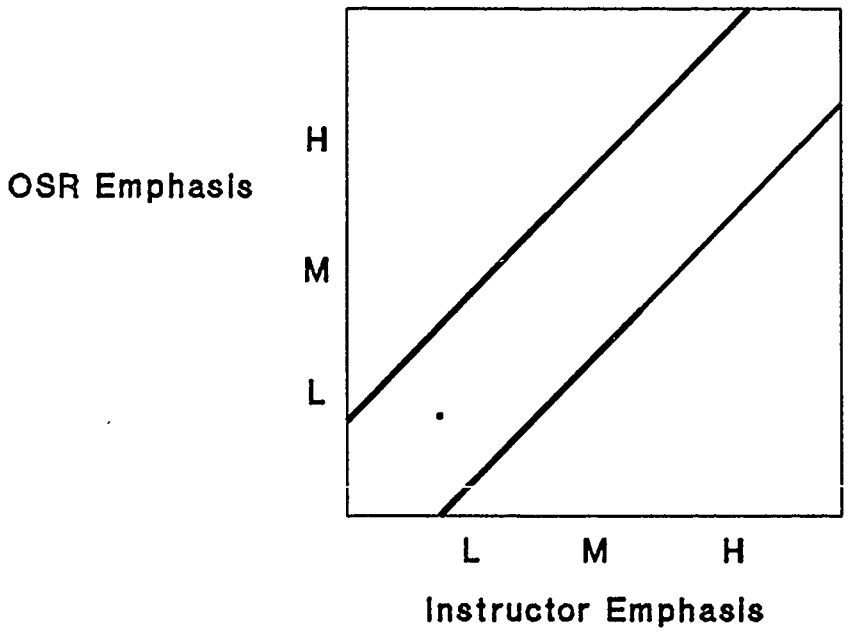


Figure 2b: Training hit, low emphasis

Appendices can be obtained from
Universal Energy Systems, Inc.

Research Initiative Program (RIP)

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

AN INTELLIGENT TOOL TO FACILITATE THE DEVELOPMENT OF QUALITATIVE PROCESS MODELS IN NOVICE PROGRAMMERS

Prepared by: Hugh P. Garraway, Ph.D.

Academic Rank: Associate Professor

Department and University: Computer Science and Statistics Box 5106
University of Southern Mississippi
Hattiesburg, MS 39406-5106

Phone: (601)-266-4949

Bitnet: GARRAWAY@USMCP6.BITNET

(Jan 1 - Dec 31, 1989 UES-AFOSR Mini-grant)

Date: December 26, 1989

AN INTELLIGENT TOOL TO FACILITATE THE DEVELOPMENT OF QUALITATIVE PROCESS MODELS IN NOVICE PROGRAMMERS

by

Hugh Garraway

ABSTRACT

This research resulted in the creation of a 'pop up' expert system-like tool installed within the Turbo Pascal programming environment. The purpose of this tool is to guide inexperienced, beginning programmers to the source of programming problems in the same manner that an expert instructor would. The objectives of this research were to 1) provide a partial solution to the problem of students who have reached an impasse at a time when no expert assistance is available and 2) facilitate the development of students' qualitative process models relative to program development and debugging.

I. INTRODUCTION

Artificial Intelligence and Training

The Air force successfully incorporates computer-based training to provide experiential learning in areas such as flight crew training and electronic systems maintenance. Experience that would be costly or dangerous to provide on real equipment may be gained through interactions with computer-based simulations. This experience helps a student build the structured knowledge, formally called the qualitative process model, necessary to function effectively in a professional role.

On-going research in the field of artificial intelligence (AI) is paving the way to integrate the knowledge of expert instructors into computer-based learning activities, providing students with a tutor or 'coach' to maximize the computer-based learning experience. This integrated component might be referred to as an intelligent tool.

Background

Let's begin with a descriptive scenario which illustrates the need for such an intelligent tool.

A student has just finished the first week of a computer programming course. He has attended all of the lectures, read the assignments, reviewed his notes, planned his algorithm, and hand-written the code to solve the problem given on his first programming assignment. Satisfied that the algorithm for his solution is correct and that the proper syntax has been followed, he begins entering the code into a microcomputer in the programming lab. After typing the code and entering the proper commands to compile the program, the system responds with the message:

Unknown identifier or syntax error.

The student checks to make sure that the variable name pointed to by the system's compiler error indicator has been properly declared. All seems to be in order so he tries to compile the program again and receives the same message. He can find nothing wrong with the program. In fact, he is quite discouraged as he was sure that this program would work. After all, he had carefully planned, entered, and commented it. At 0715 the next morning he is sitting at the door of the instructor's office, program in hand, hoping for some assistance. The instructor, upon arrival, invites the discouraged student in and asks the student a few questions about the program. The instructor thinks for a moment and then says "Are you sure that you closed all of your

comments properly?" After reminding the instructor that the error message and pointer had nothing to do with a comment, the student decides to take the advice and check the comments again. It is then that he notices that a) instead of a } was used to end a comment near the top of the program.

The result was that as far as the compiler was concerned, the comment, which started near the top of the program, continued until the first } encountered which happened to be at the end of the second comment. All of the code between the { and the } had been treated as one long comment by the compiler. The compiler therefore was unaware that the identifiers declared in the unintentionally commented code existed and the error indicated by the compiler did not describe the true source of the bug. The student had not yet developed the cause-effect relationships that might lead him to ask and investigate the question "what type of error could make the compiler 'not see' the declaration of a variable that was (to human eyes) obviously declared?."

The student thanks the instructor and returns to the lab, corrects the program, successfully compiles and commands the system to execute it. Voila! It works. The output matches his calculator-computed result exactly. He runs the program again but this time the result is different and it continues to change each time the program is executed. What is wrong? Could it be a problem with

the system?

In class, as programs are being turned in, the student explains the problem to the instructor who, recognizing the symptoms of a particular type of bug, reminds the class about the importance of initializing accumulator variables and the consequences of not doing so. Upon scanning the program, the student realizes that he failed to provide an initial value for a variable used to collect a sum. In the future, this student will most likely neither create similar bugs nor fail to successfully diagnose and correct bugs of the same type. But unfortunately, the student was not able to get expert advice in time to produce a correct program to submit for grading.

If this expert advice had been available to the student as a component of the lab programming environment then considerable time could have been saved and some of the frustration associated with learning to program could have been avoided.

Basic Scientific Issues In Qualitative Process Modeling

A qualitative process model is the mental model one has regarding cause and effect relationships in a given area of knowledge. This model is developed through hands-on or experiential learning. As a person builds an area of knowledge, for example, a student in

pilot training, declarative knowledge (facts) and procedural knowledge (how to use facts in a step-by-step fashion to produce a desired result) are often acquired through lecture and reading. After a period of rumination, hands-on flight training guided by an instructor provides experiential learning in an aircraft or flight simulator. The flight instructor is available to offer immediate feedback when a student falters, providing an ideal, although expensive, environment for the development of a student's qualitative knowledge of flying.

Students studying electronic systems technologies develop the qualitative process models needed to knowledgeably troubleshoot and repair radar and other electronic systems. Interactions with emerging computer-based training systems offer immediate direction to students as they practice fault diagnosis strategies on simulated systems.

Many of the physical systems and operational parameters for domains such as flight instruction and electronics maintenance training may be modeled in a computer because the design of the object of instruction, an aircraft or radar system, for example, is relatively static.

In domains such as computer programming, however, the student is learning to create an object, a program to perform a specified task. The solution to a programming assignment may be coded

correctly or incorrectly in a number (not necessarily finite) of ways. Since there is no single static model available for programming, the application of computer-based-learning activities for learning to program presents a research and development opportunity.

As with the domains previously discussed, declarative and procedural knowledge for programming may be gained through lecture, reading, and subsequent rumination. The practical application of this knowledge occurs when a student plans, writes, and successfully executes programs. It is the norm for programs not to work perfectly, or at all, on the first try. At this point, a student must examine his declarative and procedural knowledge and essentially trace through his knowledge and source program to establish a cause for a particular symptom or 'bug'. This tracing may include a backward path in the search for a cause or a jump to another path. A level of understanding beyond recalling memorized declarative and procedural knowledge is required to perform this task. The key to this understanding is contained in the student's qualitative process model. Novice programmers or programming students, who have yet to develop a valid qualitative process model for programming, often have difficulty completing assigned work without help from an expert. According to Air Force sources, an improvement in training and resulting job performance could be achieved if a means existed to

more efficiently allow students to develop valid qualitative process models.

II. OBJECTIVES OF THE RESEARCH EFFORT

This document presents a proposed design for ITDA (Intelligent Tool for Debugging Assistance) and a functioning operational model based on this design. ITDA was designed to be an intelligent software tool integrated into a programming environment. Some of the expertise of a human expert relative to programming, novice programmer errors, misconceptions, and debugging skills has been 'captured' so that all students will have instant access to a consultant when a programming impasse is met. ITDA incorporates artificial intelligence techniques and increases its knowledge as it interacts with students and, when necessary, a human expert.

Specific goals of this research were to:

1) Address a stated basic research need in the development of intelligent tutoring systems:

   Implement a practical intelligent tool and thus gather knowledge on how the engineering of such a system will proceed.

2) Provide an application of research that will:

a. Aid in reducing a teaching/learning bottleneck that often causes introductory computer programming courses to be less effective.

b. Function as a job aid to instructors.

c. Serve as an intelligent, embedded learning aid to practicing programmers.

A Practical Intelligent Tool. According to Anderson (1988), most of the work done with intelligent tutoring systems has had the status of basic research, more concerned with gathering basic knowledge than providing useful learning experiences. "The point has been reached," he states, "where a few applications are feasible and it might be worthwhile to pursue some of them both for the relatively immediate benefit and for some sense of how the engineering of these projects will progress."

ITDA addresses this issue in that one goal of this research was to create a practical intelligent tool to fill an existing need.

Teaching/Learning Bottleneck. The bottleneck occurs when a student, who has reached an impasse with a programming problem, must wait until an appointment with the instructor or teaching assistant can be arranged to obtain expert assistance. Even when expert assistance or perceived expert assistance (often in the form of the upperclassman) is available, another problem may arise, the 'quick fix'. The quick fix may occur when the expert simply corrects the problem for the novice without imparting the knowledge of how the problem was corrected or possibly of more importance, how the problem was diagnosed. The quick fix can also introduce confusion and frustration when the expert uses a programming construct not yet intended by the instructor to be a part of the student's knowledge. In an optimal learning environment for programming, quality expert assistance should be available on request to guide the student to the solution of a problem.

Instructor's Aid. ITDA should reduce the number of student/instructor consultations caused by problems within ITDA's knowledge base. ITDA will also point the instructor in the direction of a bug that presently eludes a correct diagnosis.

Rationale

Although out-of-lab instruction attempts to provide students with

the declarative and procedural knowledge necessary for creating correct computer programs, the experiential learning that takes place in the lab plays an important role in the development of a student's qualitative process model. Conceptual and mechanical errors often present major problems for novice programmers (students) who are trying to complete programming assignments. Novice programming students may not have yet developed the debugging skills which evolve as a programmer's qualitative process model for programming grows.

There seems to be an ironic recursive relationship between a student's qualitative process model of programming and his programming/debugging skills. A definition of this relationship might be:

> In order to successfully write programs, i.e. gain experience, a student must have a valid qualitative process model for programming.
>
> In order to have an appropriate qualitative process model, i.e., know how to program, a student must successfully write and maintain programs.

Thus a student's qualitative process model for programming evolves through the experience of successfully writing and maintaining

programs which requires a pre-existing qualitative process model.

This relationship may be the root of the learning problems that so often bring even the brightest students to their knees in the early stages of learning to program. When this problem arises, motivated students will seek a consultation with an expert, in many cases the instructor.

During a consultation, the student's responses to the expert's questions may identify a misconception held by the student or suggest the cause of an identifiable mechanical, typographic, or syntax error. Through dialog with the student, the instructor is often able to coach the student to discover the source of the bug as well as the logical path its location. If a misconception is determined to be at fault, then a direct tutoring or clarification activity can be initiated by the instructor. The student, through this guided diagnosis and resolution of the problem begins to build a valid causal relationship between the symptoms of a particular bug and possible sources of the problem thus adding to the student's qualitative process model of programming. In a sense, the expert has given a portion of his qualitative process model to the student.

A problem exists in this system of learning in that genuine expert advice is not always available to students as they begin to

develop their own expertise. In fact, some 'perceived expertise' can prolong a student's frustration and confusion. If it were possible for the instructor to be available to all students at all times in all labs and at all personal workstations, then the level of students' experiential learning might be expected to increase. It is impossible for a human expert instructor to assume this omnipresent role, and there lies the major implications for this research and development proposal.

Previous efforts at using a software tool to help debug programs has taken the approach of looking at a student's buggy program to identify problem areas for tutoring. Although ITDA will have tools to look at a student's program, its first course of action will be to interact with the student in an effort to identify bug or misconception types. A metaphor for ITDA's interaction with a student might be the dialog that takes place between a physician or physician's assistant and a patient previous to a physical examination.

Before describing ITDA in any greater detail, it is appropriate to briefly describe some of the related research and comment on its relevance to the ITDA concept.

III. REVIEW OF PUBLISHED RESEARCH

Several researchers have studied problems encountered by novice programmers and others have developed intelligent tutoring systems (ITS) or intelligent tools for programming and other areas relative to the concept for ITDA. Naturally, the results of this research will influence the development of ITDA.

Brown, Burton, and De Kleer's (1982) work with tutors (SOPHIE I, II & III)) for electronic troubleshooting provides a foundation of research and application in tutoring diagnostic skills for systems that can be represented within a tutor. Burton (1988) states:

> "It is important that these new systems be built on effective environments, that is, ones that present relevant problems and provide pedagogically appropriate tools. ... The environment in many ways defines the way the student looks at the problem. ... Empowering environments that make explicit the process the student has to do should be developed and their use explored."

MENO II (Soloway,1982) and PROUST (Johnson and Soloway, 1985) are bug-finding programs developed with support from the Office of Naval Research. Both MENO II and PROUST compare a student's program with a library of bugs and misconceptions associated with

a programming assignment dealing with rainfall. An inference is then made by MENO II or PROUST regarding the suspected underlying misconception and a report in the form of comments is generated for the student. In MENO II An attempt was made to use context-independent bug templates, but according to the developers the effort was less than successful. PROUST has been given the ability to access knowledge bases for several programming assignments and thus may be used on more than one assignment although its application is limited to the programs it "knows."

GUIDON (Clancey, 1983) was designed to teach diagnostic problem-solving to medical students. The subject material for GUIDON is the rule base for MYCIN (Shortliffe, 1976), a medical consultation system for diagnosing infectious diseases. Although diagnostic rules could be produced for particular cases, they were not presented in the hierarchical, top-down order as would generally be followed by an expert. This was a function of the unstructured set of production rules within MYCIN.

Anderson's (1985) LISP tutor is described as approaching the effectiveness of a human tutor for teaching the LISP language. It provides a structure editor which produces construct templates that are completed by the student. The tutor provides assistance when a student has problems in correctly completing functions. Immediate feedback is given to the student as each element of a

program is entered so that only a correct program can be constructed. As a program is being developed, the tutor can provide examples of correct code for the student.

Perkins (1986) classifies programming students into groups of stoppers, movers, and extreme movers. The stoppers simply give up and disengage from the problem when a programming impasse is met. The movers think and try new ideas and sometimes break the impasse and carry on to completion. The extreme movers seem to experiment without thinking. Sometimes the extreme mover will find a correct solution but often a logical path to the solution will be overlooked or possibly be prematurely abandoned. Perkins observed that some students who initially tended to disengage from a problem proved capable of solving the problem when encouraged by a researcher.

Pea (1986) has classified language-independent-conceptual bugs based on his observation of novice programmers. He has labeled three major categories of conceptual bugs as Parallelism Bugs, Intentionality Bugs, and Egocentrism Bugs.

Relevance of Published Research to the Development of ITDA

The success of Brown, Burton, and De Kleer's work in providing reactive environments for experiential learning suggests further

development of tutors or intelligent tools to provide similar environments for other types of labs where it may not be desirable or possible to model a static system such as an electronic device. This could be the case with a tutor or tool, such as ITDA, to assist in the development of debugging skills.

Soloway, in his concluding remarks about MENO II, briefly discusses the role of the human program consultant (the expert instructor) and the possibility of incorporating the expertise and dialog management of the expert into future intelligent tutoring systems.

Clancey's research with GUIDON points out the necessity of providing a data structure for rule representation that contains not only the discreet rules but the expert diagnostic paths through the rules.

The reported success of Anderson's LISP tutor supports further research and practical application of artificial intelligence to the area of learning to program.

ITDA's presence should serve as encouragement to Perkins' potential stoppers and keep them moving. Extreme movers might benefit from the direction offered by ITDA since it will keep them "on track" in a diagnostic path. In a system such as ITDA,

identifying a misconception as a member of a group, as defined by Pea, might serve as the top level starting point for a dialog with the student in which subsequent bug subset group types could be identified.

Programming Environment

Programming environments have in the past required significant cognitive overhead for students who have had to wrestle with complex system specific commands and directives to enter, compile, link, and run programs (duBoulay, 1986). To many students, learning the system "incantations" caused at least as much stress as learning to program. Today, products such as Borland's Turbo compilers have reduced this overhead as integrated program development environments combining editors, compilers, linkers, and pop-up help screens are commonly used in teaching programming. These environments provide novice and experienced programmers with easy to learn and use, efficient tools for entering, testing, and experimenting with programs.

Assuming that students have been challenged with a problem to solve, the addition of an immediately accessible expert (ITDA) to such an environment should bring the lab programming environment closer to Burton's vision. By integrating ITDA into an existing successful programming environment, its effect as a component of

an instructional system can be observed and fine-tuned. This will allow ITDA to not only fulfill a practical role, but also provide more basic research for future inclusion of such components in intelligent tutoring systems.

IV. DESIGN AND IMPLEMENTATION OF ITDA

ITDA is an intelligent tool designed to facilitate the development of qualitative process models in novice student programmers. Specifically, ITDA is designed to enhance the learning environment of novice student programmers by providing instant access to the knowledge and collected diagnostic experiences of an expert programmer/debugger. This expert knowledge is made available to students through a program using artificial intelligence techniques to manage the application of several diagnostic tools to students and their buggy programs. Students 'consult' with ITDA in a manner similar to a consultation with a live expert.

Although ADA is becoming the main programming language used by Department of Defense contractors and Air Force programming efforts, for the purpose of this research, ITDA was developed to assist novice programmers learning Pascal using Turbo Pascal. The reasons for this approach are:

1) A standardized programming environment for ADA is still in development and it will be some time before it is available.

2) Turbo Pascal offers a "tried and true" environment for program development.

3) Now that ITDA is operational, it can be refined, further developed, and applied to other language environments.

ITDA exists as a collection of programs stored on the Turbo Pascal system disk used by the student. ITDA is available to a student through a simple series of keystrokes (ALT-D).

ITDA's pedagogical strategy is to engage the student in a menu-based natural language dialog to determine the overall class of advice needed. The student's responses during the dialog help ITDA select the proper decision tree to traverse in an effort to lead the student as close to the problem as possible. Each decision tree consists of nodes in which the student is guided by ITDA to make specific observations in the program (results, error messages, etc.) and respond accordingly. If deemed appropriate, ITDA can execute one of its scanning tools to help identify a problem.

When an interaction with ITDA fails to bring the student to the successful resolution of a problem, the student will be directed to consult with the instructor. When a student is interacting with ITDA, a transcript of the dialog will be made and stored on the student's disk. This transcript is available to the instructor when a student has encountered a problem not currently covered by ITDA's expertise. The transcript serves two purposes. First, it helps the instructor quickly identify and eliminate some solution paths. Second, it indicates the need and marks the logical position for a new knowledge node in ITDA's expertise. After the student's problem is resolved, the instructor may execute a command to add a knowledge node to ITDA so that future problems of this nature can be diagnosed under ITDA's direction. ITDA therefore dynamically, and in a heuristic manner, increases its expertise as its experience grows. This constitutes a modification of what Feigenbaum (in Shea, 1983) calls a "graceful failure mode", an instance in which the system reaches the end of a solution path and still fails to successfully diagnose a problem.

ITDA consists of diagnostic, explanatory, and tool modules tied together by a management module. ITDA has access to the student's source program and can detect some possible problems such as the probability of accidentally commented code. Originally, ITDA was to have access to Turbo

Pascal's error pointer but it was decided that it would be of benefit to require the student to make note of the error type and number and supply this data to ITDA. pointer. ITDA's internal structure is designed so that additional tools may be easily added and managed.

Engineering The ITDA Knowledge Base

ITDA begins its existence as a shell containing tools and decision trees to lead students to the diagnosis of problems frequently encountered by novices. The beginning knowledge base for ITDA was created by identifying the major classes of possible bugs, i.e. runtime, syntax, I/O, inconsistent results etc. and incorporating them into a menu-based natural language command module. A selection from this menu will (if an error message type problem is selected) prompt for the specific error number and this identifies the tree to be traversed to guide the student to the problem. To gather the knowledge to build the initial trees, several instructors graciously kept a log of novice programmer errors encountered during office visits during a semester of introductory programming classes at the University of Southern Mississippi. This data was then used to construct the trees associated with each classified error type.

Interface considerations

Successful integration of ITDA into the learning environment demanded that every effort be made to keep this environment simple and easy to use. Ideally, a student should only need to know one additional command to invoke ITDA. Since ITDA will leads the student through a series of diagnostic activities and tool applications, the student's cognitive overhead is not appreciably increased.

The Coding of ITDA

Turbo Pascal was used as the development language for ITDA. ITDA is installed as a terminate and stay resident (TSR) program that executes Turbo Pascal for the student (the student simply types TURBO as with conventional Turbo Pascal). The knowledge base is maintained separately from ITDA. It is hidden from the student in a subdirectory and may be created and added to interactively by the expert instructor. Tool modules may be added as self contained procedures to ITDA so that modification of the original ITDA code is kept to a minimum of one or two lines. The portion of code that allows ITDA to function as a TSR program is actually the main program and even ITDA is a procedure stored as a separate file. This allows a programmer who is not familiar with the system ins and outs necessary for TSR to safely modify ITDA.

Student interaction with ITDA

When a programming impasse is met, the student simply types ALT-D. A menu introducing ITDA is then displayed. The researcher decided that ITDA was not a sensible acronym and has taken the liberty to present ITDA to students with the nickname BUBBA. Why BUBBA? IDTA is in a sense like a big brother standing by to help and ITDA was conceived in Texas where BUBBA is respected as a name of distinction. The student is then presented with the menu to classify the general error type. It is possible for the student to exit ITDA (BUBBA) at this point if the student needs to look more closely at the Turbo error message if one exists. The student then continues (after reentering ITDA if he exited) and through interaction with the menu causes a tree to be selected. The selected tree is binary in structure and is traversed through interaction in a manner similar to the 'Guess the Animal Game' where the computer asks questions that are responded to by Yes or No to determine the animal. The difference is that IDTA asks question that require the student to observe his program or program design and respond accordingly with Y or N. When a tree has been traversed to the end of a limb (a leaf node) ITDA is able to give advice that in many cases will point the student to the actual problem. If the advice is not of help then the student is advised to see the instructor who personally helps the student and then enters a new knowledge node in the appropriate tree to

reflect the 'newly discovered' programmer error cause.

Application of ITDA

The beginning knowledge base for ITDA was gathered during the Fall 1989 semester at the University of Southern Mississippi. It was implemented in a class for the first time in a course for teachers preparing to teach computer science in secondary schools. This class experienced Pascal for only two weeks so the chance to observe ITDA in operation was quite limited. Students who summoned ITDA seemed to have no problems using ITDA although one student commented that some of the dialog questions could be stated more clearly. One student made the comment that 'BUBBA saved my life' late one evening when no live profs were around.

V. RECOMMENDATIONS

It is obvious that ITDA needs to be tested and fine-tuned. The researcher plans to use ITDA in a full programming/problem solving course during the Spring semester of 1990 in an effort to answer the following question.

> Does ITDA improve the programming environment of novice programmers by making expert consultation more readily available?

ITDA's design is such that it should be relatively easy to attach ITDA to other programming languages or environments or perhaps non-programming applications. Anyone interested in further application of ITDA should contact the researcher at the address indicated at the front of this document.

ACKNOWLEDGEMENTS

The researcher would like to thank Lt.Col. Hugh Burns and Major Jim Parlett, who as successive chiefs of the Intelligent Systems Branch of the Human Resources Lab at Brooks AFB were of great help in developing the concept of this research. Also, much thanks to Dr. Wes Regian, Dr. Kurt Steuck, Capt. Kevin Kline, and everyone else associated with the branch. The support and commitment by the USAF Office of Scientific Research and execution of the RIP by Universal Energy Systems is appreciated. A special note of thanks goes to Mr. Danny Carter and Mrs. Mary Dayne Gregg who kept a log of programmer errors to build a knowledge base. Mr. Keith Alphonso deserves special recognition as a master programmer for the work he carried out helping to implement the code for this project.

REFERENCES

Anderson, John R. The Expert Module. in Polson, M.C. & Richardson, J.J. Foundations of INTELLIGENT TUTORING SYSTEMS. Hillsdale, N.J.:Lawrence Erlbaum Associates Publishers, 1988.

Anderson, John R. and Reiser, Brian J. "The Lisp Tutor". BYTE April 1985, 10(4)

Bonar, J. & Soloway, E. Pre-Programming Knowledge: A Major Source of Misconceptions in Novice Programmers. Pittsburgh University Learning Research and Development Center, Pa.: 1985. (ERIC Document Reproduction Service No. ED 258 805)

Brown, J.S., Burton,R.R. & De Kleer, J. Pedagogical, natural language and knowledge engineering techniques in SOPHIE I, II and III. in Sleeman, D. & Brown, J.S. Intelligent Tutoring Systems. New York, N.Y.: Academic Press, Inc., 1982.

Burton, Richard R. The Environment Module of Intelligent Tutoring Systems. in Polson, M.C. & Richardson, J.J. Foundations of INTELLIGENT TUTORING SYSTEMS. Hillsdale, N.J.:Lawrence Erlbaum Associates Publishers, 1988.

Clancey, William J. GUIDON. Technical Report #9. Stanford University, Calif.: 1983. (ERIC Document Reproduction Service No. ED 237 055)

duBoulay, Benedict. Some Difficulties of Learning to Program. Journal of Educational Computing Research, 1986, 2(1)

Johnson, W.L. & Soloway, E. PROUST. BYTE April 1985, 10(4)

Kozlov, Alex. Rethinking Artificial Intelligence. High Technology Business, May, 1988.

Pea, Roy D. Language-Independent Conceptual "Bugs" in Novice Programming. Journal of Educational Computing Research, 1986, 2(1)

Perkins, D.N. and others. Conditions of Learning in Novice Programmers. Journal of Educational Computing Research, 1986, 2(1)

Shea, T. Future Data Bases Will Be Smarter Rather Than Bigger. INFOWORLD, May 23, 1983.

Shortliffe, E. H. Computer-based Medical consultations: MYCIN. New York: 1976. American Elsevier.

Sleeman, D. Intelligent Tutoring Systems: A Review. Stanford Ca.: Stanford University, 1984. (ERIC Document Reproduction Service No. ED 257 450)

Soloway, Elliot and others. MENO-II: An AI-Based Programming Tutor. New Haven, Conn.: Yale University, 1982. (ERIC Document Reproduction Service No. ED 237 054)

Soloway, Elliot & Johnson, Lewis W. PROUST: Knowledge-Based Program Understanding. New Haven, Conn.: Yale University, 1983. (ERIC Document Reproduction Service No. ED 237 055)

1989 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

Final Report

Prepared by:	Douglas E. Jackson
Academic Rank:	Professor
Department and	Mathematical Sciences
University:	Eastern New Mexico University
Research Location:	Eastern New Mexico University
USAF Researcher:	Malcolm J. Ree
Date:	19 Aug 89
Contract No.	F49620-88-C-0053/SB5881-0378

Acknowledgments

I wish to thank the Air Force Office of Scientific Research for sponsoring this work and UES for administrative support.

I would also like to thank Malcolm J. Ree of Brooks AFB in San Antonio Texas for many helpful conversations and for making part of the Air Force database on test scores available.

On the Effect of Range Restriction on Correlation Coefficient Estimation

by

Douglas E. Jackson

ABSTRACT

Suppose it is desired to estimate the correlation coefficient between random variables X and Y in some population P and the only data available is from some population Q, where Q is a proper subset of P. X and Y are defined on P, while X^* and Y^* will denote X and Y restricted to Q. In the summer of 1988 the author wrote a simulation program to study the effect of this restricted sampling on the estimation of correlation coefficients. The work was supported by the SFRP and the results may be found in Jackson(1988).

The Air Force is attempting to implement those selection criteria that optimize their chances of recruiting the most capable individuals. Whenever a new test is suggested it must be evaluated by estimating its correlation with performance criteria and with tests that are currently part of the selection process. The difficulty is that this new test can only be administered to Air Force personnel. That is, people who have already been selected. Air Force personnel constitute the population Q and applicants constitute the population P. It is necessary to use a sample from Q to estimate correlations between tests that are to be used in P. This is called the range restriction problem. The purpose of this paper is to present the results of a study which is a continuation of the work started under the SFRP. It addresses a number of questions related to the range restriction problem. These problems arose during the summer of 1988 during numerous conservations with staff members at Brooks AFB.

I. Introduction

When certain linearity and homoscedasticity conditions are satisfied there is a theorem that shows how $\rho_{X,Y}$ (the correlation in P) may be computed from $\rho_{X\bullet,Y\bullet}$ (the correlation in Q). The result was first demonstrated by Pearson(1903) and then strengthened by Lawley(1943). $r_{X\bullet,Y\bullet}$, which is calculated using a sample from Q, is an estimate of $\rho_{X\bullet,Y\bullet}$, and Pearson's formula may be used to compute an estimate of $\rho_{X,Y}$ by using $r_{X\bullet,Y\bullet}$ in the place of $\rho_{X\bullet,Y\bullet}$ in the formula. This estimate is sometimes called the corrected correlation coefficient or Pearson's correction statistic or simply the correction statistic. The simulation program mentioned above was written to evaluate the correction statistic and it was found to work very well when the joint distribution of all tests is multinormal. The current paper investigates a number of questions related to the correction statistic using a modified version of this program, which runs on an IBM PC, and three related programs, two of which were written to operate on a mainframe computer. This section contains a list of these questions and a statement of Lawley's theorem.

When a new test is a candidate for induction into the enlistment qualification battery or some other qualification system, then a standard F test is performed to decide if the new test adds anything to the prediction capabilities of the system. An obvious question is whether or not restricted sampling might bias this F test. Section II is devoted to this question.

There are certain variables that influence personnel selection that are not part of the test battery, and hence are not included in the calculation of the correction statistic. The reason might be that the variable is not known or that it is difficult to measure. This is referred to as the unknown variable problem. In section III we look at the mathematical reasons that unknown variables degrade the accuracy of the correction statistic, try to measure the magnitude of this degradation by simulation, and discuss one proposed solution.

The Fisher Z-transformation(Z-transform) of the corrected correlation coefficient between X and Y appears to have a normal distribution when X and Y come from a bivariate normal distribution. Whereas it is provable that the Z-transform

of the ordinary sample correlation coefficient($r_{X,Y}$) has a normal distribution, we are only able to give empirical evidence of the analogous result for the corrected statistic. Evidence is also presented that the mean of the Z-transform of the corrected statistic is very close to the Z-transform of $\rho_{X,Y}$ in the multinormal case. In other words, the inverse Z-transform applied to the mean of the Z-transforms of a random sample of correction statistic observations is $\rho_{X,Y}$. These observations would lead to a method of calculating confidence intervals for $\rho_{X,Y}$ if only the variance of the Z-transform of the corrected statistic were known. In order to obtain estimates of this variance a program has been written that runs on a VAX computer. These matters are addressed in sections IV and V.

Up to now all comparisons of the uncorrected versus the corrected statistic has been made for multinormal distributions. The data to which these estimates are applied may not be multinormal or indeed even linear and homoscedastic. In section VI the uncorrected and the corrected statistic are compared on real data. These data were obtained from the Brooks AFB and consist of 3930 test records, where each record has 11 test scores. The 11 test scores are the 10 tests of the Armed Services Vocational Aptitude Battery(ASVAB) and 1 criterion score.

Two other matters should be mentioned here as they were discussed in the proposal for this grant. The paper Jackson(1988) submitted to UES as part of the SFRP program has been modified for internal publication at Brooks AFB. Extensive modifications have been made to render it acceptable for the Brooks publication system. Malcolm J. Ree of Brooks is a co-author of this paper. The other matter is a reference manual for program CORR. This is included here as appendix B, and appendix A presents a general discussion of the program and its internal structure.

It seems appropriate to include in this section a statement of the correction formula and the set of minimal assumptions necessary for its application. The following theorem is due to Lawley(1943). Variables that are part of the selection criteria are called explicit selection variables and all others are called incidental selection variables. P is called the applicant group and Q is called the selected or

the restricted group.

Let X be the p-element vector of explicit selection variables, and Y the n - p element vector of incidental selection variables on the applicant group. Then X^* and Y^* represent the explicit and incidental selection variables on the selected group. Let

$$V = \begin{bmatrix} V_{p,p} & V_{p,n-p} \\ V_{n-p,p} & V_{n-p,n-p} \end{bmatrix}$$

represent the variance-covariance matrix for X* , Y*. The first p rows and columns refer to the components of X^*. So $V_{p,p}$ is the variance-covariance matrix of X^*, $V_{n-p,n-p}$ is the variance-covariance matrix for Y^*, $V_{p,n-p}$ gives the covariances between X^* and Y^*, and $V_{n-p,p}$ is the transpose of $V_{p,n-p}$. In this discussion V refers to selected data and W refers to applicant data. In our application V will be the estimates of the variance-covariance of all tests and it is based on selected data. The restricted population consist of those who were accepted into the organization so we have data on all tests for these people. Let

$$W = \begin{bmatrix} W_{p,p} & W_{p,n-p} \\ W_{n-p,p} & W_{n-p,n-p} \end{bmatrix}$$

be the matrix of variance-covariances for the applicant data. We will estimate $W_{p,p}$ from the data since we have data for the explicit selection variables on all applicants. The $W_{p,n-p}$, $W_{n-p,p}$, and $W_{n-p,n-p}$ are the matrices that we wish to know and will be given to us by the theorem. $W_{n-p,p}$ is, of course, the transpose of $W_{p,n-p}$ so we will just give an expression for $W_{p,n-p}$ when we state the theorem. The following statement of the theorem is taken from Birnbaum,Paulson, and Andrews(1950).

Assumption 1 :(Linearity) For each j the true regression of Y_j on X is linear.

Assumption 2 :(Homoscedasticity) The conditional variance-covariance matrix of Y given X does not depend on X.

Theorem: Under assumptions 1 and 2

$$W_{p,n-p} = W_{p,p}V_{p,p}^{-1}V_{p,n-p} \quad \text{and}$$

$$W_{n-p,n-p} = V_{n-p,n-p} - V_{n-p,p}(V_{p,p}^{-1} - V_{p,p}^{-1}W_{p,p}V_{p,p}^{-1})V_{p,n-p}$$

If sample correlation coefficients are used instead of population parameters in the matrices $W_{p,p}$,$V_{p,p}$,$V_{p,n-p}$,$V_{n-p,p}$ and $V_{n-p,n-p}$ then the entries of the matrices $W_{p,n-p}$, and $W_{n-p,n-p}$ are estimates of the correlations in the applicant population. These estimates are the corrected statistics of interest and program CORR was written to study their sampling distribution. Appendix A contains a general description of this program and appendix B is a reference manual for its use.

II. The F Test

Since the general considerations are not significantly different from the one variable versus two variable model, only this special case is treated. It is assumed here that Y is the criterion variable, X_1 is the only explicit selection variable, and X_2 is a candidate to become one if it increases our ability to predict which individuals will have high Y scores. In the full model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + E_f$$

and in the reduced model

$$Y = \beta_0' + \beta_1' X_1 + E_r.$$

It is assumed in the model that

$$E(E_f|x_1, x_2) = E(E_r|x_1) = 0,$$

$$Var(E_f|x_1, x_2) = \sigma^2_{E_f},$$

and

$$Var(E_r|x_1) = \sigma^2_{E_r}.$$

In other words the mean of E for given X values is zero and the variance of E for given X values does not depend on those X values. It is also assumed that the distribution of E for any given set of X values is normal and is independent of the

distribution of E for any other set of X values. A discussion of the F test may be found in any standard text that covers multiple regression, for example Dunn and Clark(1974).

The null hypothesis for this test is

$$H_0 : \beta_2 = 0.$$

The test statistic is

$$F = \frac{(SSE_{r^*} - SSE_{f^*})/((n-2)-(n-3))}{SSE_{f^*}/(n-3)}$$

where $SSE_{r^*}(SSE_{f^*})$ is the sum of the squares due to error for the reduced(full) model and n is the sample size. The sum of squares due to error is the sum of the squares of the vertical distances between the individual data points and the corresponding points on the best least squares regression line or plane. The $*$ characters indicate that the samples used to calculate these statistics are taken from the restricted population. This is necessary, of course, due to the fact that for the X_2 and Y variables only restricted data are available. Under the assumptions on the full model, plus the null hypothesis, the sampling distribution of F is an F-distribution with 1 numerator degrees of freedom and $n-3$ denominator degrees of freedom. Notice that no assumptions are necessary concerning the distribution of X_1 or X_2. It is only required that E for given X values is normal, and the distributions of E for different fixed values of X are independent. See Chatterjee and Price(1977) for a statement of this result. These conditions are satisfied in the restricted population if they are satisfied in the applicant population. However, we would like to show in addition that applying the correction statistic to the F calculation is an identity operation. It has no effect at all. But first we need to explain what is meant be applying the correction formula to the F calculation.

The F statistic may easily be written in terms of multiple correlation coefficients, and thus it seemed appropriate to consider the application of the correction formula to these correlations. As we shall see, the correction formula works well on the multiple correlation coefficients but has no effect on the value of F.

Now define

$$\hat{Y}_{12} = \beta_0 + \beta_1 X_1 + \beta_2 X_2,$$

$$\hat{Y}_1 = \beta_0' + \beta_1' X_1,$$

$$R_f^2 = \rho_{Y,\hat{Y}_{12}} \text{ and } R_r^2 = \rho_{Y,\hat{Y}_1}.$$

R_f^2 and R_r^2 are called multiple correlation coefficients. Under the assumptions of the full model and the null hypothesis, the conditions of Lawley's theorem are meet and the correction formula may be applied to estimates of $R_{r\cdot}^2$ and $R_{f\cdot}^2$ to obtain estimates of R^2 and R_r^2. A few tests using a simulation program to be discussed shortly sl ıat the correction formula works very well on these multiple correlation coefficients and that poor estimates are obtained if the correction is not used. This is just another example of the classical range restriction problem and the correction formula should be used.

Let $S_{f\cdot}^2[S_{r\cdot}^2]$ be the standard sample statistic for estimating $R_{f\cdot}^2[R_{r\cdot}^2]$. That is, let $S_{f\cdot}^2[S_{r\cdot}^2]$ be the sample correlation coefficient between Y^* and $\hat{Y}_{12}^*$ [Y^* and $\hat{Y}_1^*$]. The reader should realize that in the previous sentence $\hat{Y}_{12}^*[\hat{Y}_1^*]$ is actually defined using the sample estimates for β_0, β_1, and β_2 [β_0',and β_1']. Again $*$ indicates the restricted population.

It is easy to show that

$$(n-1)(1-S_{f\cdot}^2)S_{Y\cdot}^2 = SSE_{f\cdot}$$

and

$$(n-1)(1-S_{r\cdot}^2)S_{Y\cdot}^2 = SSE_{r\cdot}.$$

Then a simple algebraic manipulation gives

$$F = \frac{S_{f\cdot}^2 - S_{r\cdot}^2}{1-S_{f\cdot}^2}(n-3).$$

The claim is that this value of F does not change if the corrected values of $S_{f\cdot}^2$ and $S_{r\cdot}^2$ are used in this formula instead of applying it as written. Now the correction

formula applies to population parameters instead of sample statistics, and hence it must be shown that

$$\frac{R^2_{f\cdot} - R^2_{r\cdot}}{1 - R^2_{f\cdot}} = \frac{R^2_f - R^2_r}{1 - R^2_f}. \tag{1}$$

Our corrections, of course, are calculated using sample estimates, while the correction formula express a relationship holding for population parameters. But any estimates based on this formula will satisfy the same relationships as those holding for population parameters. It follows immediately that in both the full and reduced models the variance of the error terms are the same in the applicant population as in the restricted population. It is well known that $(1 - R^2)\sigma^2_Y = \sigma^2_E$ in both the applicant and the restricted population. In some texts, for example Dunn and Clark(1974), the equation in the last sentence is taken as the definition of the multiple correlation coefficient. Our definition is equivalent. Hence

$$(1 - R^2_f)\sigma^2_Y = \sigma^2_{E_f} = \sigma^2_{E_{f\cdot}} = (1 - R^2_{f\cdot})\sigma^2_Y.$$

and

$$(1 - R^2_r)\sigma^2_Y = \sigma^2_{E_r} = \sigma^2_{E_{r\cdot}} = (1 - R^2_{r\cdot})\sigma^2_Y..$$

But equation 1 follows from these two equations by the same manipulations used above to derive a formula for F in terms of the sample multiple correlation coefficients.

Program CORR was modified to calculate the F statistic in the two versus one variable case. In the following example the program calculated 100 F values, each based on a sample of size 63. Each variable has mean zero and standard deviation one. The three variables are X_1, X_2, and Y with $\rho_{X_1,X_2} = .707$, $\rho_{X_1,Y} = .707$, and $\rho_{X_2,Y} = .5$. The formula for β_2 is

$$\beta_2 = \frac{\rho_{X_2,Y} - \rho_{Y,X_1}\rho_{X_1,X_2}}{1 - \rho^2_{X_1,X_2}} \left(\frac{\sigma_Y}{\sigma_{X_2}}\right)$$

which is zero in this case and so the null hypothesis is satisfied. Four runs were made and for each the mean of the 100 F values is given as well as the number of values exceeding 4.0, which is the .05 critical value for an F distribution

with 1 numerator and 60 denominator degrees of freedom. The expected value of this F distribution is 1.03. The means of the 100 F values in the four runs were .904,1.236,.991, and 1.122. The number of F values 4.0 or larger in the four runs were 5,7,5, and 5. Next four runs were made with the same parameters accept that we restricted the population to those observations having $X_1 \geq 0.67$. Since X_1 is a standard normal random variable this implies a selection ratio of 0.25. For these four runs the means of the 100 F values were 1.177,.981,.997, and .964. The number of F values 4.0 or greater were 6,6,5, and 4. These empirical observations tend to confirm our conclusion that the F statistic is not affected by range restriction.

III. Hidden Variables

Consider the effect of an explicit selection variable that is not included in the calculation of the corrected correlation coefficient using Lawley's theorem. The reason for exclusion of this variable might be that it is unknown to the individual doing the correction, or perhaps it is difficult to measure. Assuming that the hypotheses of Lawley's theorem would be satisfied if all explicit selection variables were included then they most likely will not be satisfied if one is omitted. This effect is easily observable when there are only three variables and so all our examples include just three variables. First it is shown mathematically why the correction formula should be expected to fail and then a number of simulations are presented to give an idea of the magnitude of the inaccuracies caused by hidden variables. The reference manual for CORR in appendix B includes a discussion of how to use this feature of the program.

Consider the model

$$Y = \beta_0 + \beta_1 X_1 + E_1, \tag{1}$$

where X_1 and E_1 are quasi independent. By quasi independent we mean that the mean of E_1 for any given value of X_1 is zero, and the variance of E_1 for any given value of X_1 does not depend on that given value. Under these circumstances it follows immediately from the definition of covariance that

$$cov(X_1, E_1) = 0.$$

Now if the above assumptions hold, and X_1 is the only explicit selection variable, then the correction formula can be used to estimate the correlation between X_1 and Y in the applicant population. Suppose, however, that there is another explicit selection variable, X_2, and that

$$Y = \beta_0' + \beta_1' X_1 + \beta_2' X_2 + E_{12}, \tag{2}$$

where E_{12} is quasi independent of X_1 and X_2. This last statement just means that the mean of E_{12} for given values of X_1 and X_2 is zero and that the variance of E_{12} for given values of X_1 and X_2 does not depend on those values. Again, it follows that

$$cov(X_1, E_{12}) = cov(X_2, E_{12}) = 0.$$

The problem with applying Lawley's theorem to model 1, when X_2 is an explicit selection variable, is that E_1 may not be quasi independent of X_2, which is required by Lawley's theorem since X_1 and X_2 are both explicit selection variables. We are assuming that E_2 is quasi independent of X_1 and X_2 but that does not imply that E_1 is quasi independent of these two random variables. Indeed it can be shown that if model 1 holds then

$$cov(X_2, E_1) = cov(X_2, Y) - \frac{cov(X_1, Y)cov(X_1, X_2)}{\sigma_{X_1}^2}$$

and if this quantity does not happen to be zero then we know that X_2 and E_1 are not quasi independent for that would imply that $cov(X_2, E_1) = 0$. So if model 1 and 2 both hold, as they do for the multinormal distribution, and X_2 is not included in the correction calculation, then the corrected values will most likely be wrong. It is exactly this scenario that is assumed in the following simulations.

The examples presented here were chosen to demonstrate that the effect of variables missing from the correction calculation can be significant. They are not presented as typical examples. The three variables are X_1,Y, and X_2, where each is a standard normal, X_1 is the known explicit selection variable, and X_2 is the hidden explicit selection variable. The correction procedure for one explicit

selection variable is used but in reality the selection criteria are

$$X_1 \geq 0$$

and

$$X_2 \geq 0.$$

In the first example the unrestricted population parameters are $\rho_{X_1,Y} = .5, \rho_{X_1,X_2} = .8$, and $\rho_{X_2,Y} = 0$. $\rho_{X_1,Y}$ is the parameter being estimated. A sample size of $n = 200$ was used. Thus, one observation of the corrected statistic involves generating multinormal observations until 200 have satisfied the selection criteria and then using this data in the correction formula. One run of CORR calculates the corrected statistic 100 times and displays the distribution of these values and several sample statistics including the mean and standard deviation of the 100 values. Two runs produced a mean corrected sample statistic of .677 for the first and .682 for the second. The standard deviation for both runs was .05. The selection ratios were .398 and .402 for the two runs. Recall that the true value of $\rho_{X_1,Y}$ is .5 and hence we clearly have an overestimate of somewhere between 35 and 37 percent.

Now the hidden variable is removed and the two selection criteria are replaced by the one

$$X_1 \geq .25.$$

The value .25 was chosen to produce a selection ratio of .4 so as to be comparable with the previous runs. Two runs produce mean corrected sample statistics of .504 and .509. The standard deviations were .08 for the first run and .09 for the second. The selection ratio was .401 for both runs. Thus it is clear that the significant overestimates of the previous two runs were caused by the hidden variable.

The following is an example showing an underestimate caused by a hidden variable. This time it is assumed that $\rho_{X_1,Y} = .3$, $\rho_{X_1,X_2} = .4$, and $\rho_{X_2,Y} = .8$. X_1 is the known and X_2 is the unknown explicit selection variable and the selection criteria are again

$$X_1 \geq 0$$

and

$$X_2 \geq 0.$$

The sample size is 200 and the parameter being estimated is $\rho_{X_1,Y} = .3$. Two runs were made and the mean corrected statistics were .160 and .174. The standard deviation of the corrected sample statistic was .11 in both cases. This means that the standard deviation of our estimate of the mean corrected statistic, based on 100 repetitions, is approximately .011. Hence there is a significant underestimate in the range of 42 to 47 percent. The selection ratio was .315 for both runs.

Again, the hidden variable is removed by replacing the two selection criteria by

$$X_1 \geq .48.$$

These two runs gave a mean corrected statistic of .297 and .284. The standard deviation of our estimate of the mean corrected statistic was approximately .012 for both runs. The selection ratios were .314 and .316. The underestimate of the previous two runs was clearly caused by the hidden variable.

Significant hidden variable effects are clearly possible. Whether or not significant inaccuracies exist in real applications is not known. The following parameter values were taken from the Air Force data base on test scores. Two cases were considered, one with, and one without a hidden variable just as in the previous examples. This example is less contrived then the last two, which were deliberately chosen to produce dramatic results. The details of the present case are exactly as in all previous cases except that $\rho_{X_1,Y} = .71$, $\rho_{X_2,Y} = .7$, and $\rho_{X_1,X_2} = .83$. The mean corrected statistic for the 100 replications was .677 for one run and .678 for the other. The standard deviation of our estimate of the mean corrected statistic was about .006 for both runs. The selection ratio was .4 both times. This slight downward bias in the estimate of $\rho_{X_1,Y} = .71$ is significant but not severe.

Now the hidden variable is removed just as before by replacing the two selection criteria by one condition on X_1 selected to produce a selection ratio of .4. The two estimates for these runs were .697 and .706 and the standard deviation of our

estimate was about .005 in both runs. The hidden variable caused a very slight downward bias.

Solutions to the hidden variable problem have been proposed, for example, by Gross and McGanney(1987). The model they assume is slightly different from ours. In their model the selection condition is a single inequality stating that a linear combination of the observable explicit selection variables plus an error term is non-negative. The error term plays the role of the hidden variable or variables. In our model the criteria consists of several equations which all must be satisfied and one of these equations involves an unobservable or hidden variable. The Gross McGanney model assumes that

$$Y = \beta_0 + \beta_1 X + E_y$$

and

$$Y_S = \alpha_0 + \alpha' X_S + E_S,$$

where $\alpha' X_S$ denotes a linear combination of the observable explicit selection variables and $Y_S \geq 0$ is the selection condition. Also X and E_y, as well as X_S and E_S, are required to be quasi independent as defined in the last section. Finally, it is assumed that E_y given X and E_S given X_S are both normally distributed. The authors display a number of simulations based on a maximum likelihood estimation of the parameters in the model. They do not have an analytic equation for the maximum likelihood estimators but rather use the Newton-Raphson numerical technique to find the maximum of the likelihood function. This method works well when the conditions of the model are meet and ρ_{Y,Y_S} is not close to zero. When ρ_{Y,Y_S} is close to zero the Pearson correction formula works well. They decide which statistic to use by doing a hypothesis test for $\rho_{Y,Y_S} = 0$. This test, of course, depends on the assumption that the error terms are normal.

There are a number of objections to using this maximum likelihood estimator. Many assumptions are necessary for its use. The selection conditions must all fit into one equation and the error terms must be normal. Both of these assumptions are false for most Air Force testing data. In addition, it is not known for certain

that significant inaccuracies are caused by hidden variables in practice. Pearson correction on the other hand requires no assumption of normality or constraints on the selection criteria.

The assumptions required to use Pearson correction are not known to hold perfectly in the absence of hidden variables. In the simulations presented earlier in this section there is that implicit assumption that they do hold in the absence of hidden variables. If this assumption is true, then the most prudent course would be to identify the hidden variables and include them in the Pearson correction.

It is difficult to imagine any statistic that attempts to correct for hidden variables that does not make substantial assumptions about the form of the selection criteria and the distribution of the error terms. For these reasons it seems that the best strategy is to simply use the Pearson correction formula. This matter is further addressed in section VII.

IV. Discussion of Confidence intervals

The following theorem is well known. See, for example, Brunk(1960).

Theorem: If r is the sample correlation coefficient of a sample of size n from a bivariate normal population, then the statistic

$$Z = \frac{1}{2} \log \frac{1+r}{1-r}$$

is asymptotically normally distributed with

$$E(Z) \doteq \frac{1}{2} \log \frac{1+\rho}{1-\rho}$$

and

$$V(Z) \doteq \frac{1}{n-3},$$

where ρ is the population correlation coefficient being estimated by r and $\doteq$ means "approximately equal to". The approximation appears adequate for most purposes for sample sizes as small as 10.

If the distribution of the Z-transform of the Pearson correction of r is also normal then maybe we have the basis for the construction of interval estimates of

ρ in the applicant population. We tested for normality of the corrected statistic with a chi-square test applied to data generated by CORR. Six cells were used for the χ^2 test. For the first test the sample mean($\overline{X}$) and the sample standard deviation(S) were used. In this case the six cells were defined as $(-\infty, \overline{X} - 2S)$, $(\overline{X} - 2S, \overline{X} - S)$,$(\overline{X} - S, \overline{X})$, $(\overline{X}, \overline{X} + S)$,$(\overline{X} + S, \overline{X} + 2S)$, and $(\overline{X} + 2S, \infty)$. If the Z-transform of the Pearson correction of r is normal then the χ^2 values come from a chi-square distribution with 3 degrees of freedom. For a test at the $\alpha = 0.1$ level of significance, the critical value of χ^2 is 6.25. Based on eight runs with 3,4, and 5 variables and various assumed correlation parameters, it seems likely that the corrected sample correlation coefficient has a distribution that is at least not significantly different from normal. The largest χ^2 observed in the eight runs was 6.17. On the eight runs the χ^2 values were 1.44, 1.67, 6.17, 2.07,0.74, 3.26, 1.32, and 1.61. One more run was made with exactly the same parameters as the run that resulted in $\chi^2 = 6.17$, and for this run $\chi^2 = 2.42$.

There may, of course, be correlation parameters for which the distribution of the Z-transform of the corrected statistic is not normal. Because of this problem, and because the standard deviation of the Z-transform of the corrected statistic is greater than predicted by the theorem, in the next section a procedure based on a simulation program is suggested. But first it is necessary to determine if the mean of the Z-transform of the corrected statistic is as predicted by the theorem, namely

$$\frac{1}{2} \log \frac{1+\rho}{1-\rho}. \tag{1}$$

Eight more runs were made with the same parameters as the previous eight. This time we tested the hypothesis that the distribution of the Z-transform of the corrected statistic is normal with mean given by equation 1, the transform of the population parameter being estimated. Now the number of degrees of freedom of the χ^2 distribution is 4. For a test at the $\alpha = 0.1$ level of significance the critical value is 7.78. On the eight runs the observed χ^2 values were 1.90, 3.58, 1.93, 7.21, 2.41, 2.76, 5.41, and 2.02. For the example that gave $\chi^2 = 7.21$ another simulation was done and this time $\chi^2 = 0.71$.

Based on these observations, and others not presented here, it seems reasonable that in at least two aspects the Z-transform of the corrected statistic behaves as the Z-transform of the sample correlation coefficient. Namely, it is approximately normal, and the mean is the Z-transform of the population parameter being estimated. The parameter being estimated in the case of the sample correlation coefficient is ρ^*, the correlation in the restricted population, and for the corrected statistic it is ρ, the correlation in the applicant population. The difficulty is that the standard deviation of the Z-transform of the corrected statistic is larger than the standard deviation of the Z-transform of the sample correlation coefficient. This last value is approximately $1/\sqrt{n-3}$, as predicted by the theorem. For the eight cases mentioned above, the standard deviation of the Z-transform of the corrected statistic ranged from $1.05/\sqrt{n-3}$ to $1.7/\sqrt{n-3}$. So an interval estimate based on a standard deviation of $1/\sqrt{n-3}$ would be too small.

The standard deviation may depend on all of the input parameters. In the instance of the Air Force testing battery there are 10 explicit selection variables and hence 10 variances and 45 correlation coefficients. We obviously need some way to estimate the standard deviation of the Z-transform of the corrected statistic if we are to obtain an interval estimate for ρ. The only approach that seems feasible at this time is as follows.

Take the observed corrected correlations, our point estimates of the correlations in the applicant population, and do a simulation as if these were the true population parameters. For the sample sizes typically used, for example at Brooks AFB, we have every reason to believe that the standard deviation of the Z-transform of the corrected statistic does not vary greatly with small changes in the population parameters. This is repeatedly observed in our simulation studies. The purpose of the simulation is to estimate the standard deviation of the Z-transform of the corrected statistic, to calculate the chi-square value to test the normality assumption, and to print the 90% and the 95% capture ratios.

Let d represent our estimate of the standard deviation and

$$t = \frac{1}{2} \log \frac{1+\rho}{1-\rho},$$

where ρ is the correlation of interest in the applicant population. Then for each sample correlation coefficient(r) observed in the simulation it is recorded whether or not it is true that

$$-Z_{.05} \leq \frac{\frac{1}{2} \log \frac{1+r}{1-r} - t}{d} \leq Z_{.05}.$$

The fraction of time that this is true is called the 90% capture ratio. The 95% capture ratio is also computed, using $Z_{.025}$ in place of $Z_{.05}$. If the capture ratios are close to .9 and .95 respectively and the chi-square value is small then it is reasonable to compute a 90% or 95% confidence interval based on d. A derivation of this computation follows.

Let $(1-\alpha) * 100$ % be the confidence coefficient of the interval being defined. It is assumed that

$$1 - \alpha = P(-Z_{\alpha/2} \leq \frac{\frac{1}{2} \log \frac{1+r}{1-r} - t}{d} \leq Z_{\alpha/2}).$$

After some algebraic manipulation it is seen that this probability is the same as

$$P(\frac{a-1}{a+1} \leq \rho \leq \frac{b-1}{b+1}) \tag{2}$$

where

$$a = \frac{1+r}{1-r} e^{-2Z_{\alpha/2} d}$$

and

$$b = \frac{1+r}{1-r} e^{2Z_{\alpha/2} d}.$$

Thus equation 2 defines the $(1-\alpha) * 100$ % confidence interval.

As an example of the use of this procedure suppose that there are 9 explicit selection variables called $V_1, V_2, \ldots, V_9$ and one implicit selection variable V_{10}. $\rho_{V_1, V_{10}}$ is to be estimated using a sample of size 100. Use the correction formula to get

point estimates of all 45 correlation coefficients as well as 10 variances. Suppose that the point estimate for $\rho_{V_1,V_{10}}$ is .5. Now plug all of these point estimates into the simulation program using the appropriate selection criteria. Suppose that the standard deviation of the Z-transform of the correction (d) is 0.150. Note that $1/\sqrt{n-3} = .102$, so d is about 50% more than the standard deviation of the Z-transform of $r_{V_1,V_{10}}$. Suppose that the capture ratio of the 90% and 95% confidence intervals are close to .9 and .95 respectively, and that the chi-square value is small. Hence

$$\frac{a-1}{a+1} = \frac{1.83-1}{1.83+1} = 0.29$$

where

$$a = \frac{1+.5}{1-.5}e^{-2(1.645)(.15)} = 1.83$$

and

$$\frac{b-1}{b+1} = \frac{4.914-1}{4.914+1} = 0.66$$

where

$$b = \frac{1+.5}{1-.5}e^{2(1.645)(.15)} = 4.914.$$

So the 90% confidence interval for ρ is $(.29, .66)$.

V. Two Programs for Interval Estimates

Two programs, NORM and TEST, have been written and implemented on the VAX computer. Norm uses a multinormal distribution for the joint distribution of the test scores. The method of generating these multinormal observations is essentially the same as the method in program CORR and this method is presented in Jackson(1988). The only difference is that the selection process has been speeded up somewhat. TEST generates test scores by random sampling from a database of test scores supplied by Brooks AFB. The main purpose of norm is to estimate the standard deviation of the Z-transform of the corrected sample correlation coefficient and to produce 90% and 95% confidence intervals based on the computations of the last section. The main purpose of TEST is to test this procedure using real test data.

The input and output for NORM is just like that described in appendix A and B except that on output more information is given on the Z-transform. Both the 90% and 95% capture ratios are given as well as the chi square value and the 90% and 95% confidence intervals. The meaning of these values and the intended use of program NORM is as follows.

Suppose that table 1 below represents the corrected correlation coefficients based on some sample of test scores for 11 tests, and an interval estimate is desired for the correlation between variables 1 and 11. The point estimate for this correlation is 0.596. The values in table 1 are input to NORM as described in appendix A, and a simulation is done as if these were the actual population parameters. The values in table 1 are the best available estimates of the population parameters. The variable REPS represents the number of times an estimate of the correlation between variables 1 and 11 is calculated by the simulation. That is to say, it is the number of random samples that are generated and for each random sample the sample correlation coefficient and the corrected sample correlation coefficient are calculated from the data of that sample. Suppose we use REPS = 100. For each of the 100 repetitions a 90%(95%) interval is constructed as explained in the previous section. The 90%(95%) capture ratio is the fraction of these 100 90%(95%) confidence intervals that contain the true value of 0.596. The chi-square value is computed by the procedure presented in the last section to test the hypothesis that the Z-transform of the corrected statistic is normal with mean

$$\frac{1}{2}\log\frac{1+.596}{1-.596} = .687.$$

Since this computation uses 6 cells and uses the sample standard deviation as the hypothesized standard deviation of the distribution, under the null hypothesis the computed value comes from a chi-square distribution with 4 degrees of freedom. For 4 degrees of freedom the critical value for $\alpha = .1$ is 7.78 and for $\alpha = .05$ it is 9.49. Finally, the program computes a 90% and a 95% confidence interval using the equations of section IV, the sample standard deviation for the 100 repetitions, and the value 0.596 as the point estimate. As stated in section IV, these confidence

intervals are based on the assumption that the transform of the corrected statistic is normal and unbiased. The purpose of the simulation is to estimate the standard deviation of the Z-transform. The assumption is that the standard deviation obtained using the corrected estimates does not differ greatly from the true value of the standard deviation.

1.000										
0.722	1.000									
0.801	0.708	1.000								
0.689	0.672	0.803	1.000							
0.524	0.627	0.617	0.608	1.000						
0.452	0.515	0.550	0.561	0.701	1.000					
0.637	0.533	0.529	0.423	0.306	0.225	1.000				
0.695	0.827	0.670	0.637	0.617	0.520	0.415	1.000			
0.695	0.684	0.593	0.521	0.408	0.336	0.741	0.600	1.000		
0.760	0.658	0.684	0.573	0.421	0.342	0.745	0.585	0.743	1.000	
0.596	0.749	0.487	0.489	0.465	0.433	0.503	0.680	0.640	0.570	1.000

Table I
Corrected Data

After NORM has been started it is just like running the program discussed in appendix B. To start NORM type "run norm" at the VMS $ prompt. The results of the replicated calculations of the corrected(uncorrected) correlation coefficients are placed in NPLTC.DAT(NPLTU.DAT). Then type "run nplot" and respond as directed in appendix B for program PLOT. One departure from the procedure outlined in appendix B is that NORM terminates after the first plot. This is because the plot information is always placed in a file named PLT.PRT and so a second plot in the same run would cause the first to be overwritten before it could be examined. For several plots just do "run nplot" several times. The program of appendices A and B runs on a PC and is written in a version of PASCAL in which the pseudo random number generator needs no human intervention to be initiated. NORM runs on a VAX and is written in a version of PASCAL in which the pseudo random number generator must be supplied a seed to begin execution. This seed

needs to be an odd integer and conventional wisdom holds that it should have about 5 digits for the best stochastic properties of the stream of random numbers. The user is queried for the seed each time NORM executes. It is a good idea to run each simulation two or three times using a different seed each time.

The mean and variance of all variables in the model must be supplied to NORM in addition to the correlation coefficients of table 1. For a given fixed subpopulation the means and variances of the variables do not influence the outcome. The corrected values depend only on the correlations between the variables. Hence any simulation can be done with variables all having mean zero and standard deviation one. So the following example includes only standard normal random variables. The program requires entry of the means and variances because if it just assumed that all variables were standard normal then the user could not use the real selection criteria. One would instead have to derive a set of criteria that would produce the same subpopulation that the real criteria would have produced, had the actual means and variances of the variables been used.

A simulation was done using the coefficients in table 1 . These are the corrected values from a sample collected at Brooks AFB. All 11 variables were assumed to have mean zero and variance one. The selected sample size was set at 200 and 100 replications were done. Thus for each of the 100 samples generated, observations were made until 200 applicants meet the selection criteria. As the resulting selection ratio for this example was about .25, this means that on the average about 800 observations were made for each of the 100 samples generated. In the program all variables must be given a name and in this case the names given were $X_1, X_2, \ldots, X_{11}$. The first three variables were designated as explicit selection variables and the selection conditions were $X_1 \geq 0.2, X_2 \geq 0.2$, and $X_3 \geq 0.2$. The variables of interest were designated as X_1 and X_{11} and so the actual value of the correlation of interest is .596. After NORM and NPLOT have executed, the file PLT.PRT contains the summarized results of the simulation. For the current simulation, table 2 gives the contents of PLT.PRT except that the histogram has been removed to conserve space.

Between X_1 and X_2 RHO=0.596 ; A complete description is in malc.dat
The mean sel ratio is 0.254 ; The mean est is 0.599 ; the std-dev is 0.06
$b_0 = 0.322$ $b_1 = 0.460$; corrected ; sample size is 200 ; # reps = 100
Smlest-Lrgest 0.451-0.744

[THE PLOT GOES HERE]

mean of the transform = 0.696 ; SD = 0.089 ; $1/sqrt(n-3) = 0.071$
90% capture ratio = 0.9 ; 95% capture ratio = 0.95 ; chi2 = 2.164
90%(0.494 , 0.682) ; 95%(0.472 , 0.697)
untransform of the mean of the transform = 0.602

Table II
PLT.PRT First Run

This file tells us that the correlation between the first and the eleventh variable, which is 0.596, is being estimated and that all the input parameters can be found in a file named MALC.DAT. The selection ratio is given, and it is stated that the sample mean of the 100 corrected estimates was 0.599. The sample standard deviation of the 100 corrected estimates was 0.06. The sample regression coefficients of X_{11} on X_1 are given, and it is noted that the values in this file pertain to the corrected statistic. Values for the uncorrected statistic may be obtained by running NPLOT and specifying the uncorrected option. By doing this, one would learn that, in this case, the sample mean of the 100 uncorrected estimates was 0.345. After the plot, omitted here, information about the Z-transform of the 100 corrected values is presented. It is seen that the sample standard deviation of the transformed values is 0.089 and that the 90% and the 95% capture ratios are exactly as they should be. The small chi-square value(2.164) indicates a good fit of the transformed values to a normal distribution with mean 0.687(the Z-transform of 0.596). Based on a standard deviation of 0.089, the 90% and 95% confidence intervals are (0.494 , 0.682) and (0.472 , 0.697) respectively. These interval estimates are associated with the point estimate 0.596. Also shown is

$1/sqrt(n-3) = 1/sqrt(197) = .071$. This is the expression given in the theorem at the beginning of the previous section as the approximate value of the standard deviation of the Z-transform of the uncorrected statistic. In all of our simulation studies this expression is a very good predictor of the observed sample standard deviation of the Z-transform of the uncorrected statistic. For example, in the present example this observed estimate was 0.074. The standard deviation of the Z-transform of the corrected statistic is always greater than the standard deviation of the Z-transform of the uncorrected statistic. However we must use the corrected statistic because the uncorrected statistic is frequently, as in the current example, very biased. For comparison, table 3 gives the results of another simulation using exactly the same input parameters but a different seed for the pseudo random number generator.

Between X_1 and X_2 RHO=0.596 ; A complete description is in malc.dat
The mean sel ratio is 0.253 ; The mean est is 0.593 ; the std-dev is 0.05
$b_0 = 0.319$ $b_1 = 0.455$; corrected ; sample size is 200 ; # reps = 100
Smlest-Lrgest 0.427-0.718

[THE PLOT GOES HERE]

mean of the transform = 0.687 ; SD = 0.085 ; $1/sqrt(n-3) = 0.071$
90% capture ratio = 0.91 ; 95% capture ratio = 0.97 ; chi2 = 2.634
90%(0.498 , 0.679) ; 95%(0.478 , 0.693)
untransform of the mean of the transform = 0.596

Table III
PLT.PRT Second Run

Program TEST is like program NORM except that the observed test scores are selected randomly from a database of test scores. If NREC represents the number of records in the database file, then a random observation is made by selecting one number from $1, 2, \ldots, NREC$ and then reading that record. The record number is selected at random in such a way that the numbers $1, 2, \ldots, NREC$ all have

an equal chance of being selected. This is done using a pseudo random number generator, so the user will be required to supply a seed each time the program runs. Sampling is done with replacement. In other words, a record may be selected more than once in a sample. This is, of course, not the way sampling is really done, but we are only interested in cases where NREC is sufficiently larger than the sample size that sampling with replacement will have no noticeable effect on the outcome.

The input to TEST differs slightly from that of NORM. The user does not need to supply the means, variances, and correlation coefficients of all the random variables representing test scores, as these are implicit in the database. However, the correlation between the two variables of interest is requested so that it can be printed in the output. Also the variables are not given names but are referred to by the position they occupy in the records of the database file. The user will be prompted for the name of the database file(scores file) and the number of records in that file. The number of test scores in each record of the database file is, of course, the number of variables in the simulation. So the user is not required to specify the number of variables. That number is represented by a constant in the programs called on by TEST. In order to use TEST with a database file having a different number of test scores in each record one program(TESTGBL) must be modified in one place and all programs must be compiled and linked again. The number is currently set to 11.

After TEST has been executed, TPLOT must be executed just as NPLOT must run after NORM. In order to calculate the capture ratio for the 90% and the 95% confidence intervals TPLOT uses the standard deviation from a corresponding run of NORM. A corresponding run means a run using the means, variances, and correlation coefficients of the variables in the database used by TEST. In this way the capture ratios reflect the procedure discussed in the previous section. The output of TPLOT is left in file PLT.PRT, which is the same file in which NPLOT leaves its output. The output of TPLOT is almost identical to that of NPLOT. At the top of the output the numbers of the variables of interest is given along with the actual value of the correlation between these two variables.

VI. Runs Using Brooks AFB Test Score Data

A set of test scores from Brooks AFB consists of 3930 records, each containing 11 test scores. The means, standard deviations, and correlation coefficients of these data appear in table 1.

Means

82.54	53.12	53.78	54.08	55.19	54.94	54.86	52.34	52.89	54.46	52.16

Standard Deviations

5.61	6.42	5.63	4.65	4.68	6.28	6.74	8.05	6.89	7.23	7.53

Correlation Coefficients

1.00										
0.37	1.00									
0.31	0.23	1.00								
0.37	0.59	0.14	1.00							
0.34	0.37	0.17	0.43	1.00						
0.09	−0.08	0.25	−0.07	0.06	1.00					
0.08	−0.10	0.13	−0.03	0.10	0.54	1.00				
0.27	0.37	0.20	0.25	0.19	−0.16	−0.15	1.00			
0.28	0.31	0.57	0.20	0.20	0.27	0.14	0.05	1.00		
0.29	0.41	0.38	0.30	0.20	−0.06	−0.05	0.50	0.31	1.00	
0.32	0.49	0.24	0.38	0.24	−0.12	−0.12	0.56	0.20	0.49	1.00

Table I
Brooks AFB Data

These data have already been selected on the basis of the first 10 scores, the ASVAB. Variable 11 is the criterion variable. So variables 1 through 10 are the explicit selection variables and variable 11 is the implicit selection variables. Notice that if variable 11 is linear and homoscedastic with respect to the first 10 variables before selection on those 10 variables has occurred, then it will be linear and homoscedastic after selection has occurred. Therefore, there is no logical difficulty with taking these data to represent an applicant population subject to more selection on the first 10 variables.

A number of runs of TEST were made to try and answer two questions. The first is : "How well does the confidence interval procedure described in section IV work in this data set?" Unfortunately, the answer is that it does not work as well

as we would like. It produced a very reasonable capture ratio for the 90% and 95% confidence intervals in the majority of the cases ran, but sometimes the capture ratios were very low. In the worst case, presented below, the capture ratio for 90% was .73 and the capture ratio for 95% was .82. The conclusion, at least for this data set, is that if one has no other way to get a feeling for the variability inherent in a sampling process then this procedure is worthwhile. The second question is : "In this data set is the corrected statistic more accurate than the uncorrected statistic?" The answer to this question is yes. The corrected statistic was too high and the uncorrected statistic was too low in every case ran. However the distance between the true parameter value and the uncorrected estimate was typically 2 or 3 times the distance between the true parameter value and the corrected statistic. The conclusion is that one should always use the corrected statistic.

The two cases presented here are selected because they represent the most extreme cases encountered with respect to the second question above. The first case, where the corrected statistic is considerably better, is certainly more typical than the second case, where the corrected and uncorrected statistics are essentially the same distance from the true parameter value. The two cases also give the extreme values of the capture ratios for the 90% and 95% confidence intervals. With respect to the first question neither case is typical. For the cases considered, the capture ratios bounced around between the two extreme values given in these two examples. In both cases the sample size was 100 and the number of repetitions was 100.

In the first there were three restrictions using all 10 explicit selection variables. The sum of the first 3 variables had to be at least 189.44, the sum of the next 4 had to be at least 219.07, and the sum of the last 3 explicit selection variables had to be at least 159.65. These 3 restrictions resulted in a selection ratio of 0.225. The correlation between variables 8 and 11, which can be seen from table 1 to be .56, was the parameter being estimated. The corrected estimate was .600 and the uncorrected estimate was .440. The capture ratios of the corrected statistic for the 90% and the 95% confidence intervals were 0.94 and 0.97 respectively.

In the second case there was only one restriction. That restriction was that the sum of all 10 explicit selection variables had to be at least 568.2. With this one restriction the selection ratio was 0.5. The correlation between variables 10 and 11, which was 0.49, was being estimated. The corrected estimate was 0.558 and the uncorrected estimate was 0.425. The capture ratios of the corrected statistic for the 90% and the 95% confidence intervals were .73 and .82 respectively.

VII. Recommendations

Under the assumptions of linearity and homoscedasticity, Pearson correction does not change the value of the F statistic. If these assumptions are not satisfied then the application of the correction formula is not appropriate. The application of the F test also requires normality of the error term for fixed X values and independence of error terms for different fixed X values. These assumptions probably do not hold exactly in the applicant population. But if they do hold in the applicant population, then they should also hold in the selected population provided the new test whose predictive value is being tested is not a hidden explicit selection variable. Range restriction may, of course, have an effect when real data is used, and there might be a fairly persistent bias in the test one way or the other. If a large data set from the application population could be found, then it might be advantageous to study the sampling distribution of the F statistic using this data set.

Section III establishes that hidden variables are potentially a real problem. Assuming that linearity and homoscedasticity hold when all explicit selection variables are included, then they may not hold when an explicit selection is excluded from the model. This was demonstrated in section III. If it can somehow be determined that hidden variables are sometimes causing inaccuracies in the Pearson correction statistic, then one must either include these variables in the correction or find some other statistic that gives correct results whether hidden variables are present or not. The application of such a statistic seems likely to require some assumption regarding the distribution of the error term in the linear model.

Furthermore, this statistic is likely to not be very robust with respect to this assumption. For example, the statistic proposed by Gross and McGanney requires the assumption of normality of the error terms. It also requires that the selection criteria fit in one equation. Finally, it has the property that it works poorly exactly when the Pearson correction statistic works well. These comments are not meant to be a criticism of the work of Gross and McGanney, but rather to point out the difficulty of developing a robust statistic that is able to detect the presents of hidden variables and correct for them. We should continue to seek a way to determine if hidden variables are a real problem. However, Pearson correction appears to work very well and should be replaced with another statistic only in the presents of overwhelming evidence that the new statistic performs better. The effect of range restriction is always present and Pearson correction does a very good job of correcting for this effect. If a large database of applicant scores, along with the values of suspected hidden variables for each record, could be obtained then it would be possible to do an empirical study to determine if the inclusion of these variables would improve the accuracy of the Pearson correction statistic.

A procedure for interval estimates is described in section IV, the programs to implement this procedure are described in section V, and some empirical observations of the procedure are discussed in section VI. It should be added to the observations of section VI that all of the runs made indicated that the sampling distribution of the corrected statistic is approximately normal even when the samples come from the database of test scores, as they do when running program TEST. It was also observed that the standard deviation of the Z-transform of the corrected statistic for a corresponding run of NORM was a fairly good estimate of this standard deviation from a run of TEST. Thus, the reason that the capture ratios of the confidence intervals are sometimes too low is mostly because the corrected statistic from the database is frequently a slight overestimate of the true population parameter. Even if the confidence intervals are not entirely accurate, when decisions about the validity of tests must be made, it is useful to have this information about the sampling distribution of corrected statistic. The capture

ratios of confidence intervals based on the uncorrected statistic were very low because the uncorrected statistic is considerably more of an underestimate than the corrected statistic is an overestimate.

APPENDIX A

General Description of the PC Simulation Program

The program was written in PASCAL and is currently running on an IBM compatible micro computer. The joint distribution of all of the random variables is assumed to be multinormal in the unrestricted population. The inputs to the program are listed here for reference and they will be explained later as we discuss the program.

The number of variables[nv] and their names[vname]
Unrestricted population mean and std-dev[mu , sig]
The correlation coefficients in the unrestricted population[rho]
The number of explicitly selected variables[nve] the first nve entered
The number of restrictions[nr]
The coefficients of the explicitly selected variables[ncoeff]
Cutoff value for each restriction[cutoff]
Size of the unrestricted population[nwp]
Size of the restricted population[nvp]
The number of times the experiment will be repeated[reps]
The two variables of interest in the list of variables[int1,int2]
The number of hidden explicit selection variables[hes]
Cutoff for the hidden explicit selection variable[hcutoff]

Figure 1 below is an example of a file describing the input to a run. The first line says that there are 3 variables in this case. The next three lines give the name, mean, and standard deviation of the three variables. In this case they each have mean 0.0 and standard deviation 1.0. The next three lines give the correlation matrix for the three variables. So the correlation coefficient for (x,y) is 0.86, for (x,z) it is 0.0, and for (y,z) it is 0.43. The next line gives the number of explicit selecti riables. There is 1 in this case and so X is the only explicit selection variable. it is specified that there is only 1 restriction(selection) and that the

restriction is $(1.0)X \geq 0.0$.

```
3
x  0.0   1.0
y  0.0   1.0
z  0.0   1.0
   1.00   0.86   0.0
   0.86   1.00   0.43
   0.00   0.43   1.0
1  # of explicitly restricted variables
1  number of restrictions
1.0  0.0
2  3  variables of interest
0  0.00 # hid exp sel vars  cutoff value
1  50  100
```

Figure 1.
An Input File

The selected group will consist of those persons getting a score of zero or greater on the X test. The third to last line says that the variables of interest are 2 and 3[Y and Z]. The second to last line says that there are no hidden variables. The other possibility is that this line could indicate that there was one hidden variable. If that were the case than that variable is assumed to be the last variable, Z in this case. If one hidden variable were specified then the second number on this line would be the cutoff value for this variable. So if the second to last line were "1 0.00" than the selected group would consist of those persons getting a score of zero or greater on the X test and on the Z test. Data and a histogram of the distribution will be given for the uncorrected r between X and Z and the same information is given for the Pearson correction statistic. The program calculates the Pearson correction statistic using the theorem from section I. The last line will be explained after the following discussion.

Creating a multinormal observation is equivalent to simulating one individual. In the above case this means getting three values, one for each of the three test scores X, Y, and Z. Each multinormal observation is part of the applicant group

and is also a member of the selected group if the scores satisfy all of the restrictions. For the present case this means that the score on the X test must be at least zero. One experiment is simulated by generating observations until two conditions are satisfied. There must be at least nwp observations in the applicant group and there must be at least nvp observations in the selected group. For most cases we set nwp = 1 and then the only restriction is that we have at least nvp observations in the selected group. One run of the program consists of simulating reps experiments. The last line of a file which describes a run gives nwp, nvp, and reps in that order. In figure 1, nwp = 1, nvp = 50, and reps = 100.

When program CORR begins it will ask if the user wants to enter the data necessary to describe a run or to give the name of a file which contains the data in the expected format. The file in figure 1 is called test4 and so we can just give that name to CORR and the run is specified by the input parameters in figure 1. The reason test4 is in the expected format is because CORR wrote the file on a previous run. It was written when CORR executed, and it was specified that data would be entered from the keyboard and that these data were to be saved in a file named test4. Now if one is familiar with PASCAL read statements they could use a text editor to change some of the parameters and use test4 for another run. After CORR executes, the data necessary to produce the histograms of the corrected and the uncorrected statistics are in two internal files and one must run program PLOT which will read these internal files and display this data on the printer.

For each experiment CORR calculates each of the following quantities.

b0 and b1 = the estimates of the regression parameters
statu = the uncorrected estimate of correlation coefficient
statc= the corrected estimate of the correlation coefficient calculated with the equations of theorem 3

Hence, CORR will generate reps copies of each of these parameters. In each case the two implied variables are int1 and int2, and the regression parameters are for

int2 on int1. In the case of b0 and b1 the only values retained are the totals so that after the reps experiments have been generated the mean values of these parameters may be calculated. In the case of statu and statc each observed value is retained and written to the files pltu.dat and pltc.dat respectively. As mentioned earlier, the user can run PLOT to have all these results displayed.

APPENDIX B

PC Simulation Reference Manual

Program CORR runs under the Turbo Pascal Version IV system on an IBM PC AT. It runs under the integrated environment Turbo.exe. This gives the user access to all of the features that make Turbo Pascal a user friendly system. For users not wanting to enter the integrated environment there is a command line version explained in chapter 12 of the Turbo Pascal Version IV owner's handbook. This appendix, however, presupposes that the integrated environment is being used.

To begin the Turbo integrated environment type "turbo" followed by the "enter" key. Perhaps it is preferable to keep the Turbo programs in one directory, say \v4, and the application programs in a subdirectory , say \v4\corr. Here "application programs" means CORR plus all of the other programs and units that are part of the simulation system. In this subdirectory one would also keep the data files that are input to the simulation programs. These data files describe the parameters of a run. For a complete description of the simulation programs and the format of the data files, see appendix A. If the the Turbo system is in \v4 and the applications are in \v4\corr then it might be useful to create in this subdirectory a file called "t.bat" that contained the single command "\v4\turbo". Then to start the system, get in directory \v4\corr and type "t" followed by the "enter" key.

Suppose that the Turbo programs and the application programs have all been loaded and the Turbo system has been initiated as discussed above. The screen is displaying the Turbo Pascal integrated environment main menu, consisting of File,Edit,Run,Compile, and Options. In normal circumstances the only selections necessary to run simulations are File and Run. There are a number of ways to select menu options, but only one is mentioned here and it is the must general,

in that it works from any place in the menu system. Simultaneously pressing the "alt" key and the first letter of any main menu option will select that option. To begin a simulation select the File option by simultaneously pressing "alt" and "f". The screen now shows the File option. Press "l" to select the Load sub-option. The screen responds with "*.pas". Hit the "enter" key and a directory of the application programs is shown. Use the arrow keys to place the highlighted rectangle over the name "corr.pas." Once the highlighted section is properly placed, hit the "enter" key. The screen shows the source listing of CORR and you are in the editor. Select the Run option by simultaneously pressing "alt" and "r". Now the simulation program begins to execute. While you are in the editor, before selecting the Run option, it is important not to change the source code and if you do then just do not save the changes and nothing will be damaged. From the editor, you can terminate Turbo Pascal by choosing the File option and then hitting "q" for quit.

Suppose that CORR has been started by loading it under the File option and then choosing the Run option as just explained. The first question asks if you want to describe the parameters for a new run(type "e" then the "enter" key) or if you want to specify the name of a data file that was created on a previous run and contains the parameter specifications(type "f" then the "enter" key). Since this is the first time we have executed the program we select "e". Now the program is asking for the name of a file into which will be written the parameter specifications given in this run. Type a file name and then the "enter" key. Later, when the program is executed again, it will be possible to take the "f" option to the first question and then give this file name.

After having taken the "e" option to indicate that parameters will be entered from the keyboard, rather than from a file, you must give values for all of the input parameters discussed in appendix A. Whenever giving a real number less than one, such as .543, Turbo Pascal insists that the number start with a zero. So you must enter 0.543 and not .543. In other words, Turbo Pascal wants one digit to precede the decimal point for all real numbers. Failure to comply with this rule will cause a runtime error and you will find yourself back in the editor. To begin

again select the File option, select the Load sub-option, load CORR, then select the Run option and start over. After all parameters have been specified from the keyboard or from a specified file, the program will start doing the simulation. It prints a digit d when it has completed d tenths of the repetitions. In this way it is possible to estimate, at any time, how much longer the run will take. There is no way to stop the program except at input/output operations. That is when the program is printing on the screen or waiting for input from the keyboard. If the "ctrl" key and the "pause" key are pressed simultaneously then the program will be interrupted at the next input/output operation.

After the simulation is complete you will be requested to strike any key to return to the Turbo Pascal menu. When you do strike a key you will be back in the editor. Enter the File option and then the Load sub-option as before and this time place the highlighted area over the name "plot.pas" and press the "enter" key. Select the Run option and then program PLOT will execute. At this point the user must choose to see frequency data and sample statistics for the corrected statistic[c], the uncorrected statistic[u], or to terminate the program[q]. Suppose the "c" option is taken. Now the program is asking if the user wants to set the plot parameters or if certain default settings are desired. The usual choice here is "y" meaning "yes" the user wants the program to set the plot parameters. These parameters include the minimum value, the maximum value, the number of divisions, and the physical width of a division for the horizontal axis of the frequency distribution plot. The best strategy is to first let the program select these values and then it is possible to learn from the resulting plot the range of values taken by the simulated statistic. Then one has some information to use in determining the optimal plot parameters and these may be specified on a subsequent pass. Next the program wants to know if a plot of the statistic or of the Z-transform of the statistic is desired. If the Z-transform is specified then extra information about the sample statistics of the Z-transformed observations is given after the plot.

The data at the top of the output from program PLOT is the same whether the Z-transform is specified or not. It gives the names of the two variables of

interest. It give the true value of the correlation between the variables of interest in the applicant population. It gives the name of the file that contains the input parameters for the run. It also gives the mean selection ratio, as well as the mean and standard deviation of the value of the corrected or uncorrected statistic. These are calculated from the sample of "reps" calculations of the corrected or uncorrected statistic, each based on a sample of size "nvp". "Reps" and "nvp" are user supplied input parameters that describe the characteristics of a simulation. The selection ratio for one calculation is the fraction of observations generated that satisfied the selection criteria. The regression parameters(b_0, and b_1) for the second variable on the first are printed. It is indicated whether this information refers to the "corrected" or the "uncorrected" statistic. The sample size("nvp") and the number of repetitions("reps") are given. Finally, the smallest and largest values of the statistic or corrected statistic that were observed in the "reps" experiments are printed.

Next comes the histogram of the distribution of the uncorrected or the corrected statistic or of the Z-transform of one of these. What is plotted here depends on what was specified when program PLOT started to execute. On a later pass through this program the scale of the horizontal axis may be changed.

If a Z-transform plot was made then extra information is printed after the histogram. The mean and standard deviation of the transform of the corrected or uncorrected statistic based on the "reps" repetitions is printed. The 95 % capture ratio is given. This ratio is based on a confidence interval calculated using the standard deviation of the Z-transform, not on $1/\sqrt{n-3}$. Finally the inverse Z-transform of the mean of the "reps" Z-transform values is printed.

Birnbaum, Z.W., Paulson, E., Andrews, F.C.(1950), "On the Effects of Selection Performed on some Coordinates of a Multi-Dimensional Population," *Psychometrika*, 15, no.2, 191-204.

Chatterjee and Price(1977), Regression Analysis by Example, John Wiley and Sons, pg 57,58

Dunn O.J. and Clark V.A.(1974), Applied Statistics: Analysis of Variance and Regression, John Wiley and Sons

Gross A.L. and McGanney M.L.(1987), "The Restriction of Range Problem and Nonignorable Selection Processes," *Journal of Applied Psychology*, 72 , No. 4, 604-610.

Jackson, D. E.(1988), "A Tool for Studying the Effect of Range Restriction on Correlation Coefficient Estimation," USAF-UES Summer Faculty Research Program, Contract No. F49620-87-R-0004.

Lawley, D.N.(1943), "A Note on Karl Pearson's Selection Formulae," *Proc. Roy. Soc. Edin.", Sect.A.(Math & Phys.Sec.), 62, Part I, 28-30.*

Pearson, Karl(1903), "On the Influence of Natural Selection on the Variability and Correlation of Organs," Phil. Trans. Roy. Soc. A, 200, 1-66.

VALIDATION OF AN ENLISTED AIR FORCE SPECIALTY TASK TAXONOMY AND CROSS-AFS EASE-OF-MOVEMENT PREDICTIONS

Charles E. Lance Ph.D.

and

David L. Mayfield

Department of Psychology

University of Georgia

20 December 1989

Final report prepared under Universal Energy Systems Project 210 (No. S-210-9MG-017), Air Force Office of Scientific Research Contract No. F49620-88-C-0053/SB5881-0378.

SUMMARY

The objectives of this research were to (a) provide evidence on the usefulness of 26-category enlisted AFS task taxonomy developed at AFHRL (HRL taxonomy) for research on cross-job ease-of-movement, (b) assess the convergent validity of ease-of-movement (EOM) indices developed on the basis of the HRL task taxonomy (HRL-EOM indices) against additional EOM predictors, (c) assess the postdictive validity of HRL-EOM indices against variables related to actual cross-AFS EOM, and (d) obtain some initial empirical evidence on the construct validity of ease-of-movement predictors and criteria. Data obtained in 1989 from 836 7- and 9-skill level respondents to a Skills/Knowledge Questionnaire, and in the late 1970's from 3135 retrained airmen and their immediate supervisors on Airman Retraining Program Surveys were used to address research objectives. Results indicated (a) reasonably strong support for the usefulness the HRL task taxonomy for supporting research on cross-AFS transferability of skills, (b) convergent validity of HRL-EOM indices against alternative predictors, and (c) potential value of EOM indices for predicting actual retraining ease. Future research needs include (a) development of validated measures of the ease with which individuals learn to perform a new job effectively, (b) predictive, longitudinal validation of EOM indices against actual retraining ease, and (c) convergent validation of alternative EOM estimates.

PREFACE

The Manpower, Personnel, and Training (MPT) Technology Branch of the Manpower and Personnel Division, Air Force Human Resources Laboratory (AFHRL/MODJ) is engaged in several research and development efforts designed to support the integration of MPT considerations early on in the weapon system acquisition process. The research reported here was designed to support work ongoing in AFHRL/MODJ on the development of a validated taxonomy of tasks performed by Air Force enlisted personnel, and on the development of procedures for estimating the transferability of skills across enlisted specialties.

ACKNOWLEDGEMENTS

This research was conducted under Universal Energy Systems (UES) Project 210 (No. S-210-9MG-017), Air Force Office of Scientific Research Contract No. F49620-88-C-0053/SB5881-0378. The authors thank Universal Energy Systems, the Air Force Office of Scientific Research, Dr. Michelle Lynskey, and Mr. Wayne Archer, and especially Dr. R. Bruce Gould for their assistance and support in this research.

I. INTRODUCTION

Cross-job transferability of skills (Kavanagh & Gould, 1988) has been defined as "the continuous use of acquired knowledge and abilities when moving from one job to another" (Fine, 1957a, 1957b, p. 938) and, more recently, "the ease with which individuals trained to proficiency on one job can apply acquired knowledge and skills in learning another job" (Lance, Kavanagh, & Gould, 1989a, p. 3). As noted by Lance et al. (1989a), there are large bodies of literature on topics related to the transferability of employee job skills such as (a) transfer of training to on-the-job performance (Baldwin & Ford, 1988), (b) psychological adjustment following job transfer (Brett, 1984; Pinder & Walter, 1984), and (c) patterns of career mobility (Louis, 1980), but as yet little literature exists on the measurement or prediction of cross-job transferability of skills.

The topic of transferability of employee job skills across job families has been ignored perhaps because most employee flows are (a) vertical, that is, within career ladders, (b) within, rather than across, job families, or (c) between, rather than within, organizations (Anderson, Milkovich, & Tsui, 1981; Brett, 1984; Burak & Mathys, 1980; Hall, 1976; Mobley, Griffeth, Hand, & Meglino, 1979; Muchinsky & Morrow, 1980; Pinder & Schroeder, 1987; Pinder & Walter, 1984). Cross-job family transferability of skills is an emerging concern, however, for human resource management (HRM) research and practice in the light of (a) increasingly rapid technological changes in the workplace and

consequent changes in employee skill mix requirements (Fossum, Arvey, Paradise, & Robbins, 1986; Kozlowski & Farr, 1988; Rumberger, 1981), (b) demographic shifts in the U.S. labor force (Ahlburg & Kimmel; Gentner, 1989; Gould, 1989), (c) increasing emphasis on long-range, strategic, human resource planning (Dyer, 1982, 1985; Scarpello & Ledvinka, 1988), and (d) increasing foreign economic competition (Lenway, 1985). Making the most efficient use of employees' current job skills can be viewed as part of a firm's overall HRM strategic planning mandate, and determining the transferability of employees' present knowledge and skills to address projected manpower and personnel needs can aid in addressing this end.

The transferability of employee job skills has become particularly important to the U.S. military in the light of Congressional mandates and Department of Defense policy that manpower, personnel, and training (MPT) considerations be intergrated along with the design, test, and production of new and modified weapon systems (WSs) (see Askren & Eckstrand, 1980; Gentner, 1989; Stephenson & Gentner, 1987; Weddle & Fulkerson, 1980). Basic MPT questions include (a) for manpower: How many individuals will be required to operate and maintain the new WS? How many eligible recruits will be available in the U.S. labor force? How many people will be needed in each occupational specialty supporting the WS? (b) for personnel: What aptitude and other entry-level requirements will be needed? What occupational structures will be required for the operation and maintenance of the WS? and (c) for training: What skills will

predictions, Lance et al. (1989a) found that EOM indices indicated the more difficult cross-AFS movements to be those (a) into AFSs having higher occupational learning difficulties (OLDs), (b) from AFSs having lower OLDs, and (c) into new AFSs having different aptitude requirements than the old, or From-AFS. However, Lance et al. (1989a) also noted several needs for additional research including needs for (a) additional corroborative evidence for the usefulness of work taxonomies for supporting EOM estimates, (b) empirical validation of EOM estimates against the ease with which individuals actually are able to retrain from various jobs into others, and (c) additional convergent and construct validation evidence on cross-job EOM estimates. These additional research needs provided the major impetus for the present research.

Purposes of the Present Research

As Lance et al. (1989a) pointed out, the validity of cross-job EOM estimates hinges, in part, on the appropriateness of the underlying taxonomic structure assumed for the description and analysis of the relevant work domain. Lance et al. (1989a) based their EOM estimates on subject matter expert (SME) ratings of months required for a typical incumbent to attain proficiency on tasks included within 26 categories of a task taxonomy developed by Bell and Thomasson (1984) at the U. S. Air Force (USAF) Occupational Measurement Center (OMC). The categories included in this task taxonomy and their definitions are listed in Appendix A. This OMC task taxonomy subsequently has proven useful for (a) SME allocation of tasks performed by enlisted

airmen into taxonomic categories (Gould, Archer, Filer, Short, & Kavanagh, 1989), (b) supporting SME direct ratings of Relative Time Spent and Months to Proficiency on tasks included within categories (Lance et al., 1989a), and (c) deriving cross-AFS EOM estimates (Lance et al., 1989a). However, Lance, Kavanagh, and Gould, (1989b) suggested that several of the OMC task taxonomy's categories either overlapped, or were not defined in sufficiently concrete terms. To address these concerns, Lance et al. (1989b) developed an alternative taxonomy for tasks performed by USAF enlisted personnel based on a synthesis of (a) Relative Time Spent rating data for the OMC task categories collected from 675 SMEs on a Skills/Knowledge Questionnaire (SKQ), (b) data collected from 2122 airmen on the General Work Inventory (GWI) (Cunningham & Ballentine, 1982), (c) data collected from 2494 airmen on the Electronics Principles Inventory (EPI) (Ruck, 1986), (d) existing work taxonomies, and (e) expert judgment. The categories included within this taxonomy, developed at the Air Force Human Resources Laboratory (HRL taxonomy), along with the category definitions, are listed in Appendix C. The first purpose of the present research was to determine the usefulness of the HRL task taxonomy for (a) obtaining reliable SME judgments of Relative Time Spent and Months to Proficiency on tasks included in the HRL task categories, (b) differentiating among AFSs in terms of their task contents, and (c) supporting systematic estimates of cross-AFS EOM.

The second purpose of this research was to assess the convergent validity of cross-AFS EOM estimates based on SME

judgments of Months to Proficiency on tasks within HRL task taxonomy categories. Previously, Lance et al. (1989a) found support for convergence between EOM estimates based on the OMC task taxonomy (OMC-EOM indices) and cross-AFS differences in occupational learning difficulty and aptitude requirements. An approach similar to Lance et al.'s (1989a) was taken here to assess convergence between HRL-EOM indices and additional predictors of actual EOM. Specifically, it was predicted that HRL-EOM estimates would indicate more difficult movements (a) into AFSs which have higher OLDs, (b) from AFSs which have lower OLDs, and (c) across, rather than within, aptitude requirement areas.

Although Lance et al. (1989a) demonstrated significant convergent validity for OMC-EOM estimates against AFS differences in occupational learning difficulties and aptitude requirements, they identified the additional need for validation evidence against the actual ease with which individuals are able to retrain and learn to perform a new job effectively. The third purpose of this research was to address this additional research need using a postdictive validation strategy by relating EOM indices to performance and skill utilization criteria obtained on retrained airmen as part of a large-scale retraining study conducted by AFHRL in the late 1970's (Skinner, 1981, 1982, 1983; Skinner & Alley, 1980, 1984). The general hypothesis was that EOM indices which indicated relatively more difficult cross-AFS movements would be associated with (a) lower airman performance levels in the new- or "TO-AFS", (b) lower ratings of incumbents'

technical abilities in the new- or "TO-AFS," and (c) lower ratings of the extent to which technical skills learned in the old- or "FROM-AFS" were helpful in the "TO-AFS."

Finally, the present research attempted to provide some initial empirical evidence on the construct validity of cross-AFS EOM predictors and criteria. Kavanagh and Gould (1988) suggested that cross-job transferability of skills is likely to be determined, at the job level, by factors such as job difficulty, differences in aptitude requirements, and FROM- and TO-AFS task commonalities. At the level of the individual, extent of prior job learning, and additional social and motivational factors can be considered as determinants. According to this conception, cross-job transferability of skills (rather than performance effectiveness in the TO-AFS per se) is the endogenous variable, and job- and individual-level predictors are the exogenous variables. This study attempted to provide some initial tests of linkages discussed by Kavanagh and Gould (1988) by examining relationships between several variables related to cross-job transferability of skills (ratings of performance in the TO-AFS, and judgments of the transfer of technical skills to the new assignment) and several predictors of cross-job ease-of-movement (EOM indices indicating task commonalities and learning time, differences in FROM- and TO-AFSs' OLDs and aptitude requirements).

Summary. The objectives of this research were to (a) provide evidence on the usefulness of the 26-category HRL task taxonomy described by Lance et al. (1989b) for research on cross-

job ease-of-movement, (b) assess the convergent validity of EOM indices developed on the basis of the HRL task taxonomy (HRL-EOM indices) against additional EOM predictors, (c) assess the postdictive validity of HRL-EOM indices against criteria related to actual cross-AFS EOM, and (d) attempt some initial tests of the construct validity of ease-of-movement predictors and criteria.

II. METHOD

The present research used data obtained (a) from several different sources, (b) at different levels of analysis (e.g., individual-, AFS-, and AFS aptitude cluster-level), (c) for different purposes, and (d) at different times. The data sets included for this research are described, followed by descriptions of the measures examined.

Data Sources

1988 Skills/Knowledge Questionnaire (SKQ)

Lance et al. (1989a) reported on the development of a Skills/Knowledge Questionnaire (SKQ) for the purposes of soliciting three sets of judgments from SMEs (7-, and 9-skill level airmen) about tasks performed by journeyman-level (5-skill level) airmen as they related to the OMC task taxonomy: (a) a binary Part-of-Job rating, (b) Relative Time Spent Performing tasks within each category, and (c) Months to Proficiency, or the number of months for a newly assigned recruit to attain proficiency on tasks within each of the OMC task categories. A copy of the 1988 version of the SKQ is included in Appendix B.

Thirty SMEs within each of the 47 AFSs which had the highest rates of cross-AFS movement (either "out-of," into another AFS, or "into," from other AFSs) for the prior two years were targeted as possible survey respondents. A total of 675 SKQs were returned. Data reported by Lance et al. (1989a), were from 560 airmen in one of 41 AFSs in which at least five potential research participants provided usable data, and which were assigned to a single aptitude requirement area. These 41 AFSs, their AFS Codes (AFSCs), titles, Occupational Learning Difficulty Indices (OLDs), aptitude area assignments and percentile cutoff scores (see below) are shown in Table 1. As part of the study reported by Lance et al. (1989a) EOM indices were computed for the 1640 possible movements across the AFSs shown in Table 1 using mean SME Months to Proficiency ratings from the 1988 SKQ (see Lance et al., 1989a for details). These estimates were based on the OMC enlisted task taxonomy and will be referred to below as OMC-EOM estimates.

1989 Skills/Knowledge Questionnaire (SKQ)

Questionnaire Development

Several revisions were made to the 1988 SKQ in developing the 1989 SKQ. First, questionnaire instructions were rewritten toward brevity and clarity. Second, and most importantly, task rating categories were amended from the OMC enlisted AFS task taxonomy categories to the HRL task categories described by Lance et al. (1989b), and as shown in Appendix C. This change reflected recommendations made by Lance et al. (1989b) that future EOM estimates be linked to the HRL enlisted task taxonomy

Table 1
Air Force Specialties Surveyed With the 1988 SKQ

AFSC	Specialty Title	MAGE Area	OLD
113x0c -	Flight Engineer	G-55	91
122x0 -	Aircrew Life Support	G-30	81
207x1 -	Morse Systems Operator	A-45	80
241x0 -	Safety Specialist	G-53	105
242x0 -	Disaster Preparedness	G-58	105
251x0 -	Weather Specialist	G-64	83
272x0 -	Air Traffic Control Operator	G-43	98
274x0 -	Command and Control Specialist	G-48	95
275x0 -	Tactical Command and Control Spec	G-48	92
304x0 -	Wideband Communication Equip Spec	E-67	122
305x4 -	Elec Component & Switching Sys Spec	E-67	120
306x0 -	Elec Commun & Crypto Equip Sys Spec	E-67	125
411x0c -	Missile Systems Maintenance Spec	E-67	126
411x1a -	Missile Maintenance Specialist	M-51	130
426x2 -	Jet Engine Mechanic	M-44	154
431x1 -	Tactical Aircraft Maintenance Spec	M-51	120
431x3 -	Airlift Aircraft Maintenance Spec	M-51	125
451x4 -	F-15 Avionics Test Stn & Comp Spec	E-67	124
451x5 -	F-16/A-10 Avionics Test Stn & Comp Spec	E-67	124
451x6 -	F/FB-111 Avionics Test Stn & Comp Spec	E-67	124
454x3 -	Fuel Systems Maintenance Specialist	M-51	131
456x1 -	Electronic Warfare Systems Specialist	E-67	119
472x4 -	Vehicle Maintenance Con & Analysis Tech	A-45	80
491x1 -	Communication-Computer Systems Operator	G-43	91
491x2 -	Communication-Computer Systems Programmer	G-53	91
492x1 -	Information Systems Radio Operator	A-45	80
493x0 -	Communication-Computer Systems Control	E-56	126
496x0 -	Comm-CompSystems Program Mgt Spec	G-58	91
603x0 -	Vehicle Operator/Dispatcher	M-44	78
645x1 -	Materiel Storage & Distribution Spec	G-30	70
645x2 -	Supply Systems Analysis Specialist	A-51	91
651x0 -	Contracting Specialist	A-56	93
661x0 -	Logistics Plans Specialist	A-61	125
702x0 -	Administration Specialist	A-32	67
705x0 -	Legal Services Specialist	A-45	77
732x0 -	Personnel Specialist	A-45	80
733x1 -	Manpower Management	G-64	121
751x1 -	Training Systems Specialist	G-56	96
811x0 -	Security Specialist	G-35	92
811x2 -	Law Enforcement Specialist	G-35	94
903x1 -	Nuclear Medicine Specialist	G-43	79

Note. AFSC = Air Force Specialty Code. MAGE scores (M = Mechanical, A = Administrative, G = General, E = Electronic) are percentile qualifying scores. OLD = Occupational Learning Difficulty index.

which described several task areas in more concrete terms than had the OMC taxonomy. Rating scales for judgments about tasks within task categories were identical, however, for the 1988 and 1989 versions of the SKQ: (a) a binary Part-of-Job rating, (b) a 9-point Relative Time Spent rating ("1 = Very small amount" to "9 = Very large amount"), and (c) a 9-point Months to Proficiency rating (1 = 0 to 1 month, 9 = nine or more months).

Third, two additional rating tasks were included along with the 1989 SKQ rating booklet. These ratings solicited global SME judgments regarding the relative retraining ease (a) from their own AFS to each of the other AFSs targeted for data collection, and (b) from each of the other AFSs into their own AFS. These data were collected for research purposes that were beyond the scope of the present study. Consequently, they are not reported here. The 1989 SKQ is included in Appendix D of this report.

Data Collection

A number of criteria were used to identify AFSs for data collection with the 1989 SKQ. First, a list was generated of all matching AFSCs between the most current Airman Classification Chart available (dated 30 August 1988) and the valid AFSCs recorded in the Airman Retraining Program Survey data (Skinner, 1981, 1982, 1983; see below). Only current AFSs whose AFSCs had apparently not changed since the time of the Airman Retraining Program Survey were considered eligible for data collection.

Second, retraining frequencies across those AFSCs satisfying the first criteria were calculated from the Airman Retraining Program Survey data set. The goal of this second criteria was to

identify those AFSs in the Airman Retraining Program Survey that had the highest frequencies of movement either "out of," to other AFSs, or "into," from other AFSs, so as to maximize the sample size for postdictive validation of HRL-EOM indices described below.

Third, current AFSs which had no 5-skill level (i.e., 3- 7-skill level AFSs) were eliminated from consideration. Other criteria used to target AFSs for data collection were that (a) an adequate number ($\geq$ 30) of 7- and 9-skill level incumbents were currently available for participation, (b) a broad range of occupational learning difficulties should be sampled, (b) the Mechanical, Administrative, General, and Electronics (MAGE) aptitude areas each should be well represented, and (c) a number of AFSs previously survey with the 1988 version of the SKQ should be sampled. Using these criteria, 43 AFSs were targeted for data collection. These AFSs along with their AFSCs, AFS titles, MAGE aptitude areas and percentile qualifying scores, are listed in Table 2. Those AFSs that were also surveyed with the 1988 SKQ are indicated with an asterisk to the left of the AFS title.

In July 1989, the 1989 SKQ, along with a cover letter, detailed rating instructions, and HRL task category definitions, was mailed to 1565 7- and 9-skill level incumbents in the 43 AFSs shown in Table 2. Usable surveys were returned by 836 respondents for a response rate of 53%. Respondent characteristics are summarized in Table 3. The typical respondent was male (91%), pay grade E6 or E7 (86%), had some college education (mean education = 13.26 years), supervi[illegible] five

Table 2
Air Force Specialties Surveyed With the 1989 SKQ

AFSC	Specialty Title	MAGE Area	OLD
113x0c	*Flight Engineer	G-55	91
114x0	Aircraft Loadmaster	M-55	115
207x1	*Morse Systems Operator	A-45	80
241x0	*Safety Specialist	G-53	105
242x0	*Disaster Preparedness	G-58	105
251x0	*Weather Specialist	G-64	83
271x1	Airfield Management Specialist	A-45	91
272x0	*Air Traffic Control Operator	G-43	98
276x0	Aerospace Control & Warning Systems	G-53	105
303x2	Aircraft Control & Warning Radar	E-77	136
304x0	*Wideband Communication Equip Spec	E-67	122
304x4	Ground Radio Communications Spec	E-67	123
324x0	Precision Msmt Equip Lab Spec	E-67	127
423x0	Aircraft Electrical Systems Spec	E-43	124
452x4	Tactical Aircraft Maintenance Spec	M-51	120
454x1	Aerospace Ground Equipment Mech	M-51&E-33	124
455x2	Avionic Communication Specialist	E-67	122
457x0	Strategic Aircraft Maint Spec	M-51	129
461x0	Munitions Systems Specialist	M-61/E-46	122
462x0	Aircraft Armament Systems Spec	M-61/E-46	122
472x1	Special Vehicle Mechanic	M-44	127
542x0	Electrician	E-33	101
545x0	Refrigeration and A/C Spec	M-51/E-33	127
551x0	Pavements Maintenance Specialist	M-44	97
551x1	Construction Equipment Operator	M-44	112
553x0	Engineering Assistant Specialist	G-48	91
571x0	Fire Protection Specialist	G-39	81
603x0	*Vehicle Operator/Dispatcher	M-44	78
631x0	Fuel Specialist	M-51&G-39	87
645x0	Inventory Management Specialist	A-45/G-43	79
645x1	*Materiel Storage & Distribution	G-30	70
645x2	*Supply Systems Analysis Specialist	A-51	91
651x0	*Contracting Specialist	A-56	93
661x0	*Logistics Plans Specialist	A-61	125
702x0	*Administration Specialist	A-32	67
732x0	*Personnel Specialist	A-45	80
741x1	Fitness and Recreation Specialist	A-27	76
791x0	Public Affairs Specialist	G-69	111
811x0	*Security Specialist	G-35	92
811x2	*Law Enforcement Specialist	G-35	94
902x0	Medical Service Specialist	G-43	81
906x0	Medical Administrative Spec	G-43	77
981x0	Dental Assistant	G-43	75

Note. AFSCs 452x4, 454x1, 455x2, and 457x0 were changed in 1989 from AFSCs 431x1, 423x5, 328x0, and 431x2, respectively.
*AFSs also surveyed with the 1988 SKQ.

Table 3

1989 SKQ Respondent Characteristics

Variable	N	Percent	Mean	SD	
Pay Grade:					
E4	1	.1			
E5	112	13.5			
E6	485	58.3			
E7	234	28.1			
Sex:					
Male	674	91.3			
Female	64	8.7			
Education:			13.26	1.40	Years
< High School	3	.4			
High School Graduate	323	39.2			
1-3 Years College	424	51.4			
> 4 Years College	75	9.1			
Number Supervised:			4.86	7.56	Airmen
0	226	27.1			
1	86	10.3			
2	97	11.6			
3	84	10.1			
4	74	8.9			
5	66	7.9			
6-10	103	12.3			
Over 10	99	11.9			
Job Tenure:			2.72	3.47	Years
Less than 1 Year	254	30.5			
1 - 2 Years	213	25.6			
2 - 3 Years	130	15.6			
3 - 4 Years	86	10.3			
4 - 5 Years	33	4.0			
5 Years or longer	117	14.0			
Total Active Federal Military Service:			15.31	3.88	Years
Less than 5 Years	1	.1			
5 - 10 Years	71	8.5			
10 - 15 Years	320	38.3			
15 - 20 Years	363	43.5			
over 20 Years	80	9.6			

others (mean = 4.86), had been in the job just under 3 years (mean = 32.72 months), and in the Services over 15 years (mean Total Active Federal Military Service = 183.63 months). These respondent characteristics were comparable to those from the 1988 SKQ survey (see Lance et al., 1989a). Descriptive statistics, calculation of EOM indices, and other results based on these data are described below.

Airman Retraining Program Survey (ARPS)

A comprehensive evaluation of the USAF Airman Retraining Program was initiated in late 1977 with the aim of "tracking the progress of retrainees in their new military specialties" (Skinner, 1981, p. 1110). As part of this program of research AFHRL scientists in the Manpower and Personnel Division developed two questionnaires for an Airman Retraining Program Survey (ARPS), (a) an 85-item instrument completed by airmen who had retrained into another specialty (retrainees), and (b) a 66-item instrument completed by retrainees' supervisors. Questionnaire items addressed "the retrainees' accommodation to the new specialty, the impact of retraining, and retrainees' job skill and performance" (Skinner, 1981, p. 1111).

Potential respondents for the ARPS were enlisted airmen who had retrained into a new AFS between July 1973 and August 1977 or between April 1978 and August 1979. Approximately 20,000 potential respondents representing 1733 different cross-AFS movements were identified (see Skinner, 1981). According to Skinner (1981) questionnaires were mailed to 18,065 retrainees and their supervisors and, after data editing, 13,070 retrainees

and 11,549 of their supervisors returned usable questionnaires (see Skinner, 1981 for details). ARPS data were made available for this research by AFHRL/MO. In the present research, a total of 7,358 retrainees were identified whose FROM-AFSCs (i.e., AFSCs prior to retraining) and TO-AFSCs (i.e., AFSCs subsequent to retraining) matched current AFSCs on the 30 April 1988 Airman Classification Chart.

The number of ARPS retrainees whose FROM-AFSCs and TO-AFSCs matched those of the AFSs surveyed with the 1988 and 1989 versions of the SKQ was 3135. Table 4 shows the number of ARPS retrainees who retrained across the AFSs surveyed with the 1989 SKQ. Table 4 shows that in the 1970's, some AFSs were more often "sources" for retraining (e.g., 207x1, 452x4, and 811x0) while other AFSs were more often retraining "targets" (e.g., 114x0, 242x0, 251x0, and 661x0). Table 5 shows ARPS retrainee sample demographic characteristics. The typical retrainee in the ARPS subsample generated for the present study was male (85.5%), E4 or higher (81%), a high school graduate (98%), had been in the current AFS at least one year (87%) and in the Services over 5 years (83%).

The retrainee and supervisor ARPS questionnaires contained items soliciting standard background information, items requesting information regarding retraining, a number of items pertaining to retrainee adjustment to the retraining AFS (the TO-AFS) (e.g., attitudinal, motivational, and satisfaction measures), and items addressing retrainee performance, and skill utilization in the retraining AFS (see Skinner, 1981, 1982). In

Table 4
Number of ARPS Retrainees in AFSs Surveyed With 1989 SKQ

AFSC	Specialty Title	No. Respondents FROM-AFS	TO-AFS
113x0c -	Flight Engineer	11	64
114x0 -	Aircraft Loadmaster	20	144
207x1 -	Morse Systems Operator	111	0
241x0 -	Safety Specialist	3	90
242x0 -	Disaster Preparedness	3	133
251x0 -	Weather Specialist	6	146
271x1 -	Airfield Management Specialist	5	36
272x0 -	Air Traffic Control Operator	113	105
276x0 -	Aerospace Control & Warning Systems	20	45
303x2 -	Aircraft Control & Warning Radar	37	8
304x0 -	Wideband Communication Equip Spec	25	24
304x4 -	Ground Radio Communications Spec	28	35
324x0 -	Precision Msmt Equip Lab Spec	7	103
423x0 -	Aircraft Electrical Systems Spec	43	15
452x4 -	Tactical Aircraft Maintenance Spec	329	41
454x1 -	Aerospace Ground Equipment Mech	51	14
455x2 -	Avionic Communication Specialist	17	12
457x0 -	Strategic Aircraft Maint Spec	54	53
461x0 -	Munitions Systems Specialist	41	9
462x0 -	Aircraft Armament Systems Spec	126	21
472x1 -	Special Vehicle Mechanic	15	33
542x0 -	Electrician	20	75
545x0 -	Refrigeration and A/C Spec	13	32
551x0 -	Pavements Maintenance Specialist	38	12
551x1 -	Construction Equipment Operator	22	27
553x0 -	Engineering Assistant Specialist	2	60
571x0 -	Fire Protection Specialist	79	28
603x0 -	Vehicle Operator/Dispatcher	184	53
631x0 -	Fuel Specialist	122	15
645x0 -	Inventory Management Specialist	148	78
645x1 -	Materiel Storage & Distribution	153	27
645x2 -	Supply Systems Analysis Specialist	5	72
651x0 -	Contracting Specialist	4	135
661x0 -	Logistics Plans Specialist	0	132
702x0 -	Administration Specialist	386	217
732x0 -	Personnel Specialist	72	202
741x1 -	Fitness and Recreation Specialist	18	54
791x0 -	Public Affairs Specialist	1	88
811x0 -	Security Specialist	307	34
811x2 -	Law Enforcement Specialist	70	74
902x0 -	Medical Service Specialist	30	83
906x0 -	Medical Administrative Spec	19	74
981x0 -	Dental Assistant	11	66

Note. FROM-AFS refers to AFSs from which retrainees retrained: TO-AFS refers to current or retraining AFS.

Table 5
ARPS Sample Demographic Characteristics

Variable	N	Percent	Mean	SD	
Pay Grade:					
E1	6	<.1			
E2	50	3.7			
E3	198	14.6			
E4	667	49.2			
E5	278	20.5			
E6	122	9.0			
E7	36	2.7			
Sex:					
Male	2679	85.5			
Female	456	14.5			
Education:					
< High School	31	1.9			
High School Graduate	1197	74.9			
1-3 Years College	331	20.7			
> 4 Years College	39	2.4			
Job Tenure:			2.65	1.03	Years
Less than 1 Year	357	12.8			
1 - 2 Years	988	35.4			
3 - 4 Years	818	29.3			
5 - 6 Years	545	19.5			
7 or more Years	84	3.0			
Total Active Federal Military Service:			9.27	4.78	Years
Less than 5 years	543	17.3			
5 - 10 years	1460	46.6			
10 - 15 years	662	21.1			
15 - 20 years	409	13.0			
over 20 years	61	1.9			

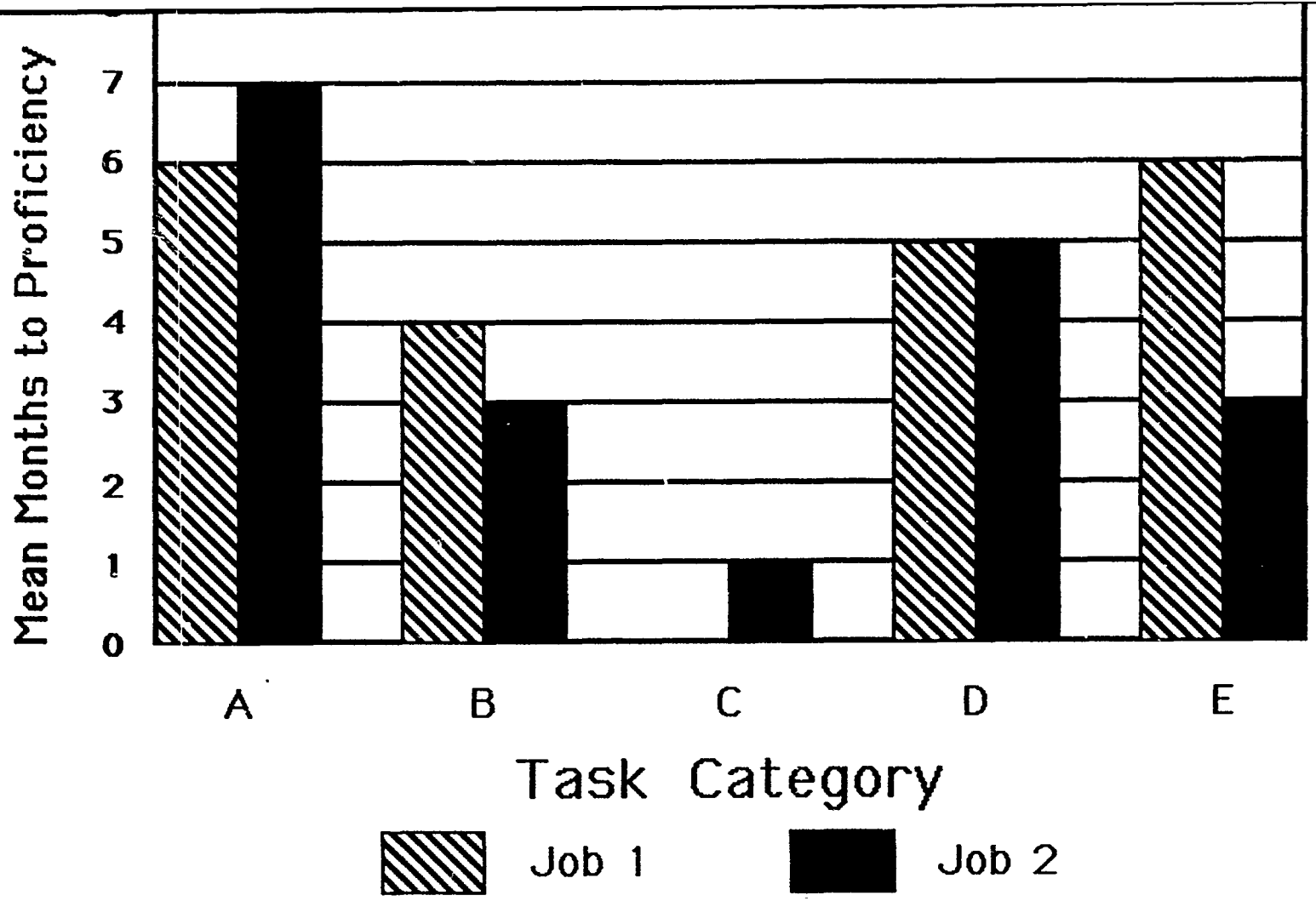
Note. Frequencies vary due to differing degrees of missing data and data source. Pay Grade, Sex, Education, and TAFMS were obtained from Airman Locator Files. TAFMS was estimated from ALF record of enlistment date with assumed current date of June 1990. Job Tenure was obtained from ARPS data.

AFS2. This procedure is illustrated in Figure 1 (from Lance et al., 1989a) where, for example, one "unit" of skills would need to be acquired in Categories A and C in moving from Job 1 to Job 2. In moving from Job 2 to Job 1, incumbents would need to acquire 1 "unit" of Category B skills and 3 "units" of Category E skills.

This procedure generated 1803 unique HRL-EOM indices (i.e., $43^2 - 43$) for movement across the 43 AFSs listed in Table 2. It should be noted that these indices (a) actually indicate the relative <u>difficulty</u> of movement, since they index the amount of task skills to be acquired in moving from AFS$_i$ to AFS$_j$, (b) are asymmetric with respect to movements between AFS$_i$ and AFS$_j$, and (c) reflect the <u>relative</u> difficulty of movement, since no absolute metric (e.g., actual number of months of job learning time) was established.

<u>Occupational Learning Difficulty Indices</u>

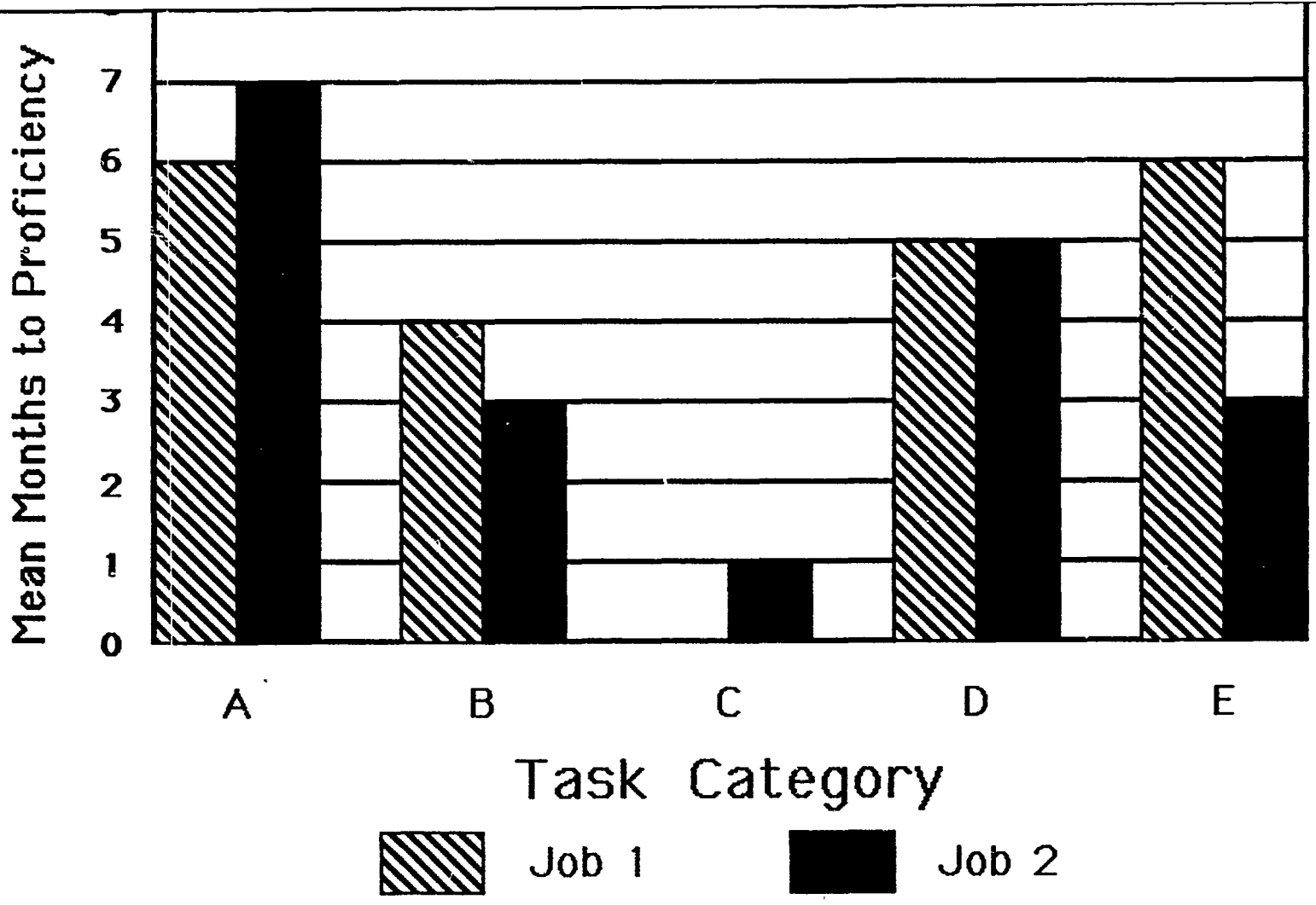
For the past 20 years, the USAF has maintained a program of research on the measurement of job and task learning difficulty (Burtch, Lipscomb, & Wi sman, 1982; Garcia, Ruck, & Weeks, 1985; Lecznar, 1971; Mead & Christal, 1970; Mumford, Weeks, Harding, & Fleishman, 1987; Ramage, 1987; Weeks, 1984). Task learning difficulty ratings of the length of time required for a typical incumbent to learn to perform tasks satisfactorily are collected as part of periodic occupational surveys (Christal, 1974). Occupational learning difficulty indices (OLDs) are specialty-level indices of the degree of specialty learning difficulty and are aggregates of tasks' difficulties that represent the average



Ease-of-Movement (EOM): Job 1 --> Job 2 = 2

Job 2 --> Job 1 = 4

Figure 1. Computation of Ease-of-Movement (EOM) Indices



Ease-of-Movement (EOM): Job 1 --> Job 2 = 2

Job 2 --> Job 1 = 4

Figure 1. Computation of Ease-of-Movement (EOM) Indices

task difficulty per unit time spent performing them. OLDs have recently been benchmarked across AFSs on a common 25-point reference scale (see Ramage, 1987) and range, theoretically, from a low of 10 to a high of 250 (Davis, Archer, Gould, & Kavanagh, 1989). OLDs for the AFSs included in the Lance et al. (1989a) study and in the present study are shown in Tables 1 and 2, respectively. It should be noted that these AFS OLDs represent the occupational learning difficulty of current AFSs but may not precisely indicate the learning difficulty of AFSs having the same AFSC at the time of the ARPS study. However, the assumption was made that differences in AFSs' OLDs substantially reflected accurate rank-ordering of the AFSs included for this research in terms of their learning difficulties.

Aptitude Requirements

The U. S. military bases accession and classification decisions, in part, on applicants' Armed Services Vocational Aptitude Battery (ASVAB) scores (Department of Defense, 1984). Research and development of the ASVAB has recently been called "without question the most significant historical collaboration between psychology and the military" (Armor, 1989, p. 4). Each Service uses a somewhat different configuration of ASVAB composites for classification of enlistees but all of the Services use the Armed Forces Qualification Test (AFQT) composite (an indicator of general cognitive ability) to establish applicants' qualifications for enlistment. The USAF uses four ASVAB subtest composites: Mechanical, Administrative, General, and Electronic for classification of enlistees (Department of

Defense, 1984) and each AFS is classified as drawing upon one (or more) of these aptitude areas (Alley, Treat, & Black, 1988). MAGE area assignments and selector area percentile qualifying scores for the AFSs included in the Lance et al. (1989a) study and in the present study are shown in Tables 1 and 2, respectively. These indices were used to index AFS-level aptitude requirements in results described below. As with the OLDs described above, AFS MAGE area assignments and percentile qualifying scores may not have been identical between the time of the 1988 and 1989 SKQ surveys and the ARPS study, but it was assumed that they were substantially similar.

Individuals' standardized AFQT scores were also available from the ARPS data to index individual-level aptitude levels. However, most, if not all, of the ARPS retrainees' AFQT scores were computed from ASVAB subtest scores during a time when the ASVAB had been misnormed (Ree, Mathews, Mullins, & Massey, 1982). Consequently, relations between ARPS retrainees' AFQT scores and other measures reported below should be interpreted with caution.

Ease-of-Movement Criteria

The primary EOM criteria were retrainee and supervisor ratings on ARPS questionnaire items. Criteria addressed three constructs related to actual ease-of-movement: (a) the extent to which previously acquired technical skills were useful in performing the TO-AFS, (b) present technical skills and abilities, and (c) performance effectiveness in the TO-AFS.

Previous Skill Utilization

Two ARPS questionnaire items concerned the extent to which

technical skills learned in the retrainee's previous AFS (the "old" or FROM-AFS) were helpful in the current assignment. The retrainee rating item asked retrainees to indicate how helpful the technical job skills learned in the previous specialty (FROM-AFS) were to them in the current specialty (TO-AFS) (ARPS Retrainee Questionnaire Item 36, recoded in the present study as "1 = Not at all helpful" to "5 = Very helpful). A similar supervisor rating item asked retrainees' supervisors to indicate how helpful the retrainee's technical job skills learned in the previous specialty were to retrainees in the retrainee's current specialty (ARPS Supervisor Questionnaire Item 46, Coded "1 = Not at all helpful" to "5 = Very helpful").

Current Skills/Ability

Two ARPS questionnaire items solicited evaluations of retrainees' current job-related skills and abilities. The retrainee rating asked retrainees to rate their own "technical ability to do job" (ARPS Retrainee Questionnaire Item 51, coded in the present study as "1 = Poor" to "5 = Excellent"). The supervisor rating asked retrainees' supervisors to rate the retrainees' "job skills and knowledge" (ARPS Supervisor Questionnaire Item 40, coded "1 = Poor" to "5 = Excellent").

Performance Effectiveness

Two ARPS Supervisor Questionnaire items (Items 43 and 44) asked retrainees' supervisors to rate retrainees' "quality of work performance" and "amount of work performance" (coded in the present study as "1 = Poor" to "5 = Excellent"). These two items were highly correlated in the sample of retrainees' supervisors

generated for the present research (n = 2939, r = .88, p < .01). Consequently, responses were combined into an additive retrainee performance composite.

Analyses

The general analytic approach in the present study was correlational. However, the specific analyses and data sets analyzed varied with the four main research objectives of this study. Analyses are described along with the four sets of results reported below.

III. RESULTS I: EVALUATION OF THE USEFULNESS OF THE HRL TASK TAXONOMY FOR CROSS-JOB EASE-OF-MOVEMENT RESEARCH

Demographic characteristics of respondents to the 1989 SKQ were presented in Table 3. Descriptive statistics for 1989 SKQ Part-of-Job ratings are shown in Table 6. Also shown in Table 6 are AFSCs and titles of AFSs representative of those having the highest proportion of respondents who indicated that tasks within each of the HRL taxonomy categories were "part of" their job. Table 6 shows that the overall part-of-job percentages varied considerably across the HRL task categories from a low of 2% for Animal Care to a high of 87% for Training, indicating that tasks within some categories (e.g., Clerical, Physical/Manual Labor, Oral/Written Communication, Managing Others, and Training) are widely performed by incumbents of the AFSs surveyed, while tasks in other categories (e.g., Food Preparation, Animal Care, Fabric/Rope Work and the Medical categories) are less frequently performed. Most of the representative AFSs shown in Table 6 as

Table 6
1989 SKQ Descriptive Statistics: Part-of-Job Ratings

Task Category	Part of Job	Representative AFS with High Percent "Part-of-Job"
Clerical	73%	702x0 - Administration Specialist
Personnel	49%	732x0 - Personnel Specialist
Maint Inventories	58%	645x0 - Inventory Management
Computational	52%	113x0c- Flight Engineer
Mech Syst Maint	47%	454x1 - Aerospace Ground Equip Mech
Mech Syst Operation	52%	545x0 - Refrigeration & A/C Maint
Complex Elec Circuit	21%	324x0 - Precision Measurement Equip
Digital Syst Maint	17%	304x0 - Wideband Communications Equip
Commun Syst Maint	18%	304x4 - Ground Radio Communications
Basic Elec Repair	32%	423x0 - Aircraft Electrical Syst
Elec Peripherals	13%	304x4 - Ground Radio Communications
Physical/Man'l Labor	68%	551x1 - Construction Equip Operator
Manufacturing/Fabr	16%	551x0 - Pavements Maintenance
Construction	16%	551x0 - Pavements Maintenance
Med-Patient Care	8%	902x0 - Medical Service Spec
Med-Technical	5%	902x0 - Medical Service Spec
Oral/Written Commun	83%	791x0 - Public Affairs Spec
Planning/Prob Solv	70%	241x0 - Safety Specialist
Science/Engineering	22%	251x0 - Weather Specialist
Artistic - A/V	17%	242x0 - Disaster Preparedness
Food Preparation	3%	114x0 - Aircraft Loadmaster
Animal Care	2%	811x2 - Law Enforcement Spec
Fabric/Rope Work	3%	571x0 - Fire Protection Spec
Managing Others	68%	303x2 - Aircraft Control & Warning
Training	87%	242x0 - Disaster Preparedness Spec
Surveillance	24%	811x2 - Law Enforcement Spec

Note. See Appendix C for detailed definitions of task categories. AFSs shown are representative of those having the highest proportion of respondents indicating that tasks within each category were part of their job.

having the highest rates of part-of-job endorsements were those that might have been expected to have been listed, indicating that AFSs were predictably differentiated on the basis of their job duties.

Table 7 shows descriptive statistics for 1989 SKQ relative time spent (RTS) and months to proficiency (MTP) ratings. RTS and MTP rating means were calculated including zero values assigned to responses that were "missing" because respondents indicated that task categories were not "part of" their job. Values of zero were assigned to represent "zero time spent" and "zero months to proficiency" in these instances. RTS and MTP means indicated considerable cross-category variability in the average time spent and learning times for the HRL task categories.

Table 7 also shows intraclass correlation coefficients (ICCs) (Lahey, Downey, & Saal, 1983; Shrout & Fleiss, 1979) for SKQ RTS and MTP ratings. ICCs are interpreted as indices of interrater reliability and index between-group rating variance relative to within-group variance. Consequently, ICCs vary both as a function of within-group interrater agreement and of between-group differences in rating means (James, Demaree, & Wolf, 1984). ICC(1,$\underline{k}$) indexes the reliability of the mean of $\underline{k}$ judges' ratings, and in the present case, the mean $\underline{k}$ across AFSs equaled 19.4. RTS and MTP ICCs generally were high. Most of the lower ICCs were associated with task categories which either were endorsed as "part of" the job very frequently (Managing Others and Training) or very infrequently (Animal Care and Fabric/Rope

Table 7
1989 SKQ Descriptive Statistics and Intraclass Correlations

Task Category	Relative Time Spent			Months to Proficiency		
	Mean	SD	ICC (1,k)	Mean	SD	ICC (1,k)
Clerical	3.47	2.95	.901	2.66	2.65	.769
Personnel	1.97	2.47	.782	2.20	2.99	.440
Maint Inventories	2.40	2.64	.866	1.96	2.47	.764
Computational	2.36	2.80	.868	1.91	2.57	.748
Mech Syst Maint	2.85	3.47	.972	2.66	3.41	.967
Mech Syst Operation	2.96	3.39	.923	2.32	3.02	.921
Complex Elec Circuit	1.92	2.79	.979	1.58	3.24	.970
Digital Syst Maint	1.03	2.47	.979	1.23	2.88	.976
Commun Syst Maint	1.16	2.68	.987	1.25	2.88	.980
Basic Elec Repair	1.97	3.21	.987	1.87	3.09	.974
Elec Peripherals	0.63	1.89	.926	0.65	1.93	.920
Physical/Man'l Labor	3.56	3.22	.910	1.44	1.93	.804
Manufacturing/Fabr	0.58	1.57	.779	0.61	1.75	.800
Construction	0.72	2.01	.962	0.69	2.03	.940
Med-Patient Care	0.45	1.77	.982	0.42	1.65	.965
Med-Technical	0.29	1.45	.974	0.28	1.37	.941
Oral/Written Commun	4.85	3.01	.868	4.36	3.19	.741
Planning/Prob Solv	3.91	3.07	.748	4.46	3.64	.642
Science/Engineering	1.17	2.48	.912	1.41	2.92	.866
Artistic - A/V	0.74	1.92	.920	0.58	1.69	.853
Food Preparation	0.07	0.47	.923	0.08	0.58	.870
Animal Care	0.06	0.50	.754	0.07	0.65	.622
Fabric/Rope Work	0.11	0.73	.714	0.13	0.84	.682
Managing Others	3.44	2.87	.605	4.60	3.81	.471
Training	5.14	2.71	.763	5.64	3.24	.439
Surveillance	1.34	2.74	.926	1.07	2.37	.855

Note. "Missing" responses for categories that were not considered "part of" the job were coded "0" to indicate "zero time spent" and "zero months to proficiency." ICC(1,k) indexes the reliability of the mean of k judges' ratings (mean k = 19.4, for ratings in this table).

Work). That is, restricted between-group mean differences on these categories may have attenuated the ICC reliability estimates. ICCs also were lower for the Personnel and Planning/Problem Solving categories. This may have been due to the slightly more abstract nature of these categories, or to actual cross-job, but within-AFS, differences in the extent to which incumbents perform tasks in these categories. Finally, MTP ICCs were generally lower than were RTS ICCs. This may be because "months-to-proficiency" judgments generally are less familiar to raters than are judgments of the amount of time spent on job tasks, or because the former requires greater inference on the part of the rater.

In summary, results shown in Tables 6 and 7 indicate that ratings on the 1989 SKQ's HRL task categories generally were reliable and were useful for distinguishing among AFSs' task contents and learning times. This supports the usefulness of the HRL task taxonomy according to two criteria for a useful classificatory system: reliable measurement of relevant characteristics of objects (AFSs) to be classified, and meaningful differentiation among the objects on the basis of measured characteristics.

Another desirable feature of classificatory systems is that the dimensions used for classification be nonredundant. Lance et al. (1989b) suggested that several of the OMC task taxonomy categories (Bell & Thomasson, 1984) appeared to overlap considerably, and identified this as one limitation to the OMC taxonomy. Correlations were computed among the 1989 SKQ RTS

ratings to determine potential sources of redundancy among rating categories. These correlations are shown in Appendix E. Correlations were near zero between most of the rating categories, indicating near orthogonality. Correlations were significantly positive, however, among categories relating to electronic, mechanical, medical, and verbal/quantitative areas (see Appendix E). Correlations among rating categories may signal some redundancy among the categories, or co-performance of tasks within categories that relate to the same broad category of job tasks.

Principal components analyses were conducted on the matrix of correlations among 1989 SKQ RTS ratings to summarize relationships among the category ratings. Scree tests (Cattell, 1966; Zwick & Velicer, 1986) suggested either a four- or seven-component solution. Varimax-rotated four- and seven-component solutions are shown in Tables 8 and 9, respectively. The components in Table 8 are clearly interpretable as I: Electronic, II: General and Administrative, III: Mechanical, and IV: Medical, and capture the major sources of category intercorrelation in Appendix E. The first three of these four components are similar to AFS clusters identified by Alley et al. (1988) by clustering multiple regression equations in which training criteria were regressed on ASVAB subtests. The four components in Table 8 may be interpreted in terms of a broader, higher-order taxonomy of enlisted AFS tasks. Note, however, that three of the less frequently endorsed HRL task categories (Food Preparation, Animal Care, and Fabric/Rope Work) failed to have significant loadings

Table 8

Varimax Rotated Principal Components Results for 1989 SKQ Relative Time Spent Ratings: Four Component Solution

	Principal Component			
Task Category	I	II	III	IV
Clerical		.583	-.330	
Personnel		.634		
Maint Inventories		.333	.338	
Computational		.523		
Mech Syst Maint			.788	
Mech Syst Operation			.828	
Complex Elec Circuit	.924			
Digital Syst Maint	.926			
Commun Syst Maint	.915			
Basic Elec Repair	.794			
Elec Peripherals	.792			
Physical/Man'l Labor			.788	
Manufacturing/Fabr			.549	
Construction			.384	
Med-Patient Care				.910
Med-Technical				.900
Oral/Written Commun		.683		
Planning/Prob Solv		.679		
Science/Engineering		.497		
Artistic - A/V		.536		
Food Preparation				
Animal Care				
Fabric/Rope Work				
Managing Others		.611		
Training		.545		
Surveillance		.327		

Note. See Appendix C for detailed definitions of task categories. Only loadings $\geq$ |.300| are shown.

on any of the four components.

The first six components of the seven-component solution shown in Table 9 also are clearly interpretable as I: Electronics, II: General and Administrative, III: Mechanical, IV: Medical, V: Technical, and VI: Protective Services. However, the seventh component, on which the Computational, Food Preparation, and Training rating categories had significant loadings, was not readily interpretable. In combination, principal component results in Tables 8 and 9 indicate that the majority of the HRL taxonomy task categories relate to four to seven broader task content dimensions, at least three of which (Electronics, Mechanical, and General and Administrative) correspond to the USAF's MAGE aptitude areas.

Perhaps the most important practical criterion for a classificatory system is whether it leads to reasonable classifications of the objects (AFSs) to be classified. This question was addressed for the HRL task taxonomy by clustering the 43 AFSs in Table 2 on the basis of 1989 SKQ mean RTS profiles. First, RTS means were calculated within each task category for each AFS. Next, a 43 x 43 dissimilarity matrix was calculated by computing pairwise Euclidean distances between each combination of AFSs across the 26 1989 SKQ category RTS means. Ward's (1963) hierarchical clustering algorithm was used to cluster AFSs. A plot of the sum of squared error at each clustering stage was used to determine the number of clusters and discontinuities were indicated both for a nine- and a three-cluster solution (this criterion is similar to the Scree test for

Table 9

Varimax Rotated Principal Components Results for 1989 SKQ Relative Time Spent Ratings: Seven Component Solution

	Principal Component						
Task Category	I	II	III	IV	V	VI	VII
Clerical		.666					
Personnel		.725					
Maint Inventories		.455	.407				
Computational		.382			.369		.493
Mech Syst Maint			.819				
Mech Syst Operation			.835				
Complex Elec Circuit	.925						
Digital Syst Maint	.926						
Commun Syst Maint	.918						
Basic Elec Repair	.791						
Elec Peripherals	.794						
Physical/Man'l Labor			.796				
Manufacturing/Fabr			.484		.385		
Construction					.671		
Med-Patient Care				.924			
Med-Technical				.919			
Oral/Written Commun		.639					
Planning/Prob Solv		.582					
Science/Engineering					.718		
Artistic - A/V		.400			.508		
Food Preparation							.748
Animal Care						.577	
Fabric/Rope Work						.370	
Managing Others		.670					
Training		.587				.308	-.303
Surveillance						.655	

Note. See Appendix C for detailed task category definitions. Only loadings $\geq$ |.300| are shown.

determining the number of common factors or principal components to retain).

Clustering results are shown in Table 10. The three-cluster solution, indicated by Roman numerals, was clearly interpretable as: I: Mechanical, II: Electronic, and III: General and Administrative. These results again are consistent with Alley et al.'s (1988) results from clustering ASVAB subtest prediction equations and with the USAF's MAGE aptitude area classification. The nine-cluster solution also identified meaningful subgroups of AFSs (identified by capital letters suffixed to Roman numerals in Table 10) within the broader three-cluster structure. Within the Mechanical cluster, three sub-clusters of AFSs were identified: (a) AFSs involved in manufacturing and construction (IA: Mechanical/Construction), (b) AFSs in which mechanical systems maintenance and operation were core activities (IB: Mechanical), and (c) AFSs in which incumbents perform both mechanically- and electrically-oriented tasks (IC: Mechanical/Electrical). On the other hand, five subclusters were identified within the broader General and Administrative (G&A) cluster: IIIA: G&A - Medical, IIIB: G&A - Surveillance, IIIC: G&A - Clerical, IIID: G&A - Technical, and IIIE: G&A Logistics. AFSs were not clearly clustered according to General versus Administrative aptitude areas, but this has not been an uncommon finding (e.g., Alley et al., 1988). In general, however, both the broader three-cluster solution and the more specific nine-cluster solution grouped AFSs into meaningful and relatively homogeneous specialty groups.

Table 10

Air Force Specialty Clusters Based on Mean 1989 SKQ Relative Time Spent Profiles

---

Cluster IA: Mechanical/Construction

551x0 - Pavements Maintenance Specialist
551x1 - Construction Equipment Operator

Cluster IB: Mechanical

452x4 - Tactical Aircraft Maintenance Specialist
457x0 - Strategic Aircraft Maintenance Specialist
461x0 - Munitions Systems Specialist
571x0 - Fire Protection Specialist
631x0 - Fuel Specialist

Cluster IC: Mechanical/Electrical

423x0 - Aircraft Electrical Systems Specialist
454x1 - Aerospace Ground Equipment Mechanic
462x0 - Aircraft Armament Systems Specialist
472x0 - Special Vehicle Mechanic
542x0 - Electrician
545x0 - Refrigeration & Air Conditioning Specialist

Cluster II: Electronics

303x2 - Aircraft Control & Warning Radar Specialist
304x0 - Wideband Communications Specialist
304x4 - Ground Radio Communications Specialist
324x0 - Precision Measurement Equipment Laboratory Spec
455x2 - Avionic Communication Specialist

Cluster IIIA: General & Administrative - Medical

902x0 - Medical Service Specialist
981x0 - Dental Assistant

Cluster IIIB: General & Administrative - Surveillance

207x1 - Morse Systems Operator
272x0 - Air Traffic Control Operator
276x0 - Aerospace Control & Warning Systems Operator
811x0 - Security Specialist
811x2 - Law Enforcement Specialist

Table 10 (continued)

Cluster IIIC: General & Administrative - Clerical

651x0 - Contracting Specialist
661x0 - Logistics Plans Specialist
702x0 - Administration Specialist
732x0 - Personnel Specialist
791x0 - Public Affairs Specialist
906x0 - Medical Administrative Specialist

Cluster IIID: General & Administrative - Technical

241x0 - Safety Specialist
242x0 - Disaster Preparedness Specialist
251x0 - Weather Specialist
553x0 - Engineering Assistant

Cluster IIIE: General & Administrative - Logistics

113x0c- Flight Engineer
114x0 - Aircraft Loadmaster
271x1 - Airfield Management Specialsit
603x0 - Vehicle Operator/Dispatcher
645x0 - Inventory Management Specialist
645x1 - Materiel Storage & Distribution Specialist
645x2 - Supply Systems Analysis
741x1 - Fitness & Recreation Specialist

Note. Mean relative time spent profiles are given in Appendix G.

Summary

Among the possible criteria for a useful classificatory system (taxonomy) are the following: (a) relevant characteristics of objects to be classified should be measured reliably, (b) objects to be classified should be meaningfully differentiated from one another on the basis of measured characteristics, (c) classificatory dimensions should be nonredundant, and (d) classifications should result in meaningful, homogeneous, groups of objects. Results presented in this section relate to these criteria for the usefulness of the HRL enlisted task taxonomy for classifying AFSs as follows: (a) interrater agreement generally was high for relative time spent and months to proficiency ratings on the HRL taxonomy's 26 task categories, (b) there was considerable cross-AFS variability in SMEs' part-of-job and relative time spent judgments on the HRL taxonomy's 26 task categories that was consistent with differences in AFSs' task content, (c) correlations among the HRL task categories generally were low, but results of principal components analyses indicated that most categories could be related to one of four to seven broader task dimensions, and (d) cluster analyses based on mean relative time spent profiles resulted in nine specific and three broader clusters of AFSs that were both meaningful and relatively homogeneous.

IV. RESULTS II: CONVERGENT VALIDATION OF EASE-OF-MOVEMENT INDICES AGAINST ADDITIONAL EASE-OF-MOVEMENT PREDICTORS

As described earlier, mean months to proficiency rating profiles were calculated across the HRL taxonomy's 26 task categories from SME ratings on the 1989 SKQ. A sample of mean months to proficiency profiles representative of those that were generated are shown in Figures 2 through 9. In these Figures, the height of each bar indicates mean SME months to proficiency judgments for each of the HRL taxonomy's task categories. The absence of a bar for some task categories indicates that SMEs judged that no tasks within these categories are performed. Mean months to proficiency profiles for all AFSs surveyed with the 1989 SKQ are included in Appendix G.

Ease-of-movement (EOM) indices were generated using an adaptation of a procedure described by Gould et al. (1989) that has also been reported by Lance et al. (1989a). The following results assess the convergent validity of the HRL-EOM estimates with (a) the occupational learning difficulties of the FROM-AFSs and the TO-AFSs referenced in moving from AFS$_i$ to AFS$_j$, (b) differences in the FROM- and TO-AFSs' aptitude requirements, and (c) EOM indices developed earlier by Lance et al. (1989a) on the basis of 1988 SME SKQ ratings on the OMC taxonomy's 26 task categories (OMC-EOM indices).

Earlier, Lance et al. (1989a) presented rationales arguing that, in general, cross-job movements should be more difficult to the extent that the new job is difficult to learn. They confirmed a significant relationship between OMC-EOM indices and

Figure 2
AFSC 242x0
Disaster Preparedness

Figure 3
AFSC 324X0
Precision Msmt.
Equip. Lab Spec.

Figure 4.
AFSC 457x0
Strategic
Aircraft
Maintenance

Figure 5
AFSC 542X0
Electrician

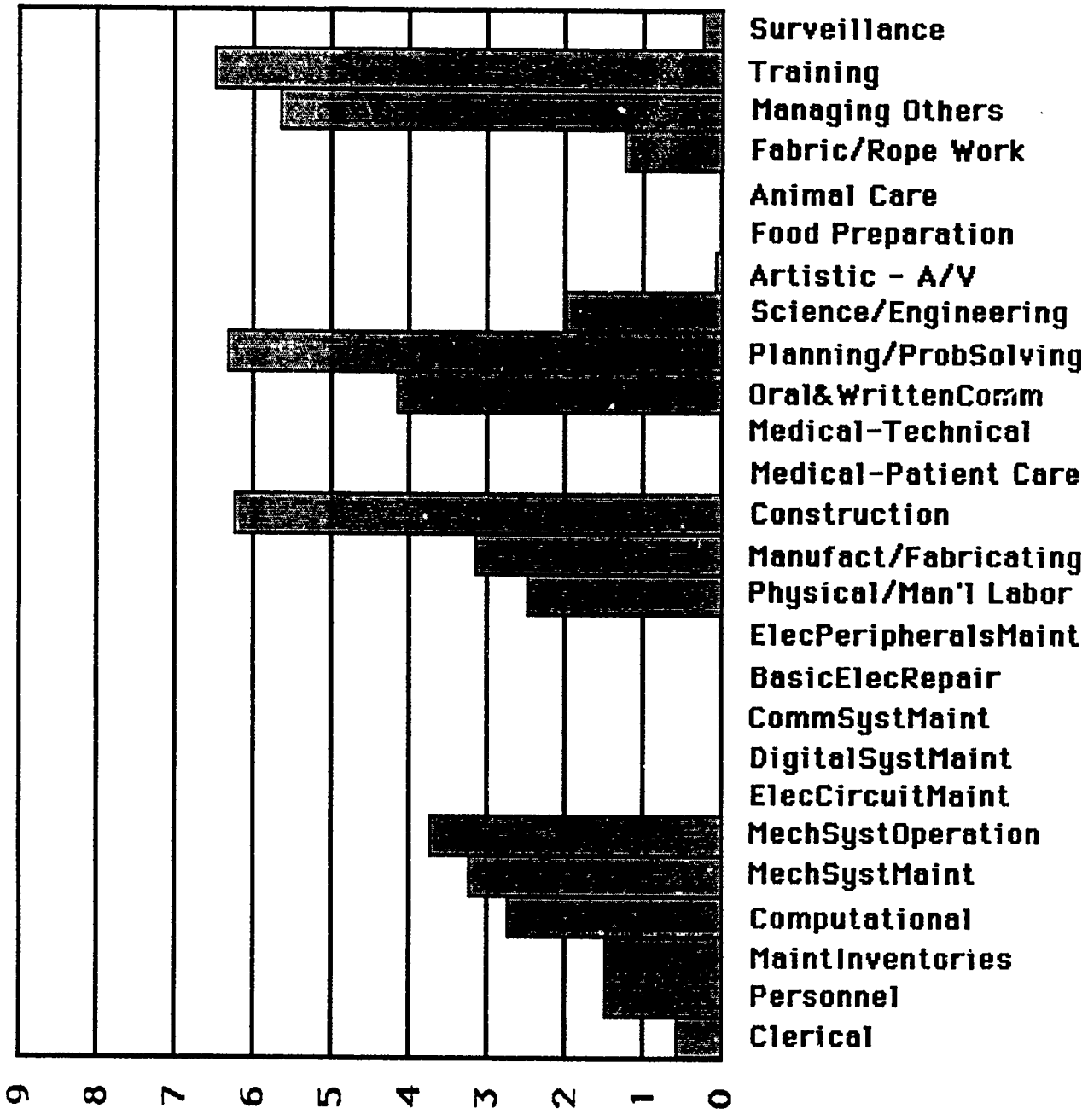


Figure 6
AFSC 551x0
Pavements
Maintenance

Figure 7
AFSC 645X2
Supply Systems
Analysis

Figure 3
AFSC 81 1X2
Law Enforce.
Specialist

Figure 9
AFSC 902X0
Medical Service Specialist

TO-AFS occupational learning difficulties (OLDs) in a sample of AFSs surveyed with the 1988 SKQ. Recalling that the EOM indices actually reflect the difficulty of cross-job movement, it was hypothesized that the HRL-EOM indices would correlate positively with TO-AFS OLDs of the AFSs surveyed with the 1989 SKQ. The correlation between TO-AFS OLDs and the HRL-EOM indices shown in Table 11 (r = .427, $p \leq .01$) supported this hypothesis.

On the other hand, Lance et al. (1989a) argued that movements from AFSs that are more difficult to learn should generally be easier than from less difficult AFSs, because skills and knowledge acquired in more difficult AFSs would more likely be generalizable to learning other jobs. Lance et al. (1989a) supported this hypothesis in finding a significant negative correlation between FROM-AFS OLDs and OMC-EOM indices among AFSs surveyed with the 1988 SKQ. The correlation between FROM-AFS OLDs and HRL-EOM indices shown in Table 11 (r = -.154, $p \leq .01$) also supported this hypothesis for the AFSs surveyed with the 1989 SKQ.

Lance et al. (1989a) also found support for their hypothesis that cross-job movements should be indicated to be easier to the extent that the FROM- and TO-AFSs require similar, versus different, aptitudes. As in the Lance et al. (1989a) study, this hypothesis was tested by correlating HRL-EOM indices with a dummy-coded variable (MAGE-F/T) that indicated whether the FROM-AFS to TO-AFS movement was within (= 1) or across (= 2) MAGE aptitude areas. The positive correlation between HRL-EOM indices and the binary MAGE-F/T variable shown in Table 11 (r = .210, $p \leq$

Table 11

Descriptive Statistics and Correlations With Ease-of-Movement Indices

	Mean	Standard Deviation	Correlation With Ease-of-Movement Indices
1. To-AFS OLD	101.30	19.64	.427**
2. From-AFS OLD	101.30	19.64	-.154**
3. MAGE-F/T	1.75	0.44	.210**

** $p < .01$

.01) supported this hypothesis for the AFSs surveyed with the 1989 SKQ.

Analysis of variance (ANOVA) results on the relationship between HRL-EOM indices and FROM- and TO-AFSs' MAGE aptitude areas are shown in Table 12. Consistent with Lance et al.'s (1989a) earlier findings, (a) the relationship of EOM indices was much stronger with the TO-AFS's aptitude area than with the FROM-AFS's, and (b) a significant interaction between the FROM- and TO-AFS aptitude areas was supported in the prediction of EOM indices. Cell means corresponding to the ANOVA results in Table 12 are plotted in Figure 10. Recalling again that the the HRL-EOM index actually reflects the difficulty of movement, Figure 10 shows that (a) movement is indicated to be generally most difficult into Electronic AFSs and least difficult into Administrative AFSs, (b) movement is indicated to be most difficult from General and Administrative AFSs, and least difficult from Electronic AFSs, and (c) movements tend to be easier within, rather than across, MAGE areas.

The final hypothesis tested by Lance et al. (1989a) was of a moderating effect of differences between FROM- and TO-AFS aptitude requirements on the relationship between TO-AFS OLDs and EOM indices. The general hypothesis which was supported in the Lance et al. (1989a) study was of a strong relationship between EOM indices and the learning difficulty of the TO-AFS (TO-AFS OLDs) for movements across different MAGE areas, but not within the same MAGE area. As in the Lance et al. (1989a) study, this hypothesis was tested here using moderated regression (Cohen &

Table 12

Analysis of Variance Results for Ease-of-Movement Indices

Source	DF	Mean Square	F
1. From-AFS MAGE (F)	3	2489.599	43.299**
2. To-AFS MAGE (T)	3	27931.844	485.791**
3. F x T	9	1644.856	28.607**
4. Residual	1790	57.498	

** $p < .01$

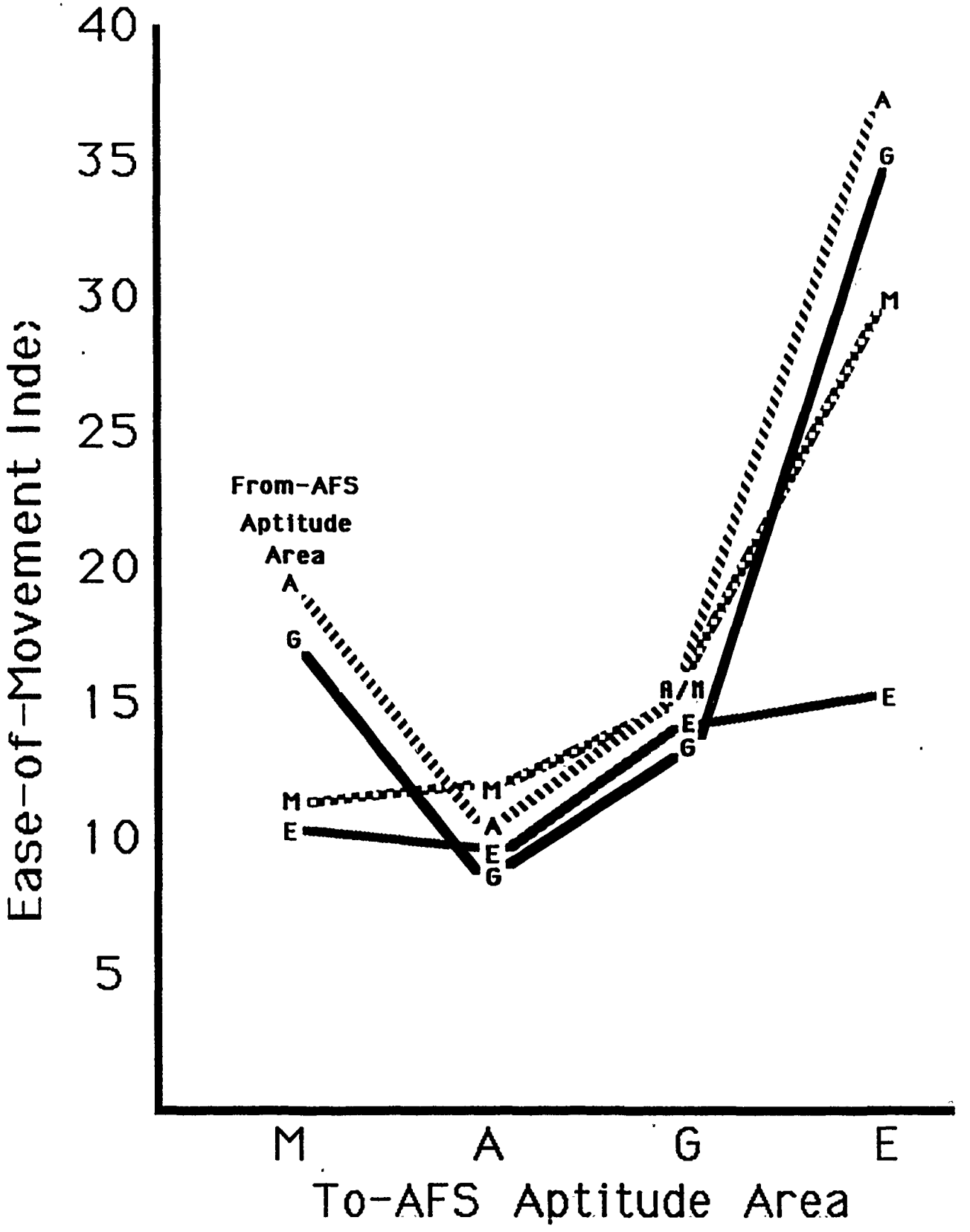


Figure 10 FROM-AFS x TO-AFS MAGE Area ANOVA Results for HRL-EOM Indicies.

Cohen, 1975; Lance, 1988) by first regressing HRL-EOM indices on TO-AFS OLDs and the binary MAGE-F/T which indicated whether movement was within or across MAGE areas. In a second step the cross-product between the TO-AFS OLD and MAGE-F/T variables was added. A significant increment in the R^2 ($\Delta R^2 = .040$, $F(1,2765) = 136.18$, $p \leq .01$) supported the presence of a significant TO-AFS OLD x MAGE-F/T interaction in the prediction of HRL-EOM indices. Subgroup regressions of the relationship between HRL-EOM indices and TO-AFS OLDs for movements across different MAGE areas and within the same MAGE area are shown in Figure 11. Results supported hypotheses and were consistent with those reported earlier by Lance et al. (1989a): there was a strong relationship between HRL-EOM indices and the learning difficulty of the TO-AFS (TO-AFS OLD) but only for movements across different MAGE areas.

Finally, the convergent validity of HRL-EOM indices was assessed against EOM indices developed earlier by Lance et al. (1989a) (OMC-EOM indices) which were based on the OMC enlisted task taxonomy (Bell & Thomasson, 1984). Recall earlier that one of the criteria for selecting AFSs to be surveyed with the 1989 SKQ was that they were also surveyed by the 1988 version of the SKQ. A total of 16 AFSs (indicated with asterisks in Table 2) were surveyed both with the 1988 and 1989 versions of the SKQ. Consequently, ease-of-movement indices had been calculated for the same 240 movements across these 16 AFSs both from the HRL task taxonomy (HRL-EOM indices) and the OMC task taxonomy (OMC-EOM indices). Although there were several similarities between

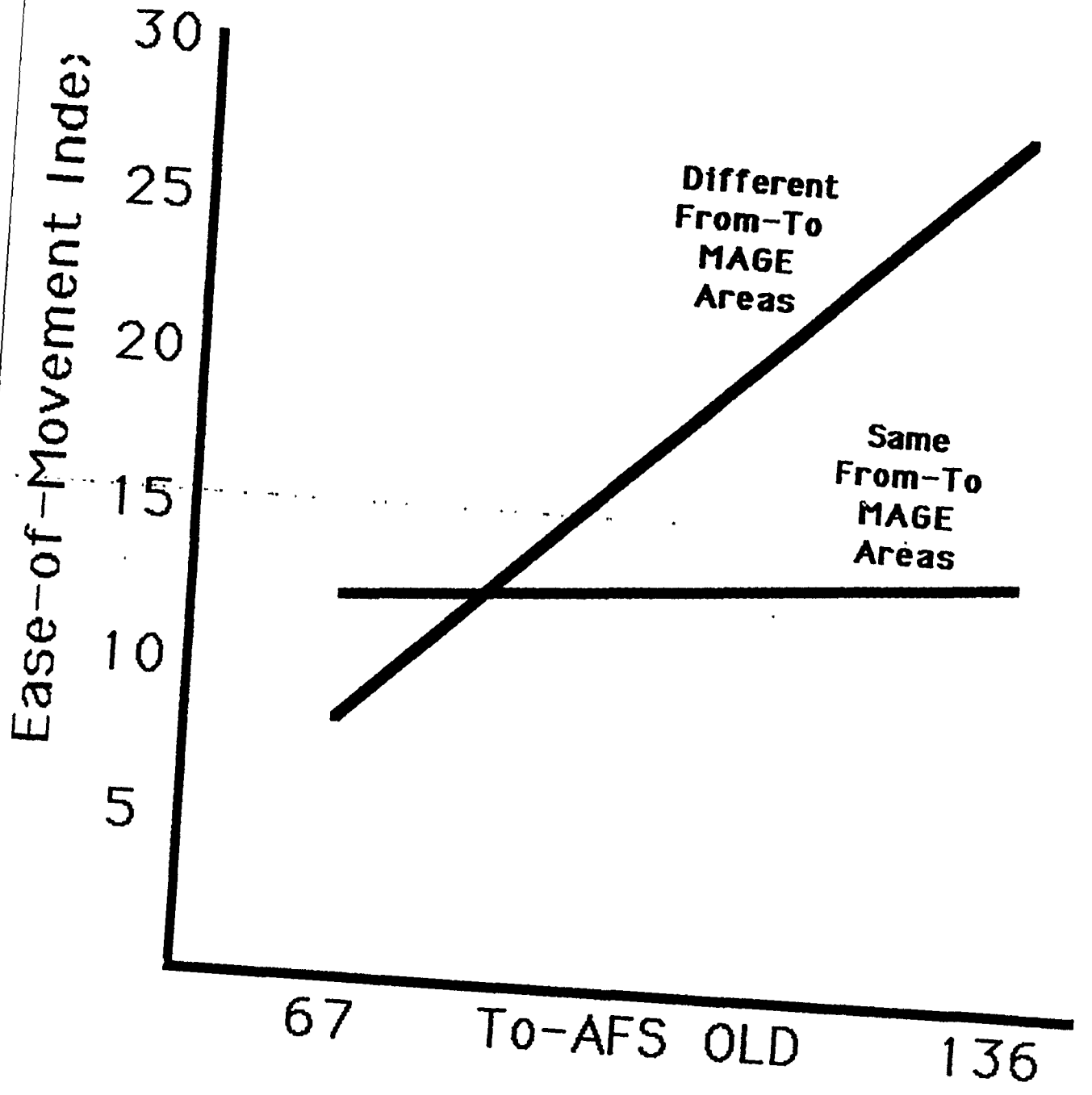


Figure 11 Subgroup Regression Results for TO-AFS OLD - HRL-EOM Index Relation.

the two taxonomies, there were also substantial differences (see Lance et al., 1989b). The correlation between HRL-EOM indices developed here and the OMC-EOM indices developed earlier by Lance et al. (1989a) was $r = .613$ ($p \leq .01$). This indicates significant, but far from perfect, agreement in rank-ordering cross-AFS movements according to their retraining difficulty. Because these two sets of EOM indices were computed (a) for movements across identical AFSs, (b) from identical months to proficiency rating scales, (c) from rating questionnaires containing substantially similar rating instructions, differences between the HRL-EOM and OMC-EOM indices can be attributed primarily to differences in the supporting task taxonomies.

Summary

HRL-EOM indices exhibited significant convergent validity against (a) occupational learning difficulties of the FROM- and TO-AFSs referenced in moving from AFS$_i$ to AFS$_j$, (b) differences in FROM- and TO-AFSs' aptitude requirements, and (c) EOM indices developed earlier on the basis of the OMC task taxonomy (OMC-EOM indices, Lance et al., 1989a).

V. RESULTS III: POSTDICTIVE VALIDATION OF EASE-OF-MOVEMENT INDICES AGAINST EASE-OF-MOVEMENT CRITERIA

The third goal of this research was to use a postdictive validation design to link EOM estimates to criteria related to the actual ease with which individuals were able to retrain into another AFS. As explained earlier, this was accomplished using data collected as part of an Airman Retraining Program Survey (ARPS) conducted in the late 1970's and made available for the

present research by AFHRL/MO. A subset of the ARPS data were selected in which retrainees' FROM-AFSCs and TO-AFSCs matched AFSCs of specialties surveyed either by the 1988 or 1989 version of the SKQ. Using over 7,000 lines of SPSS-X code, appropriate HRL-EOM indices, OMC-EOM indices, or both, depending on whether the retrained airman's FROM-AFSC and TO-AFSC matched AFSCs of specialties surveyed by the 1988 or 1989 SKQ, were written to the subset of the ARPS data selected for this study.

The primary hypothesis tested was that EOM indices would correlate with criteria related to the actual ease with which retrained airmen learned to perform their new jobs effectively. As indicated above, three categories of criteria were included. The first, retrainees' and their supervisors' judgments of the extent to which the retrainees' previous technical skills were helpful in performing the retrainees' new jobs, is related to a global assessment of the transferability of the retrainees' skills from the old to the new assignment. The prediction was that these judgments would correlate negatively with both HRL-EOM indices and OMC-EOM indices, since both EOM indices actually reflect difficulty of retraining.

The second category of criteria included retrainees' self-ratings and their supervisors' ratings of retrainee technical ability and job skills. These criteria were also expected to correlate negatively with EOM indices because retrainees' current technical skills and abilities should have been augmented to the extent that training and experience in the previous (FROM-) AFS were relevant to performing the current job.

Finally, retrainees' supervisors provided two global ratings of retrainee quantity and quality of performance. As indicated earlier, these two ratings were highly correlated and consequently were summed to form a composite supervisor performance rating criterion. A negative correlation was also expected between this criterion and EOM indices because it was expected that performance would be augmented to the extent that training and experience in the previous (FROM-) AFS prepared the retrainee to perform effectively in the new (TO-) AFS.

Results are shown in Table 13. Correlations of HRL-EOM indices with criteria all were statistically significant and in the predicted direction, supporting hypotheses. However, correlations were small in absolute magnitude. Results were not as encouraging for OMC-EOM indices. Only two of five correlations between OMC-EOM indices and criteria reached conventional levels of statistical significance, and these were opposite in sign to the hypothesized relationship between EOM indices and criteria.

Summary

Negative correlations were predicted between EOM criteria related to the ease with which retrained airmen learned to perform in their new AFS, and HRL-EOM and OMC-EOM indices. Hypotheses were disconfirmed for OMC-EOM indices. Results confirmed predictions for HRL-EOM indices but the absolute magnitudes of the correlations were small.

Table 13

Correlations between EOM Indices and EOM Criteria

EOM Criteria	Ease-of-Movement Indices: HRL-EOM Index	Ease-of-Movement Indices: OMC-EOM Index
Skill Utilization:		
Previous Tech Skill Utilization: Retrainee Rating	-.141** (2640)	.085** (918)
Previous Tech Skill Utilization: Supervisor Rating	-.111** (2588)	.043 (957)
Current Skills/Ability:		
Technical Ability: Retrainee Rating	-.066** (2454)	.071* (918)
Retrainee Job Skills: Supervisor Ratings	-.070** (2592)	.051 (961)
Retrainee Performance:		
Retrainee Performance: Supervisor Ratings	-.048** (2592)	.026 (960)

*$p < .05$; **$p < .01$

Note. Sample sizes are shown in parentheses.

VI. RESULTS IV: RELATIONSHIPS AMONG INDICATORS OF EASE-OF-MOVEMENT PREDICTOR AND CRITERION CONSTRUCTS

Assessing the construct validity of measures is a complex effort that requires substantive theory development and multiple lines of empirical evidence obtained in a hypothesis testing framework (Landy, 1986; Landy & Vasey, 1984). One source of evidence for the construct validity of measures is that the measures relate to one another as the underlying constructs theoretically relate among themselves (Binning & Barrett, 1989; Cook & Campbell, 1979; Lance, Hedge, & Alley, 1987). The purpose of this section is to extend results in the previous section which linked EOM indices to criteria related to actual EOM, to examinations of additional relationships between indicators of EOM predictor and criterion constructs.

Kavanagh and Gould (1988, 1989) identified factors hypothesized to determine cross-job EOM, defined in terms of the ease with which individuals can apply previously acquired knowledge and skills in learning to perform a new job effectively. Among these are social and motivational factors, interjob similarity (i.e., between the FROM- and TO- jobs) in task content, the learning difficulty of the new (TO-) job, and interjob similarity in aptitude requirements. Although social and motivational factors have been shown to relate to adjustment following job transfer (e.g., Brett, 1982, 1984; Pinder & Schroeder, 1987; Pinder & Walter, 1984) examination of these factors' relations with indicators of cross-job EOM was beyond the scope of this study. Rather, relations of EOM criteria with

other putative determinants of EOM identified by Kavanagh and Gould (1989) were examined using the subset of the ARPS data described in the previous section.

Besides the HRL-EOM and OMC-EOM indices, two additional variables were hypothesized to relate negatively with EOM criteria. First, as discussed by Kavanagh and Gould (1989), the learning difficulty of the TO-AFS generally should relate negatively to the ease of cross-job transfer. Consequently, it was expected that TO-AFS OLDs would correlate negatively with EOM criteria. Second, as explained above, and by Lance et al. (1989a), cross-job transfer should be easier if the new job draws upon similar, rather than different, sets of aptitudes compared to the old job. Thus, negative relationships were also expected between EOM criteria and the binary MAGE-F/T variable which indicated whether ARPS retrainees had retrained across AFSs within the same MAGE area or across different MAGE areas.

On the other hand, positive relationships were expected between EOM criteria and retrainee general cognitive ability and the learning difficulty of the old (FROM-) AFS. Results of hundreds of test validity studies support the idea that general cognitive ability tests effectively predict job performance (e.g., Hunter & Hunter, 1984; Schmitt, Gooding, Noe, & Kirsch, 1984). In the present study it was also expected that retrainee general cognitive ability, as indexed by retrainees' standardized AFQT scores, would relate positively to supervisor ratings of retrainee performance level, as well as to other EOM criteria. Finally, as Lance et al. (1989a) suggested, the learning

difficulty of the previous job (FROM-AFS) might be expected to relate positively to ease-of-movement because previous experience in more difficult jobs might be more broadly applicable in learning a new job than if the previous job is a less difficult one. Thus, a positive relationship was predicted between EOM criteria and FROM-AFS OLDs.

Finally, three categories of EOM criteria have been discussed so far, each of which relates to cross-job ease-of-movement construct somewhat differently. The first, retrainees' and their supervisors' judgments of the extent to which previous knowledge and skills were helpful in the new (TO-) AFS perhaps bears the closest relationship to "cross-job ease-of-movement" as defined in this report. The second and third categories of criteria, retrainees' and their supervisors' ratings of retrainee use of current skills and abilities, and supervisor ratings of ratee performance level, are perhaps better viewed as reflecting outcomes or consequences of retraining ease, rather than cross-job ease-of-movement per se. Nonetheless, it was expected that since all criterion variables relate to cross-job ease-of-movement, they would be positively intercorrelated.

Correlational results are shown in Table 14. For comparison, results shown previously in Table 13 (correlations between HRL-EOM and OMC-EOM indices and EOM criteria) are also shown in Table 14. Results were inconclusive in establishing the construct validity of EOM predictors or criteria. First, in addition to the HRL-EOM and OMC-EOM indices, it was predicted that the TO-AFS OLD and MAGE-F/T variables would relate nega-

tively with EOM criteria. Results supported predictions for the MAGE-F/T variable for the two EOM criteria reflecting retrainees' and their supervisors' global judgments of cross-AFS technical skill transfer. In effect, results support the idea that technical skills are more easily transferable across jobs within, rather than across, MAGE aptitude areas. Counter to expectations, TO-AFS OLDs were positively, rather than negatively, related to retrainees' and their supervisors' global judgments of cross-AFS technical skill transfer. In effect, these unexpected results suggested that individuals' technical skills are more easily transferable into more difficult jobs than they are into less difficult ones. One possible post hoc interpretation for these results is that more difficult TO-AFSs presented a greater opportunity to apply previously acquired knowledge and skills than if retraining was into a less demanding AFS.

It was also hypothesized that FROM-AFS OLDs and retrainees' AFQT scores would relate positively to EOM criteria. Again, results were mixed. FROM-AFS OLDs correlated positively with EOM criteria reflecting utilization of prior skills, and indicated a tendency for movements from more difficult jobs to be easier than from less difficult ones. On the other hand, statistically significant negative, rather than positive, correlations were obtained between retrainees' AFQT scores and three of the EOM criterion measures. That is, there was a tendency for higher-ability retrainees to perform less effectively in their new assignments than lower-ability retrainees. These results were unexpected.

Table 14

Ease-of-Movement Predictor-Criterion Relationships

	EOM Criteria				
EOM Predictors:	6	7	8	9	10
1. OMC-EOM Index	.085** (918)	.043 (957)	.071* (918)	.051 (961)	.026 (960)
2. HRL-EOM Index	-.141** (2640)	-.111** (2588)	-.066** (2454)	-.070** (2592)	-.048** (2592)
3. MAGE F/T	-.174** (2790)	-.135** (2935)	-.021 (2785)	-.026 (2940)	-.028 (2939)
4. From-AFS OLD	.036* (2790)	.034* (2935)	.027 (2785)	.011 (2940)	.001 (2939)
5. To-AFS OLD	.113** (2790)	.118** (2935)	-.000 (2785)	-.006 (2940)	-.006 (2939)
6. Retrainee AFQT	-.124** (2790)	-.061** (2935)	.036* (2793)	-.106** (2793)	-.013 (2939)
EOM Criteria:					
Skill Utilization:					
6. Previous Tech Skill Utilization: Retrainee Rating	1.000				
7. Previous Tech Skill Utilization: Supervisor Rating	.296** (2601)	1.000			
Current Skills/Ability:					
8. Technical Ability: Retrainee Rating	.129** (2779)	.091** (2597)	1.000		
9. Retrainee Job Skills: Supervisor Ratings	.052** (2607)	.263** (2930)	.188** (2603)	1.000	
Retrainee Performance:					
10. Retrainee Performance: Supervisor Ratings	.032* (2605)	.239** (2930)	.179** (2601)	.811** (2937)	1.000

Note. Sample sizes are shown in parentheses. *$p < .05$; **$p < .01$

Patterns of correlations among EOM criteria were substantially consistent with expectations. First, all correlations among EOM criterion measures were significantly intercorrelated and positively signed (see Table 14). This evidence is consistent with the idea that all measures were tapping the same underlying ease-of-movement construct. However, the absolute magnitudes of many of these criterion intercorrelations were small, suggesting either that (a) the measures were unreliable, (b) in fact they reflect different underlying constructs, and/or (c) they reflect different aspects of the same construct. Of course, is was not possible to decide among these interpretations simply on the basis of the observed criterion intercorrelations. Correlations between cross-source ratings of the same EOM criterion (i.e., retrainee vs. supervisor ratings of previous skill utilization, and use of current skills and ability) reflect levels of cross-source agreement, and observed correlations (rs = .296 and .188, see Table 14) were consistent with prior research on cross-source convergence in performance rating (Dickinson, Flynn, Hassett, & Tannenbaum, 1985; Harris & Schaubroeck, 1988). Finally, same-source different-criterion category correlations (e.g., retrainee judgments of previous skill utilization vs. current use of technical ability) generally were low, indicating that raters distinguished between different aspects of the criterion space. One exception to this pattern is the high correlation between supervisors' ratings of retrainee job skills and retrainee performance (r = .811), but a case could be made that both of

these ratings reflect an assessment of retrainee performance effectiveness.

Summary

In an earlier section (see RESULTS II), evidence was reported that supported the convergent validity of HRL-EOM and OMC-EOM indices against other putative predictors of cross-job ease-of-movement. In this section, an initial attempt was made to extend these findings to examination of the construct validity of EOM predictors and criteria within the context of a theoretical framework suggested by Kavanagh and Gould (1988, 1989). Intercorrelations among EOM criteria suggested some convergence upon a common ease-of-movement construct, but observed predictor - criterion relationships did not consistently support hypothesized linkages.

For example some evidence indicated that lower ability retrainees, rather than higher ability retrainees, performed more effectively in the retraining AFS, and other evidence suggested that retraining was generally easier into more difficult, rather than into less difficult, AFSs. These results were counter to predictions but perhaps could be explained post hoc. For example, lower ability airmen may have (a) disqualified in their previous (FROM-) AFS, and (b) retrained into less difficult AFSs, more often than higher ability airmen, while higher ability airmen may have more often retrained into more difficult AFSs than the FROM-AFS. Thus, the negative relationship between retrainee ability and performance in the retraining AFS could reflect self- or USAF-selection of low ability airmen into low

difficulty AFSs and high ability airmen into high difficulty AFSs. The positive relationship between retrainee performance and TO-AFS OLDs may reflect greater opportunity for prior skill utilization (as conjectured above), or may reflect raters' assumptions that performance effectiveness covaries with the difficulty of the work performed. Of course, these explanations are speculative.

In general, correlational evidence on the construct validity of EOM predictors and criteria examined in the present research suggests (a) fairly strong support for convergence among EOM predictors, (b) moderate support for convergence among EOM criteria, and (c) tentative support for EOM predictor - criterion linkages.

VII. DISCUSSION

Among the challenges facing the human resource manager of the 1990's and beyond are (a) declining numbers of skilled workers entering the U.S. labor force, (b) increasingly rapid technological changes in the workplace, (c) increasing demands for strategic human resource planning, and (d) increasing demands for accountability, that is, demonstrating that human resource management interventions yield economic payoffs (Ahlburg & Kimmel, 1986; Algera & Koopman, 1989; Baird & Meshoulam, 1988; Burak, 1985; Dyer, 1982; Fossom et al., 1986; Gentner, 1989; Gott, 1988; Gould, 1989; Kozlowski & Farr, 1988; Lengnick-Hall & Lengnick-Hall, 1988; Steffy & Maurer, 1988; Tornatzky, 1985). One of the specific challenges to HRM will be to match workforce knowledge and skills and organizational skill mix requirements.

The present study was designed to support research on the cross-job transferability of skills as one key component to job retraining for addressing skill mix imbalances (Mitchell & Oltrogge, 1989).

To date, little research exists on cross-job ease-of-movement (Lance et al., 1989a). The emerging literature on this topic suggests that (a) interjob similarity relates positively to cross-job EOM (Gordon et al., 1986; Gordon & Fitzgibbons, 1982; Mitchell & Oltrogge, 1989), (b) task taxonomies can be developed to support cross-job EOM estimates (Lance et al., 1989b), and (c) cross-job EOM indices which have been developed converge with other putative determinants of actual EOM (Lance et al., 1989a). The present study sought to extend research in this area by (a) validating the usefulness of an enlisted AFS task taxonomy described by Lance et al. (1989b) to support cross-job EOM research, (b) providing additional convergent validation evidence for EOM indices against additional EOM predictors, (c) validating EOM indices against criteria related to actual EOM across a number of Air Force enlisted specialties, and (d) attempting an initial assessment of the construct validity of indicators of EOM predictor and criterion constructs.

Several conclusions can be drawn from results presented above. First, the taxonomy of enlisted AFS tasks developed at AFHRL (HRL taxonomy) was useful in the present research for (a) obtaining reliable SME part-of-job, relative time spent, and months to proficiency judgments relating to tasks performed within the HRL task taxonomy categories, (b) differentiating

among AFSs on the basis of their task contents and learning times, (c) clustering AFSs into relatively homogeneous job groupings, and (d) supporting the development of cross-AFS EOM indices. However, some limitations to the HRL taxonomy include (a) several task categories (e.g., Fabric/Rope Work) were infrequently endorsed as being "part of" the AFSs surveyed by the 1989 SKQ, (b) interrater reliability for some of the HRL task categories was low, and (c) there may exist relevant task domains in the actual enlisted AFS structure that are not specified by the HRL taxonomy. Together, this supporting evidence, along with potential limitations of the HRL taxonomy, reinforce Lance et al.'s (1989b) claim that <u>any</u> task taxonomy should be considered as tentative in the light of (a) future research findings, (b) alternative uses of the taxonomy, and (c) actual changes in the nature of tasks performed in the relevant work domain.

Second, results support the feasibility of developing estimates of cross-job EOM from SME estimates of learning times for tasks within taxonomic task categories. Results for EOM estimates developed on the basis of a taxonomy developed at the USAF Occupational Measurement Center (OMC taxonomy, Bell & Thomasson, 1984) were substantially replicated in the present research for HRL-EOM indices: HRL-EOM indices indicated more difficult cross-AFS movements (a) <u>into</u> AFSs having higher occupational learning difficulty indices (OLDs), (b) <u>from</u> AFSs having lower OLDs, and (c) across, rather than within, MAGE aptitude requirement areas.

Third, EOM indices appear to hold promise for predicting, at

the job or AFS level of analysis, the actual ease with individuals are able to retrain and learn to perform a new job effectively. One of the major goals of this research was to link AFS-level EOM indices to criteria related to actual retraining ease, and some evidence presented above supported the idea that EOM indices (at least the HRL-EOM indices) were useful predictors. However, there are several potential limitations to the postdictive validation strategy adopted here, and to the criterion measures themselves. First, and most obviously, ten years had intervened between the collection of the ARPS data (containing criterion measures) and the 1989 SKQ data (which furnished the basis for developing HRL-EOM indices). During this time, the Airman Classification Structure had undergone a number of modifications so that (a) only about half of the original ARPS sample had FROM-AFSCs and TO-AFSCs that matched those in the current Airman Classification Structure, and (b) job duties even in AFSs having the same AFSC at both data collection times may have changed appreciably (e.g., with the widespread adoption of computerized technologies). Relatedly, there was no assurance that the HRL taxonomy, developed in 1989, was as useful for describing enlisted AFS tasks that existed in 1979.

Perhaps the most important limitations to the postdictive validation of EOM indices against ARPS data concern the criterion measures themselves. First, most of the criterion measures were single-item rating measures which may be subject to response biases and measurement unreliability (however, see Ganster, Hennessey, & Luthans, 1983; Larsen, Diener, & Emmons, 1985;

Scarpello & Campbell, 1983; and Spector, 1987 for alternative views). But perhaps the most important limitation concerns the construct validity of the criterion measures adopted from the ARPS survey data for the present research. The endogenous variable in ease-of-movement models presented by Kavanagh and Gould (1988, 1989) is the ease with individuals can retrain and learn to perform a new job effectively. The criteria adopted for the present study relate theoretically to this construct, but indirectly. Retrainees' and their supervisors' ratings of the extent to which previous skills were useful in the retraining (TO-) AFS represent global estimates of retraining ease but (a) a large proportion of these ratings were of retraining that had occurred several years earlier, (b) the ratings do not directly reflect the time required to learn to perform the new job proficiently, and (c) these ratings are not directly tied to the level of retrainee performance effectiveness in the new AFS. The remaining ratings of retrainee performance quantity and quality, and of current utilization of skills and ability (a) are more closely related to potential outcomes of retraining ease, rather than ease-of-movement per se, and (b) are based on a large proportion of airmen who had been in their TO-AFS longer than three years (see Table 5). Thus these latter performance criteria likely reflect the influences of performance determinants whose importance extends beyond the transfer of previously learned job skills. In short, the EOM criteria adopted for postdictive validation of EOM indices in the present study should be (a) suspected of significant criterion

contamination (Brogden & Taylor, 1950), due to the likely influences of additional performance determinants, (b) suspected of significant criterion deficiency (Brogden & Taylor, 1950), because the theoretical criterion - ease of retraining in a new assignment, was not measured directly, and (c) interpreted as providing a very conservative, lower-bound, estimate of the predictive efficiency of EOM indices in forecasting actual transferability of skills.

The fourth general conclusion suggested by results presented above is that neither the construct validity of EOM predictors nor of criteria has been firmly established. This is the second study in which significant convergence has been demonstrated between EOM indices and alternative predictors of actual ease-of-movement. However, evidence on the criterion side is more tentative, and evidence on linkages between EOM predictors and criteria is even weaker. The last two conclusions point to the most pressing needs for future research on cross-job transferability of skills.

VIII. FUTURE RESEARCH NEEDS

Three of the most important research needs are (a) for the development of appropriate criterion measures to assess cross-job transferability of skills, (b) for longitudinal collection of performance data prior to, during, and subsequent to enlisted airman retraining, and (c) predictive validation of AFS-, job-, and individual-level cross-job ease-of-movement estimates against actual retraining ease.

A research strategy to address the first of these needs may borrow both technologically and conceptually from an ongoing stream of research on the measurement of job performance (Hedge & Teachout, 1986; Kavanagh, Borman, Hedge, & Gould, 1986, 1987). The original goals of the job performance measurement (JPM) project were to develop prototype procedures for the measurement of enlisted personnel job performance, and to link enlistment standards to job performance criteria (Wigdor & Green, 1986). Subsequently, JPM technology has been adapted for the evaluation of training outcomes (e.g., Blackhurst, Ballentine, & Pellum, 1987). The feasibility of adapting JPM technologies and on-line training evaluation measures (e.g., end-of-course examinations) for the purposes of assessing cross-job transferability of skills in terms of the time to learn new job skills, and in terms of time required to become proficient in the new job following retraining (Pinder & Schroeder, 1987) should be investigated.

Assuming that appropriate measures can be developed to assess the job performance of airmen undergoing retraining, the next important research need is for longitudinal assessment of job performance indicators prior to, during, and following retraining into the new AFS. Longitudinal measurement of performance indicators is particularly important for research on transferability of skills because (a) effects of prior performance effectiveness on subsequent performance in the new AFS should be quantified, (b) it should be determined whether EOM predictors (e.g., EOM indices such as the HRL-EOM indices developed above) are more useful in predicting learning times

during retraining, the transfer of learning outcomes to on-the-job performance, or the maintenance of skills acquired during training, and (c) of statistical advantages in controlling for extraneous, unmeasured effects on criterion variables (e.g., Johnston, 1984; Lance & Cornwell, 1988).

Third, the predictive efficiency of EOM indices should be evaluated at multiple levels of analysis. The OMC-EOM and HRL-EOM indices described above are specialty-level EOM indices in that they are designed to assess the average ease with which individuals retrain from AFS $_i$ and become proficient in AFS $_j$. Research reported here and by Lance et al. (1989a) support the general feasibility of estimating EOM at the specialty-level of analysis, and validated specialty-level EOM indices may prove useful, for example, as a component of the USAF's retraining Person-Job Match (PJM) algorithm (Roberts & Ward, 1982), or as input to decisions for consolidating AFSs. However, EOM can also be conceptualized at more micro levels such as the job- or position-level, or even at the level of the individual. For example, the potential transferability of incumbent $_k$'s skills in AFS $_i$ to an array of AFSs $_{j=1,...J}$ could be evaluated either on the basis of the average transferability of AFS $_i$ incumbents' skills to the J other AFSs, or on the basis of the skills that incumbent $_k$ has specifically acquired during his or her tenure in AFS $_1$. Cross-level and multi-level research designs (Dansereau & Markham, 1987; Rousseau, 1985) are becoming more widely used in organizational research (see, for example, Kohler & Mathieu, 1989 and Lance, Hedge, & Alley, in press), and hold promise for

supporting cross-disciplinary research for cases in which different supporting disciplines have traditionally taken macro or micro orientations on similar research problems.

Future research should also investigate convergence among alternative procedures for estimating cross-job transferability of skills. For example, Gould et al. (1989) described one procedure based on analysis of the learning difficulties of tasks allocated to task categories by SMEs. Also, Short, Ware, Archer, and Kavanagh (1989) described an alternative method for estimating ease-of-movement based on SMEs' and novices' global estimation of retraining times from examining AFSs' job descriptions. Still other means for estimating cross-AFS transferability of skills could be based on job overlap as indicated by (a) general job analysis instruments such as the General Work Inventory (GWI, Cunningham & Ballentine, 1982), or (b) principles inventories such as the Electronics Principles Inventory (EPI, Ruck, 1977, 1986). Future research should address (a) the extent to which EOM estimates derived on the basis of different data bases or estimation procedures converge in estimating transferability of skills, (b) whether different EOM estimation procedures produce estimates of transferability of skills that are differentially useful for different purposes (e.g., consolidating AFSs, consolidating training courses, forecasting future MPT requirements for emerging WSs), and (c) the validities of alternative EOM estimation procedures in predicting actual retraining ease.

Summary

Results of the present research indicate (a) reasonably strong support for the usefulness of an enlisted task taxonomy developed at AFHRL (HRL task taxonomy) for supporting research on cross-AFS transferability of skills, (b) convergent validity of EOM estimates based on the HRL task taxonomy against alternative predictors of actual ease-of-movement, and (c) potential value of EOM indices for predicting actual retraining ease. Future research needs include (a) development of validated measures of the ease with which individuals learn to perform a new job effectively, (b) predictive, longitudinal validation of EOM indices against actual retraining ease, and (c) convergent validation of alternative EOM estimates.

IX. REFERENCES

Ahlburg, D. A., & Kimmel, L. (1986). Human resources management implications of the changing age structure of the U.S. labor force. Research in Personnel and Human Resources Management, 4, 339-374.

Algera, J. A., & Koopman, P. L. (1989). Coping with new technology: Central issues in perspective. Applied Psychology: An International Review, 38, 1-13.

Alley, W. E., Treat, B. R., & Black, D. E. (1988). Classification of Air Force jobs into aptitude clusters. (AFHRL-TR-88-14). Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks AFB, TX.

Anderson, J. C., Milkovich, G. T., & Tsui, A. A. (1981). A model of intra-organizational mobility. Academy of Management Review, 6, 529-538.

Archer, W. B., & Gould, R. B. (1989, August). Analysis tools under development to support MPT decisions. Paper presented at the meeting of the American Psychological Association, New Orleans, LA.

Armor, D. J. (1989, August). Applied psychology and defense manpower policy. Paper presented at the meeting of the American Psychological Association, New Orleans, LA.

Askren, W. B., & Eckstrand, G. A. (1980). Human resource considerations from concept through deployment. Defense Management Journal, 16(2), 12-19.

Baird, L., & Meshoulam, I. (1988). Managing two fits of strategic human resource management. Academy of Management

IX. REFERENCES

Ahlburg, D. A., & Kimmel, L. (1986). Human resources management implications of the changing age structure of the U.S. labor force. Research in Personnel and Human Resources Management, 4, 339-374.

Algera, J. A., & Koopman, P. L. (1989). Coping with new technology: Central issues in perspective. Applied Psychology: An International Review, 38, 1-13.

Alley, W. E., Treat, B. R., & Black, D. E. (1988). Classification of Air Force jobs into aptitude clusters. (AFHRL-TR-88-14). Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks AFB, TX.

Anderson, J. C., Milkovich, G. T., & Tsui, A. A. (1981). A model of intra-organizational mobility. Academy of Management Review, 6, 529-538.

Archer, W. B., & Gould, R. B. (1989, August). Analysis tools under development to support MPT decisions. Paper presented at the meeting of the American Psychological Association, New Orleans, LA.

Armor, D. J. (1989, August). Applied psychology and defense manpower policy. Paper presented at the meeting of the American Psychological Association, New Orleans, LA.

Askren, W. B., & Eckstrand, G. A. (1980). Human resource considerations from concept through deployment. Defense Management Journal, 16(2), 12-19.

Baird, L., & Meshoulam, I. (1988). Managing two fits of strategic human resource management. Academy of Management

Review, 13, 116-128.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. Personnel Psychology, 41, 63-105.

Bell, J., & Thomasson, M. (1984). Job categorization project. Randolph AFB, TX: Occupational Analysis Program, United States Air Force Occupational Measurement Center.

Bennett, C. A. (1971). Toward an empirical, practicable, comprehensive task taxonomy. Human Factors, 13, 229-235.

Binning, J. F., & Barrett, G. V. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. Journal of Applied Psychology, 74, 478-494.

Blackhurst, J. L., Ballentine, R. D., & Pellum, M. W. (1987). Air Force job performance measurement technology applied to training. In H. G. Baker & G. J. Laabs (Eds.), Proceedings of the Department of Defense/Educational Testing Service conference on job performance measurement technologies. Washington, DC: Office of the Assistant Secretary of Defense.

Brett, J. M. (1982). Job transfer and well being. Journal of Applied Psychology, 67, 450-463.

Brett, J. M. (1984). Job transitions and personal and role development. Research in Personnel and Human resources Management, 2, 155-185.

Brogden, H. E., & Taylor, E. K. (1950). The theory and classification of criterion bias. Educational and

Psychological Measurement, 10, 159-186.

Burak, E. H. (1985). Linking corporate business and human resource planning: Strategic issues and concerns. Human Resource Planning, 8, 133-145.

Burak, E. H., & Mathys, N. J. (1980). Career management in organizations: A practical human resource planning approach. Lake Forest, IL: Brace-Park.

Burtch, L. D., Lipscomb, M. S., & Wissman, D. J. (1982). Aptitude requirements based on task difficulty: Methodology for evaluation. (AFHRL-TR-81-34). Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks AFB, TX.

Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245-276.

Christal, R. E. (1974). The United States Air Force occupational research project. (AFHRL-TR-73-75) Air Force Human Resources Laboratory, Occupational Research Division, Lackland AFB, TX.

Cohen, J., & Cohen, P. (1975). Applied multiple regression/ correlation analysis for the behavioral sciences. New York: Wiley.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design & analysis issues for field settings. Boston, MA: Houghton Mifflin.

Cunningham, J. W., & Ballentine, R. D. (1982). The general work inventory. Raleigh, NC: Authors.

Dansereau, F., & Markham, S. E. (1987). Levels of analysis in

personnel and human resources management. Research in Personnel and Human Resources Management, 5, 1-50.

Davis, P. A., Archer, W. B., Gould, R. B., & Kavanagh, M. J. (1989, April). Development of a cost effective methodology to estimate occupational learning difficulty. Paper presented at the meeting of the Society for Industrial and Organizational Psychology, Boston, MA.

Department of Defense (1984). Test manual for the armed services vocational aptitude battery. North Chicago, IL: United States Military Entrance Processing Command.

Dickinson, T. L., Flynn, J. B., Hassett, C. E., & Tannenbaum, S. I. (1985, August). A meta-analysis of multitrait-multimethod studies of performance ratings. Paper presented at the meeting of the American Psychological Association, Los Angeles, CA.

Dyer, L. (1982). Human resource planning. In K. M. Rowland & G. R. Ferris (Eds.), Personnel management. Boston, MA: Allyn and Bacon.

Dyer, L. (1985). Strategic human resources management and planning. Research in Personnel and Human Resources Management, 3, 1-30.

Fine, S. A. (1957a). A reexamination of "transferability of skills" - Part I. Monthly Labor Review, 80, 803-810.

Fine, S. A. (1957b). A reexamination of "transferability of skills" - Part II. Monthly Labor Review, 80, 938-948.

Fleishman, E. A. (1984). Systems for linking job tasks to personnel requirements. Public Personnel Management

Journal, 13, 395-408.

Fleishman, E. A., & Quaintance, M. K. (1984). Taxonomies of human performance. Orlando, FL: Academic Press.

Forbes, J. B. (1987). Early intraorganizational mobility: Patterns and influences. Academy of Management Journal, 30, 110-125.

Fossum, J. A., Arvey, R. D., Paradise, C. A., & Robbins, N. E. (1986). Modeling the skills obsolescence process: A psychological/economic integration. Academy of Management Review, 11, 362-374.

Ganster, D. C., Hennessey, H. W., & Luthans, F. (1983). Social desirability effects: Three alternative models. Academy of Management Journal, 26, 321-331.

Garcia, S. K., Ruck, H. W., & Weeks, J. (1985). Benchmark learning difficulty technology: Feasibility of operational implementation. (AFHRL-TP-85-33) Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks AFB, TX.

Gentner, F. C. (1989, August). Including MPT considerations early in the design process: User's perspective. Paper presented at the meeting of the American Psychological Association, New Orleans, LA.

Gordon, M. E., Cofer, J. L., & McCullough, P. M. (1986). Relationships among seniority, past performance, interjob similarity, and trainability. Journal of Applied Psychology, 71, 518-521.

Gordon, M. E., & Fitzgibbons, W. J. (1982). Empirical test of the validity of seniority as a factor in staffing decisions.

Journal of Applied Psychology, 67, 311-319.

Gott, C. D. (1988). Exploratory models to link job performance to enlistment standards. (AFHRL-TP-88-39) Brooks AFB, TX: Air Force Human Resources Laboratory, Training Systems Division.

Gould, R. B. (1989, August). Challenges of developing MPT tools and data bases: Researcher's perspective. Paper presented at the meeting of the American Psychological Association, New Orleans, LA.

Gould, R. B., Archer, W., Filer, J., Short, L. O., & Kavanagh, M. J. (1989, April). Development of a methodology to estimate common task overlap. Paper presented at the meeting of the Society for Industrial and Organizational Psychology, Boston, MA.

Hall, D. T. (1976). Careers in organizations. Pacific Palisades, CA: Goodyear.

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. Personnel Psychology, 41, 41-62.

Hedge, J. W., & Teachout, M. S. (1986). Job performance measurement: A systematic program of research and development. (AFHRL-TP-86-37) Brooks AFB, TX: Air Force Human Resources Laboratory, Training Systems Division.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.

James, L. R. (1973). Criterion models and construct validity for

criteria. Psychological Bulletin, 80, 75-83.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. Journal of Applied Psychology, 69, 85-98.

Johnston, J. (1984). Econometric methods. New York: McGraw-Hill.

Kavanagh, M. J., Borman, W. C., Hedge, J. W., & Gould, R. B. (1986). Job performance measurement classification scheme for validation research in the military. (AFHRL-TP-85-51) Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Kavanagh, M. J., Borman, W. C., Hedge, J. W., & Gould, R. B. (1987). Job performance measurement in the military: A classification scheme, literature review, and directions for research. (AFHRL-TR-87-15) Brooks AFB, TX: Air Force Human Resources Laboratory, Training Systems Division.

Kavanagh, M. J., & Gould, R. B. (1988). Task, job, manpower, skill, and training requirements for emerging technology in the Air Force. (Unpublished manuscript) Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Kavanagh, M. J., & Gould, R. B. (1989, April). Transferability matrix: Ease of movement across occupational classifications. Paper presented at the meeting of the Society for Industrial and Organizational Psychology, Boston, MA.

Kohler, S. S., & Mathieu, J. E. (1989, August). A cross-level

examination of group absence influences on individual absences. Paper presented at the meeting of the American Psychological Association, New Orleans, LA.

Kozlowski, S. W. J., & Farr, J. L. (1988). An integrative model of updating and performance. Human Performance, 1, 5-29.

Lahey, M. A., Downey, R. G., & Saal, F. E. (1983). Intraclass correlations: There's more there than meets the eye. Psychological Bulletin, 93, 586-595.

Lance, C. E. (1988). Residual centering, exploratory and confirmatory moderator analysis, and decomposition of effects in path models containing interactions. Applied Psychological Measurement, 12, 163-175.

Lance, C. E., & Cornwell, J. M. (1988, August). The job satisfaction - job performance relationship: New answers to an old question. Paper presented at the meeting of the American Psychological Association, Atlanta, GA.

Lance, C. E., Hedge, J. W., & Alley, W. E. (1987). Ability, experience, and task difficulty predictors of task performance. (AFHRL-TP-87-14) Brooks AFB, TX: Air Force Human Resources Laboratory, Training Systems Division.

Lance, C. E., Hedge, J. W., & Alley, W. E. (1989). Joint relationships of task proficiency with aptitude, experience, and task difficulty: A cross-level, interactional study. Human Performance, 2, 249-272.

Lance, C. E., Kavanagh, M. J., & Gould, R. B. (1989a). Development and convergent validation of cross-job ease-of-movement indices. Manuscript submitted for publication.

Lance, C. E., Kavanagh, M. J., & Gould, R. B. (1989b). Development of a task taxonomy for United States Air Force ease-of-movement research. Manuscript submitted for publication.

Landy, F. J. (1986). Stamp collecting versus science: Validation as hypothesis testing. American Psychologist, 41, 1183-1192.

Landy, F. J., & Vasey, J. (1984). Theory and logic in human resources research. Research in Personnel and Human Resources Management, 2, 1-34.

Larsen, R. J., Diener, E., & Emmons, R. A. (1985). An evaluation of subjective well-being measures. Social Indicators Research, 17, 1-17.

Lecznar, W. B. (1971). Three methods for estimating difficulty of job tasks. (AFHRL-TR-71-30) Air Force Human Resources Laboratory, Personnel Division, Lackland AFB, TX.

Lengnick-Hall, C. A., & Lengnick-Hall,M. L. (1988). Strategic human resources management: A review of the literature and a proposed typology. Academy of Management Review, 13, 454-470.

Lenway, S. A. (1985). The politics of U.S. international trade: Protection, expansion and escape. Boston, MA: Pitman.

Louis, M. R. (1980). Career transitions: Varieties and commonalities. Academy of Management Review, 5, 329-340.

Mead, D. F., & Christal, R. E. (1970). Development of a constant standard weight equation for evaluating job difficulty. (AFHRL-TR-70-44). Air Force Human Resources Laboratory,

Personnel Division, Lackland AFB, TX.

Mitchell, M. E., & Oltrogge, C. G. (1989, August). Retraining of nonexempt workers: Successful solution for job obsolescence. Paper presented at the meeting of the American Psychological Association, New Orleans, LA.

Mobley, W. H., Griffeth, R. W., Hand, H. H., & Meglino, B. M. (1979). Review and conceptual analysis of the employee turnover process. Psychological Bulletin, 86, 493-522.

Muchinsky, P. M., & Morrow, P. C. (1980). A multidisciplinary model of voluntary employee turnover. Journal of Vocational Behavior, 17, 263-290.

Mumford, M. D., Weeks, J. L., Harding, F. D., & Fleishman, E. A. (1987). Measuring occupational difficulty: A construct validation against training criteria. Journal of Applied Psychology, 72, 578-587.

Pinder, C. G., & Schroeder, K. G. (1987). Time to proficiency following job transfers. Academy of Management Journal, 30, 336-353.

Pinder, C. G., & Walter, G. A. (1984). Personnel transfers and employee development. Research in Personnel and Human Resources Management, 2, 187-218.

Ramage, J. A. (1987). Task learning difficulty: Interrelationships among aptitude-specific benchmarked rating scales. (AFHRL-TP-86-56) Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks AFB, TX.

Ree, M. J., Mathews, J. J., Mullins, C. J., & Massey, R. H. (1982). Calibration of the Armed Services Vocational

Aptitude Battery forms 8, 9, and 10. (AFHRL-TR-81-49) Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Roberts, D. K., & Ward, J. H. (1982). General purpose person-job match system for Air Force enlisted accessions. (AFHRL-SR-82-2) Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks AFB, TX.

Rousseau, D. M. (1985). Issues of level in organizational research: Multi-level and cross-level perspectives. Research in Organizational Behavior, 7, 1-37.

Ruck, H. W. (Ed.). (1977). The development and application of the electronics principles inventory. Lackland AFB, TX: USAF Occupational Measurement Center, Air training Command.

Ruck, H. W. (1986). Skill/knowledge commonalities in selected electronics specialties. (AFHRL-TP-86-20) Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks AFB, TX.

Ruck, H. W. (Chair) (1989, August). Decision support system for training: Interdisciplinary perspectives. Symposium conducted at the meeting of the American Psychological Association, New Orleans, LA.

Rumberger, R. W. (1981). The changing skill requirements of jobs in the U. S. economy. Industrial and Labor Relations Review, 34, 578-590.

Scarpello, V., & Campbell, J. P. (1983). Job satisfaction: Are all the parts there? Personnel Psychology, 36, 577-600.

Scarpello, V. G., & Ledvinka, J. (1988). Personnel/human

resource management. Boston, MA: PWS-Kent.

Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. Personnel Psychology, 37, 407-422.

Short, L. O., Ware, G., Archer, W., & Kavanagh, M. J. (1989, April). Development of a methodology to estimate retraining time. Paper presented at the meeting of the Society for Industrial and Organizational Psychology, Boston, MA.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing interrater reliability. Psychological Bulletin, 86, 420-428.

Skinner, M. J. (1981). An evaluation of the Air Force airman retraining program. Proceedings of the Military Testing Association, 1109-1120.

Skinner, M. J. (1982). Retrained airmen: Volunteers versus non-volunteers. Proceedings of the Military Testing Association, 645-651.

Skinner, M. J. (1983). Retraining program for Air Force enlisted personnel: An evaluation. (AFHRL-SR-83-31) Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks AFB, TX.

Skinner, M. J., & Alley, W. E. (1980). Performance of retrained airmen in Air Force technical schools. (AFHRL-TR-80-7) Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Skinner, M. J., & Alley, W. E. (1984). Job Aptitude requirement

waivers for retrained airmen. (AFHRL-TR-83-42) Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Spector, P. E. (1987). Method variance as an artifact in self-report affect and perceptions at work: Myth or significant problem. Journal of Applied Psychology, 72, 438-443.

Steffy, B. D., & Maurer, S. D. (1988). Conceptualizing and measuring the economic effectiveness of human resource activities. Academy of Management Review, 13, 271-286.

Stephenson, R. W., & Gentner, F. C. (1987). Manpower, personnel, training, and safety guidance and control for weapon system acquisitions. (AFHRL-TP-87-31) Brooks AFB, TX: Air Force Human Resources Laboratory, Special Projects Office.

Tornatzky, L. G. (1985, August). Technological change and the structure of work. Paper presented at the meeting of the American Psychological Association, Los Angeles, CA.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58, 236-244.

Weddle, P. D., & Fulkerson, G. D. (1980). Forecasting the human-resource costs in Navy weapon systems. Defense Management Journal, 16(2), 6-11.

Weeks, J. (1984). Occupational learning difficulty: A standard for determining the order of aptitude requirement minimums. (AFHRL-SR-84-26) Air Force Human Resources Laboratory, Manpower and Personnel Division, Brooks AFB.

Wigdor, A. K., & Green, B. F. Jr. (Eds.) (1986). Assessing the

performance of enlisted personnel: Evaluation of a joint-service research project. Washington, DC: National Academy Press.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. Psychological Bulletin, 99, 432-442.

Appendices can be obtained from
Universal Energy Systems, Inc.

1988-89 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Universal Energy Systems, Inc.

FINAL REPORT

Prepared by:	Thomas L. Landers, Ph.D., P.E.
Academic Rank:	Associate Professor
Department and	Industrial Engineering
University:	University of Arkansas
Research Location:	AFHRL/LRL
	Wright-Patterson AFB, OH 45459
USAF Researcher:	Captain Michael Hanuschik
Date:	7 July 1989
Contract No:	F49620-88-C-0053/SB5881-0378

ACKNOWLEDGEMENTS

The researcher wishes to recognize the valuable contributions by Air Force personnel in the accomplishment of this work. Captain Michael Hanuschik has provided support and guidance as Effort Focal Point in the Air Force Human Resources Laboratory. Personnel in the F-16 System Program Office (Aeronautical Systems Division) have been very cooperative in providing case study data for use in the project. Special thanks are due Gary Arnold and Ruth Brewer of ASD/YP.

PROPORTIONAL INTENSITY RELIABILITY ANALYSIS FOR REPAIRABLE ITEMS

by

Thomas L. Landers, Ph.D., P.E.

ABSTRACT

This research project investigated alternative methods of reliability modeling, including explanatory variables, for the case of repairable items. Two approaches were considered: the parametric estimators reported by Lawless (1981) and the semi-parametric estimators proposed by Prentice, Williams and Peterson (PWP, 1981). The central objective was to evaluate the robustness of the PWP model, for the case where the true underlying model is a Nonhomogeneous Poisson Process (NHPP) with power law intensity function. Simulation was used to generate data sets from known NHPPs with power law proportional intensity functions. The study included both increasing and decreasing rates of occurrence of failures (ROCOF). The alternative models were compared based on point estimates of the instantaneous times to successive failures. The measures of merit were relative bias, mean absolute deviation and mean squared error. The research demonstrated the potential of the PWP model for analysis

of reliability failure data on repairable items, particularly for the case of increasing ROCOF. The project also evaluated software for performing PWP analysis and defined the requirements for integrated software to support further research on the robustness of PWP estimates. The researchers also defined the knowledge base for an expert system to implement covariate modeling for repairable systems reliability. Future research should expand the range of cases investigated under the alternate hypothesis of a parametric power law NHPP. Data sets for other parameter values should be generated and analyzed, sample sizes should be varied further, and the case of time truncation should be addressed. The researchers also recommend implementation of the knowledge base, for repairable item reliability, in a prototype expert system.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

ABSTRACT

I. INTRODUCTION

II. OBJECTIVES

III. REVIEW OF LITERATURE

IV. APPROACH

V. ANALYSIS AND RESULTS

VI. RECOMMENDATIONS

VII. REFERENCES

APPENDIX

KNOWLEDGE BASE SPECIFICATION

I. INTRODUCTION

Reliability is an important engineering consideration throughout the life of systems. Most systems are designed to be repairable, yet most reliability theory and practice has concentrated on nonrepairable systems. The little work done in the reliability field, on the modeling of repairable systems, has been in applying stochastic point processes [2]. Stochastic point processes are parametric approaches used for probabilistic modeling of reliability failures [46].

Proportional Hazards

A class of statistical models called the proportional hazards models has generated much recent interest in the reliability field. Proportional hazards modeling is a powerful technique which can be used for examining reliability data sets where the failure data may be heterogeneous due to the presence of concomitant variables. These explanatory variables, also known as covariates, may represent the effects of differences in environment, design or operating conditions.

Proportional hazards modeling has mainly been applied in the medical field. However, recent attempts have been made to apply the proportional hazards modeling technique in reliability engineering. The papers in the reliability literature reporting attempts at proportional hazards modeling appear primarily in conference proceedings or internal reports (e.g., [3,10,16,21,22,35]). Bendell [6,7], Baxter, et al. [5],

Dale [15], Landers and Kolarik [27] and Wightman and Bendell [48] are the only known journal articles available dealing with engineering applications of proportional hazards modeling for reliability data.

Background

Under the AFOSR/UES Summer Faculty Research Program (SFRP) the researcher investigated an expert system approach for analysis of repairable systems reliability. The results of this research contribute to the knowledge base for an expert system which could assist in reliability data analysis and statistical modeling in a RAMCAD (reliability and maintainability in computer-aided design) environment.

An expert system which performs reliability data analysis and modeling requires decision rules to perform statistical tests. The outcome of a statistical test determines the next appropriate course of analysis [20]. One of the statistical tests incorporated into the expert system will be the Laplace statistic. The Laplace statistic tests for trends in the rate of failure occurrence. Ascher and Feingold [4] describe the Laplace statistic as testing the following hypothesis:

H_0: Homogeneous Poisson Process

H_a: Montonic Trend

The Laplace test statistic approaches the standard normal curve rapidly; thus, permitting the use of the standard normal tables.

If the Laplace statistic indicates rejection of the null hypothesis, then the reliability failure data exhibits a monotonic trend and must be modeled by some nonstationary model such as the Nonhomogeneous Poisson Process (NHPP). The resulting reliability model will provide valuable information, such as:

1) The probability of next failure

2) The rate of occurrence of failures

3) The expected number of failures in a given time interval

4) The instantaneous and cumulative mean times between failures (MTBFs), or expected times of successive failures.

Such information aids in predicting reliability, scheduling maintenance and managing spare parts inventories.

II. OBJECTIVES

The objective of the RIP was to extend the SFRP effort to consider the case of failure data from a repairable system, where explanatory variables are involved and the tests for trend (e.g., Laplace test [4] and Lewis-Robinson [31] test) have indicated a trend. The primary goal was to compare alternative modeling approaches, including the Lawless parametric estimates for a power law Nonhomogeneous Poisson Process [29] and the semi-parametric estimates by the Prentice, Williams and Peterson (PWP) approach [38]. The focus of this evaluation was on the robustness of the PWP model, for the case where the true underlying model is a Nonhomogeneous Poisson Process with power law intensity function.

A secondary goal of the research was to define the knowledge base for an expert system to assist engineers in statistical analysis and modeling of failure data from repairable systems, including possible covariates. Another secondary objective was application of the method in this research to analyze actual field failure data from F-16 aircraft radar power supplies.

III. REVIEW OF LITERATURE

This section reviews the major published work on covariate models relevant to this project. Refer to the Appendix for more detailed discussion of models and tests and for definitions of notation and terminology.

Proportional Hazards Models

The method of proportional hazards modeling has been well documented in the statistical literature. References include Cox [11], Cox and Lewis [13], Kalbfleisch and Prentice [24], Lawless [29], Bendell [6,7], Baxter, et al. [5], Dale [15], Landers and Kolarik [27] and Wightman and Bendell [48].

Proportional hazards models are a class of regression models that examine the dependence of the lifetime distribution on concomitant, or explanatory, variables. The technique allows for the decomposition of the variation in the life lengths into independent factors. The significant factors are then identified, and the model is reconstituted for prediction purposes.

The proportional hazards model assumes that the explanatory variables have a multiplicative (rather than additive) effect on the hazard function, which in many cases is a more realistic assumption. The model also assumes that the hazard functions for different individuals are proportional to one another.

The advantages of the proportional hazards model over standard least-squares multiple regression techniques are that

the model handles censored data, tied values and failure times equal to zero. The model also allows for the pooling of data under heterogeneous conditions to improve estimates [7,24,26]. The basic proportional hazards model is:

$$h(t;\underset{\sim}{z}) = h_0(t)e^{\underset{\sim}{z}\underset{\sim}{\beta}} \qquad (1)$$

where

$h_0(t)$ is the baseline hazard function (i.e. the value of the hazard function if $\underset{\sim}{z}$ takes the value 0)

$\underset{\sim}{z}$ is the vector of explanatory variables

$\underset{\sim}{\beta}$ is the vector of model parameters which is to be estimated.

The exponential formulation is used because it assures a nonnegative hazard function. The explanatory variable can be a continuous measured variable, such as temperature or time, or a discrete indicator variable, such as presence or absence of a design change.

When an explanatory variable does not satisfy the proportional hazards assumption it may be used as a stratifying variable dividing the individuals under study into a number of strata. Within each stratum, the other covariates form proportional hazards. Model (1) is then modified to:

$$h_s(t;z) = h_{0s}(t)e^{\underset{\sim}{z}\underset{\sim}{\beta}s} \qquad (2)$$

where s denotes the sth stratum.

The proportional hazards modeling approach involves examining the data to provide estimates of the parameters β_1, β_2,... and investigating whether or not the parameters are significantly different from zero [1,25,26,27,31,32].

When the baseline hazard function $h_o(t)$ takes on a parametric form, such as a Weibull hazard, then the proportional hazards model becomes a parametric model. If $h_o(t)$ is left arbitrary the proportional hazards model becomes a semi-parametric model.

Proportional Intensity Models.

The semi-parametric statistical approach to reliability data analysis was developed by D.R. Cox [11]. Prentice, Williams and Peterson [38] extended the Cox model to the case of a stochastic process where individuals experience repeated failure (e.g., repairable systems). They considered the following conditional intensity function:

$$\rho\{t|N(t), \underset{\sim}{z}(t)\}=\lim_{\Delta t\to 0} \Pr\{t\le T_{n(t)+1} < t + \Delta t|N(t), \underset{\sim}{z}(t)\}/\Delta t \quad (3)$$

where

$N(t)$ is the complete failure history prior to time t

$\underset{\sim}{z}(t)$ is the complete history of covariates up to and including time t.

The two special cases of model (3) considered by Prentice, Williams and Peterson were:

$$\rho\{t|N(t),\underset{\sim}{z}(t)\} = \rho_{o_s}(t)\ \exp(\underset{\sim}{z}(t),\ \underset{\sim}{\beta}_s) \quad (4)$$

and

$$\rho\{t|N(t),\underset{\sim}{z}(t)\} = \rho_{o_s}(t-t_{n(t)})\exp(\underset{\sim}{z}(t),\beta_s) \quad (5)$$

where

ρ_{o_s} is the arbitrary baseline intensity function of the sth stratum

$\underset{\sim}{z}(t)$ is a vector of explanatory variables at time t

$\underset{\sim}{\beta}_s$ is a vector of parameters estimated for stratum s.

Individuals move from stratum to stratum as N(t) and $\underset{\sim}{z}(t)$ change. Model (4) uses a time scale measured from the beginning of the study, and model (5) uses a time scale measured from the immediately preceding failure (backward recurrence time). The stratification variable permits the study of times to first failure, times between first and second failure, and so on [38,48].

The merits of the Prentice, Williams and Peterson model are that no a priori distribution need be assumed for ρ_{0_s} (a semi-parametric approach) and each stratum may have a different baseline intensity function (a generalization from the standard assumption of a renewal process).

Lawless [28] has proposed a parametric proportional intensity model for the case of a nonhomogeneous Poisson process (NHPP) with power law intensity:

$$\rho(t) = \nu\delta t^{\delta-1} \tag{6}$$

where

t is the cumulative operating time

$\rho(t)$ is the intensity function or time-dependent rate of occurrence of failures (ROCOF)

ν is the scale parameter

δ is the shape parameter.

Covariates are incorporated through the scale parameter

$$\nu = e^{z_1\beta_1+z_2\beta_2+\cdots} \tag{7}$$

and the intensity function becomes

$$\rho(t) = \delta t^{\delta-1} e^{\underset{\sim}{z}\underset{\sim}{\beta}}. \tag{8}$$

Plots of the intensity function $\rho(t)$ versus time or of instantaneous mean time between failures IMTBF $[\rho(t)]^{-1}$ versus time are scaled apart based on the values $\beta_1, \beta_2 \ldots$ and have the same shape determined by the pooled estimate of δ. Plots of $\ln[\rho]$ versus $\ln(t)$ or $\ln[\text{IMTBF}]$ versus $\ln(t)$ are linear and parallel [18,28].

IV. APPROACH

The PWP model was tested on an actual as well as simulated data sets. The actual data set was obtained from the U. S. Air Force and is from F-16 radar power supply failures. The simulated data set was generated using a Monte Carlo method [9,30]. SAS procedure RANUNI [40] and Proschan's aircraft data set [39] were used to generate simulated failure times from a Nonhomogeneous Poisson Process with power law intensity.

Initially, the data set was investigated for significant explanatory variables using SAS procedure LIFETEST [40]. This procedure provides formal statistical tests useful in determining whether strata, defined by values of explanatory variables, are significantly different. The procedure also provides tests of association for other explanatory variables. LIFETEST was used to plot ln(t) versus ln(-ln [R(t)]). This plot allowed visual inspection of the test data set for the proportional hazards assumption. A Nonhomogeneous Poisson Process (NHPP), with power law intensity function was fit to the data using Crow's procedure, for the case of failure truncation. Crow's modified Cramer-von-Mises goodness-of-fit test was also performed [14].

When significant explanatory variable(s) were identified by LIFETEST, then the SAS procedure PHGLM was used to fit a PWP model to the test data set [41]. The Lawless method provided a parametric fit of the power law proportional intensity model [28].

Literature review, on tests for the proportional hazards assumption, indicated the test statistic reported by procedure PHGLM (Z:PH) to be the best formal statistical test. The test is based on the linear correlation between the residuals and the rank order of the failure time. Simulations have shown this test to have reasonable power when the hazard ratio is monotonic in time [41].

Schoenfeld's Chi-squared statistic is a potential omnibus test for the proportional hazards model, but requires that the time axis and the range of the covariates be divided into intervals [42]. This procedure involves subjective judgement and is not appropriate for use in the research or inclusion into a rule-based expert system.

Early versions of the statistical package BMDP provided a graphical method for verifying the proportional hazards assumption. This was the residual plot available in the procedure P2L [8]. If the proportional hazards assumption holds, then hazard plots for the residuals are assumed to be exponential order statistics. However, this statistic has been severely criticized in the literature [44] and subsequently removed from BMDP. Other tests have been proposed but not thoroughly evaluated [19,34,35,36,42,43].

No satisfactory statistical goodness-of-fit procedure was found in the literature for the PWP model. Consequently, it was not possible to perform size and power studies. The PWP model was evaluated by comparing it to the proportional

intensity NHPP, with power law intensity function (Lawless [28]). The parametric and semi-parametric models were compared graphically and measured on the basis of nearness criteria, relative to the true (theoretical) model. The mean absolute deviation and mean squared error measured dispersion, while the mean of the signed error measured bias.

SAS procedure PHGLM provided estimates for the parameter $\underset{\sim}{\beta}$ using Cox's partial likelihood method [12,23]. The parameters for the NHPP model were estimated by maximizing the log-likelihood reported by Lawless [28]. A SAS program was developed to generate simulated data from a NHPP and to estimate the parameters by the method of maximum likelihood, using a Newton-Raphson algorithm.

The researchers developed additional software to obtain estimates of the mean time to failure (MTTF) from the PWP estimates of the reliability function in each stratum defined by the failure count. The MTTF is defined by

$$MTTF = \int_0^\infty R(t)dt. \tag{9}$$

and was estimated by numerically integrating the Kaplan-Meier reliability function estimated in each stratum:

$$\hat{E}(t-t_n)_{s,PWP} = \hat{MTTF}_s = \int_0^\infty \hat{R}_s(x;\underset{\sim}{z};\underset{\sim}{\beta})dx \tag{10}$$

where

t_n is the time of the nth failure (transition into stratum s = n + 1

t is the time of the n + 1st failure (transition into stratum s = n + 2

x is a random variable denoting time to next (n + 1st) failure in stratum s.

$\hat{E}(t-t_n)_{s,PWP}$ is the expected time to next (n + 1st) failure while in stratum s = n + 1, estimated by the semi-parametric PWP method

$\hat{R}_s(x;\underset{\sim}{z};\underset{\sim}{\beta})$ is the reliability function for time to next (n+1st) failure while in stratum s = n + 1, estimated by the semi-parametric PWP method.

The researchers compared both the PWP model and the parametric estimated NHPP to the true (theoretical) NHPP from which the simulated data was produced. The comparison was based on the instantaneous mean time to next failure and was evaluated at the expected times of failure (transition to the next higher stratum defined by failure counts). For the underlying theoretical NHPP, the expected number of failures in the time interval (0,t) is

$$E(n) = \nu t^{\delta}. \tag{11}$$

The expected time of the nth failure $E(t_n)$ is then

$$E(t_n) = \left[\left(\frac{n}{\nu} \right)^{1/\delta} \right]. \tag{12}$$

At that expected cumulative operating time, the instantaneous time to next failure, estimated by the PWP model, is the MTTF for stratum n+1.

The true instantaneous time to next failure for the underlying NHPP is

$$E(t-t_n) = IMTBF(t_n) = \frac{1}{\rho(t_n)} = [\nu\delta t^{\delta-1}]^{-1}. \tag{13}$$

Similarly, the estimated IMTBF, by the Lawless method, is

$$\hat{E}(t\text{-}t_n)_{NHPP} = \widehat{IMTBF}(t) = \frac{1}{\hat{\rho}(t)} = [\hat{\nu}\hat{\delta}t^{\hat{\delta}-1}]^{-1}.$$

$$= \left[\hat{\delta}t^{\hat{\delta}-1}e^{z\hat{\beta}}\right]^{-1}. \qquad (14)$$

where

$\hat{E}(t\text{-}t_n)_{NHPP}$ is the expected time to next (n + 1st) failure, estimated by the parametric Lawless method

$\hat{\nu}, \hat{\delta}, \hat{\beta}$ are parameters of the NHPP estimated by the Lawless method.

The relative error of the estimates can be measured:

$$e_n = \frac{\hat{E}(t\text{-}t_n) - E(t\text{-}t_n)}{E(t\text{-}t_n)}. \qquad (15)$$

If there are m failures per item, then we can measure these errors for failures 1 through m-1.

Then the bias is

$$BIAS = \left(\sum_{n=1}^{m-1} e_n\right)/(m-1). \qquad (16)$$

The mean absolute deviation (MAD) is

$$MAD = \left(\sum_{n=1}^{m-1} |e_n|\right)/(m-1). \qquad (17)$$

The mean squared error (MSE) is

$$MSE = \left(\sum_{n=1}^{m-1} e_n^2\right)/(m-2). \qquad (18)$$

V. ANALYSIS AND RESULTS

This section reports the analysis and results performed on the simulated and actual data sets.

Analysis

Simulated Data Sets

In order to develop realistic data sets, the researcher fit power law NHPP models to Proschan's aircraft air conditioning data set [38]. Aircraft 7908 had twenty-three failures through 2201 hours of operating life. After the thirteenth failure, the system underwent a complete overhaul. Unfortunately, the reliability was substantially lower after the overhaul. We fit a power law proportional intensity NHPP model to the two data sets (before and after overhaul) to obtain the estimates of the shape parameter δ = 1.2324 and the scale parameters for the two strata (ν_0 = 0.0009 and ν_1 = 0.0093). The zero subscript denotes CLASS 0 (before overhaul) while the subscript one denotes CLASS 1 (after overhaul). The resulting model contained a shape parameter greater than one, which gives an increasing rate of occurrence of failures (IROCOF). To investigate the performance of estimators for the case of a decreasing ROCOF, we generated simulated data sets with δ = 1.2324 / 2 = 0.6162.

The analytical procedure was to:

(a) Generate simulated data set

(b) Estimate the parameters of an NHPP (Eq. (14))

(c) Estimate strata functions (Eq. (5))

(d) Estimate strata MTTFs (Eq. (10))

(e) Plot theoretical and estimated functions

(f) Calculate relative bias and errors (Eq. (16), (17), (18)).

The researchers developed a SAS program to generate the simulated data using equations for power law NHPP variates proposed by Blanks and Tordon [9]. The SAS program also provides maximum likelihood estimators of power law NHPP parameters using Lawless's likelihood function [28] and a Newton-Raphson optimization algorithm [29].

A separate SAS program uses a common random-number generator and seed value to reproduce the simulated data set. The program also uses SAS PROC PHGLM [41] to give the PWP estimates of regression parameters and reliability functions using the Cox likelihood [11]. Since PHGLM is in the SAS user's group library not supported by the SAS Institute, validation tests were performed on PHGLM against the widely accepted BMDP routine for Cox regression, P2L. The researchers tested PHGLM versus P2L using the Prentice, Williams and Peterson data set for bone marrow transplants[32,38]. Since PHGLM and P2L gave identical results, we standardized on the SAS PHGLM for commonality with our other programs.

The analyses employed PHGLM to block on a stratification variable (FAILURES) and regress on one discrete covariate (CLASS). The SAS program also produces predictions and plots for the resulting PWP model. Plots of reliability function, log

transformed reliability function and hazard function are produced. Finally, the program performs numerical integrations to provide estimates of the mean time to failures (expected time of the next failure) in each stratum defined by the count (stratification variable FAILURES).

Table 1 summarizes the characteristics of the data sets analyzed.

TABLE 1

DATA SETS SIMULATED AND ANALYZED

DATA SET	SHAPE δ	FAILURES N	----INDIVIDUALS---- CLASS 0	CLASS 1
1	1.2324	5	30	30
2	1.2324	5	60	60
3	0.6162	5	30	60
4	0.6162	5	60	60

The studies addressed cases with both increasing and decreasing ROCOF. In all simulations, there were five failures per individual item. In two of the simulations, there were 60 individuals equally divided between CLASS 0 and CLASS 1. In the other two simulations, there were 60 items in CLASS 0 and 60 items in CLASS 1.

Actual Data Set

The F-16 System Program Office (Aeronautical Systems Division, Wright-Patterson Air Force Base) provided actual

failure histories on a random sample of 13 F-16 radar power supplies. We analyzed the data using the methods of the Appendix, to identify a useful covariate. There was one candidate covariate definable in the data set: primary air base of operations. Such a covariate could yield insight to reliability as a function of maintenance and operational environment (which can vary among different organizations). However, this covariate did not form proportional intensity functions and the sample size was inadequate for analysis by the parametric or semi-parametric methods used in this project.

Results

PWP Estimates

Figures 1 through 3 illustrate the results of PWP analysis on a simulated data set. For the model with increasing ROCOF (δ = 1.2324) and sample size of 120 items, these Figures give the results for stratum 1 (the risk set of individuals awaiting and then experiencing first failure). A "0" denotes the stratum for CLASS 0 and a "1" denotes the stratum for CLASS 1. Figure 1 contains the product-limit (Kaplan-Meier) estimates of the reliability functions for the two classes. These functions were numerically integrated to obtain the expected time to first failure. Figure 2 contains plots of the increasing hazard functions for strata 0 and 1. Figure 3 contains the plots of ln(-ln[R(t)]) versus ln(t). The plots appear parallel, indicating that the CLASS covariate forms proportional hazards in stratum one. The plots also appear

Figure 1. Stratum 1 reliability estimates

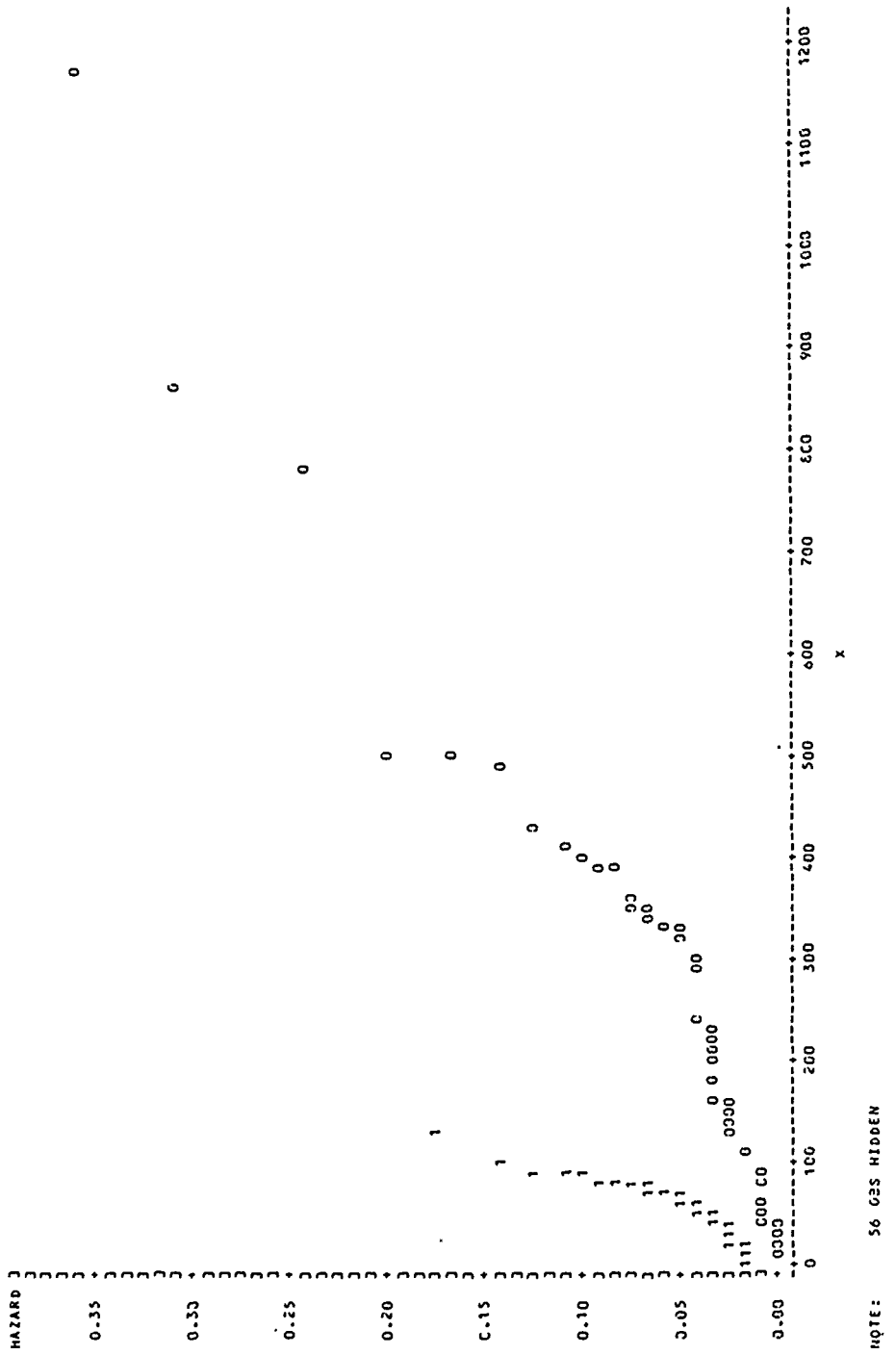


Figure 2. Stratum 1 hazard estimates

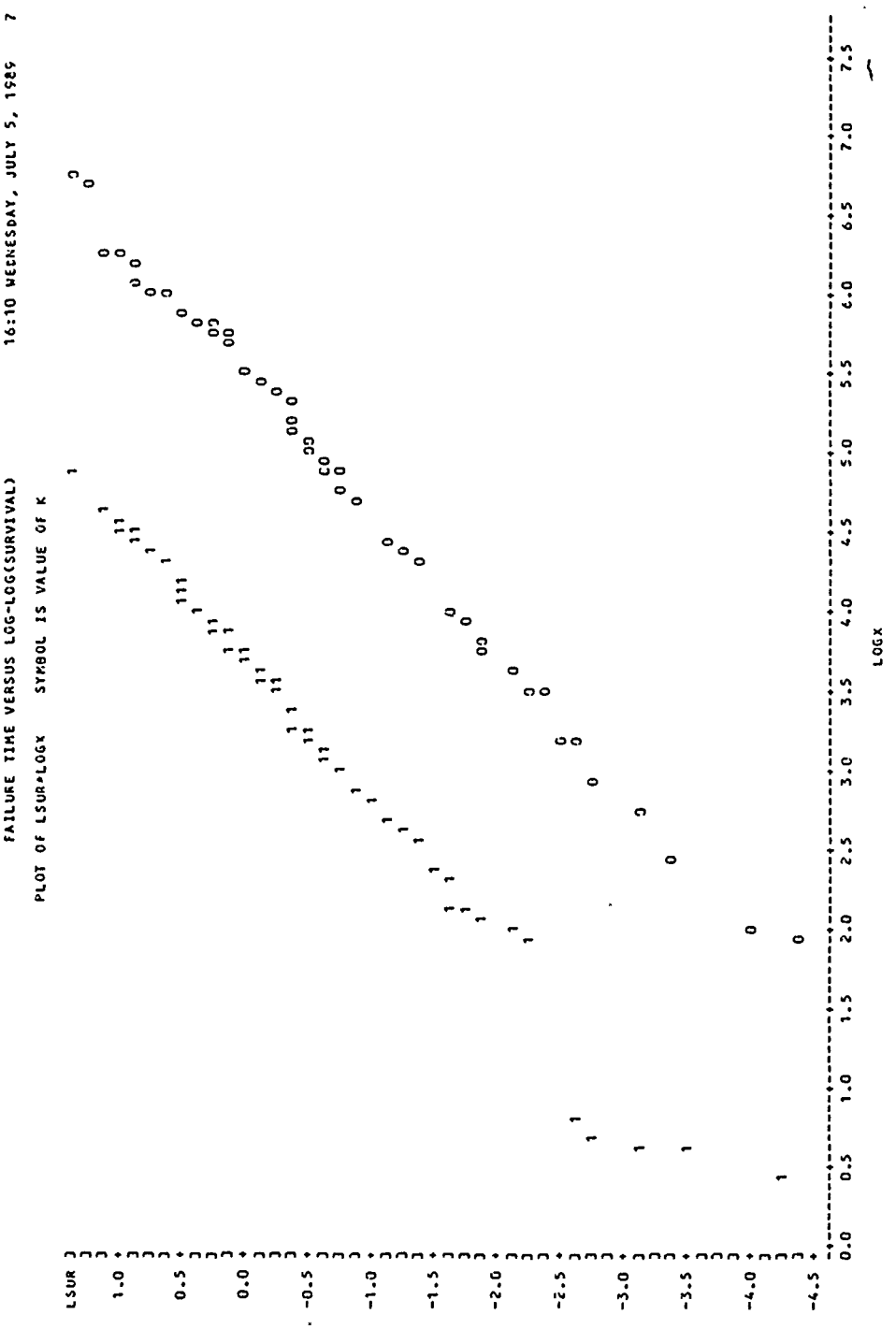


Figure 3. Stratum 1 plots of ln([-ln[R(t)]) versus ln(t)

linear, suggesting the Weibull as a possible distribution for a time-to-first-failure random variable. A NHPP with power law intensity function does have a Weibull distribution for time to first failure. Figure 4 contains the plots of ln(-ln[R(t)]) versus ln(t) for stratum 4, which also appear parallel and roughly linear with a shape tending slightly concave upward.

Robustness

Figures 5 through 8 summarize the results of the robustness studies. Figure 5 relates to data set 1 (Table 1). This data set was generated from an NHPP with power law proportional intensity and increasing ROCOF. The sample size was 60 items, divided equally between CLASS 0 AND CLASS 1. The figure contains plots of log-transformed instantaneous MTBF (IMTBF, or expected time to next failure) versus ln(t).

The plots of IMTBF are linear, illustrating that the underlying failure process is a power law NHPP. Plots of true IMTBF for the two classes are also parallel since the covariate CLASS forms proportional intensities. The estimated IMTBFS were obtained by the parametric Lawless method and are also linear and parallel, but with slope and locations differing slightly from the true models. For the sample of 60 items, the estimate of the shape parameter was 1.45504 compared to the theoretical value of 1.2324. Consequently, the slope of the estimated plots is greater than the slope of the true plots.

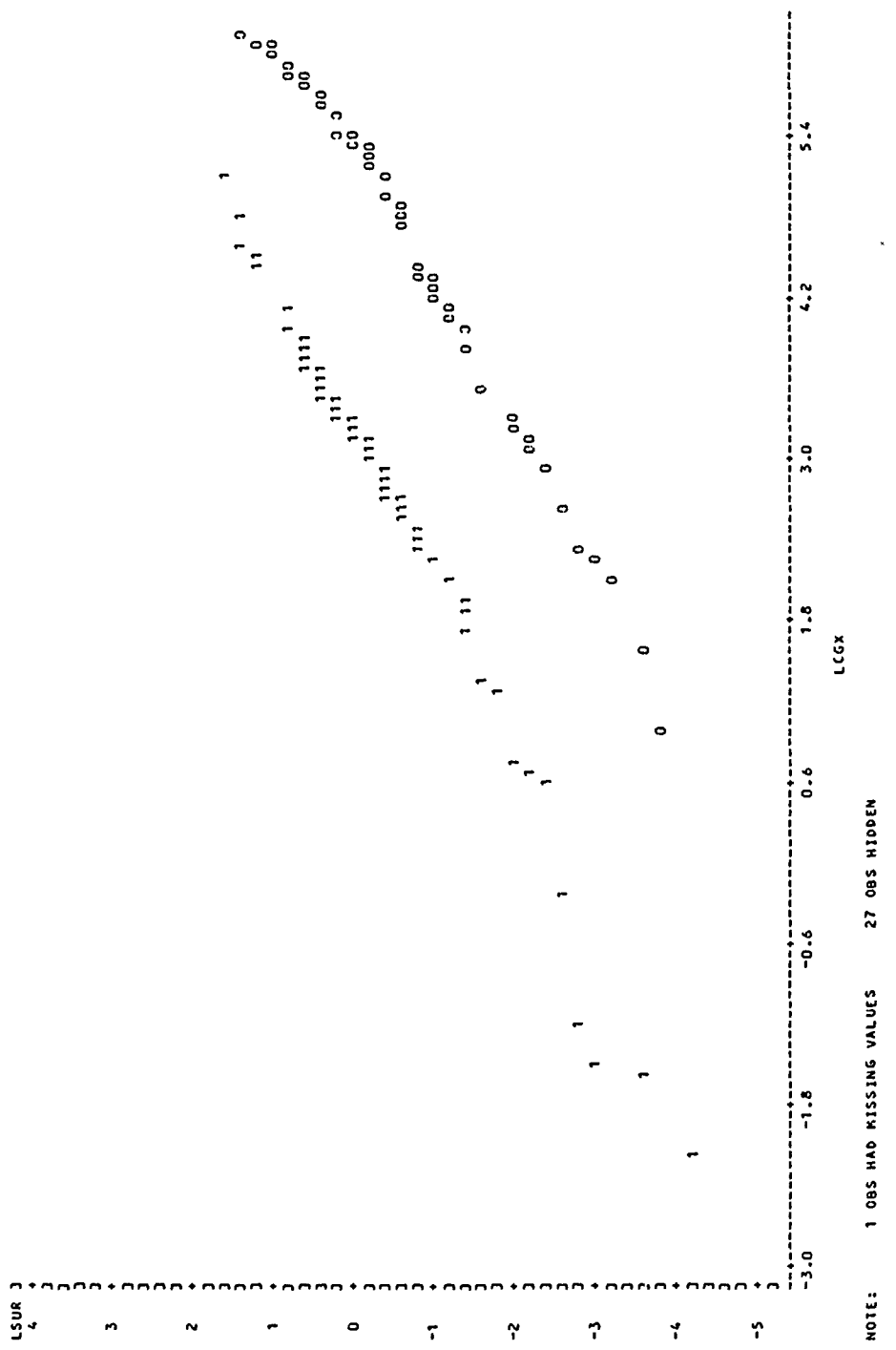


Figure 4. Stratum 4 plots of ln{-ln[R(t)]} versus ln(t).

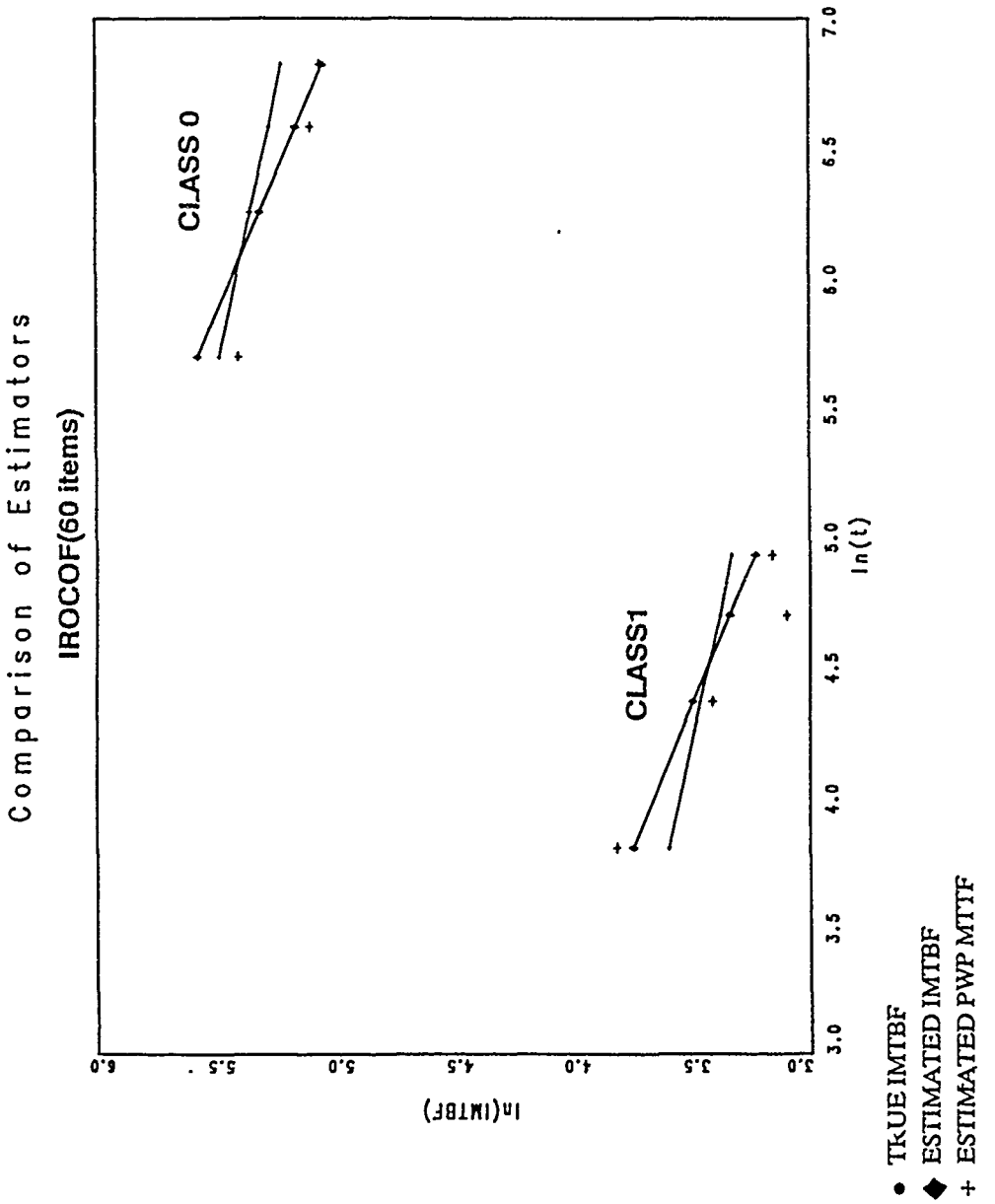


Figure 5. Comparison of estimators, IROCOF (60 ITEMS).

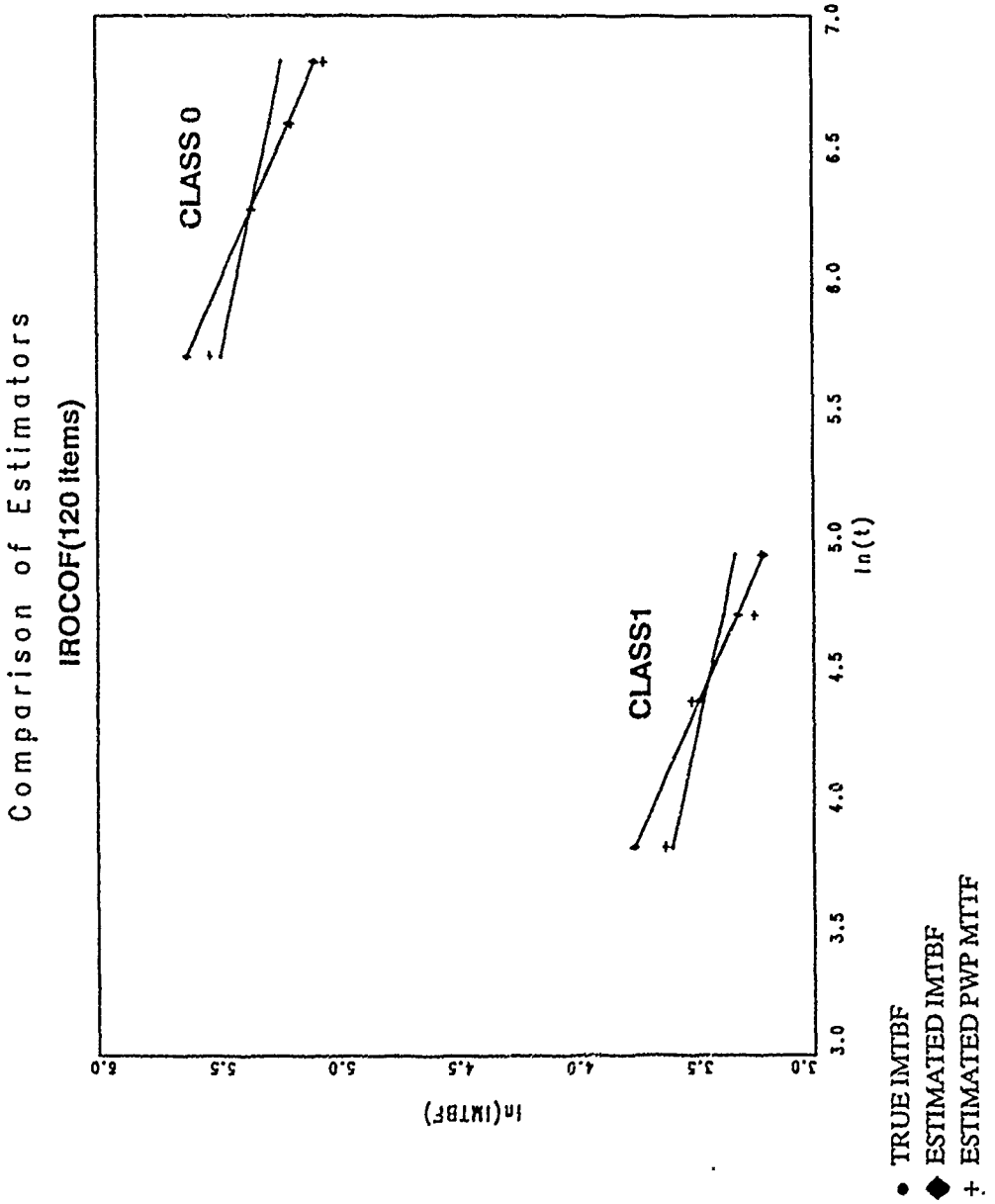


Figure 6. Comparison of estimators, IROCOF (120 items).

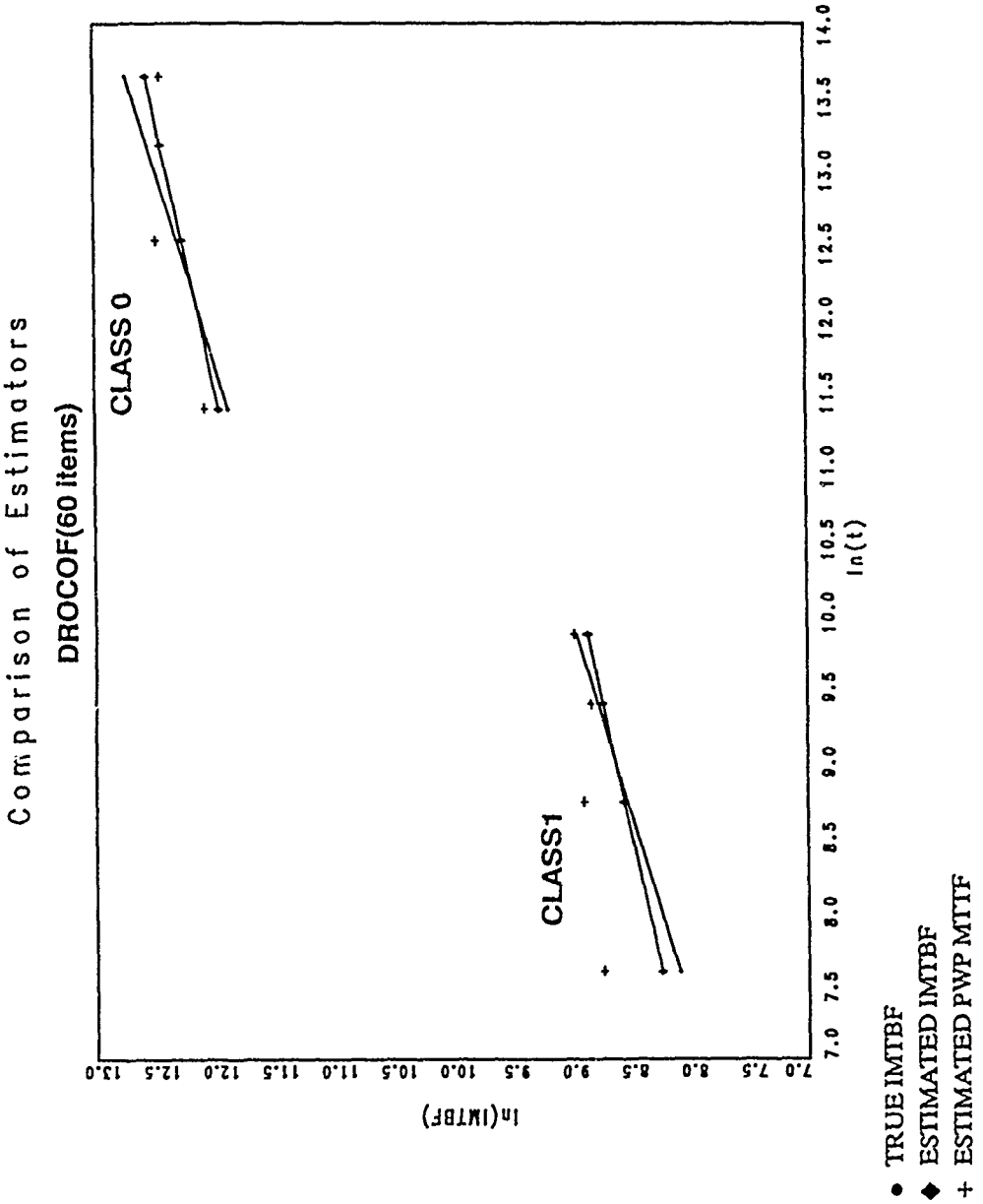


Figure 7. Comparison of estimators, DROCOF (60 items).

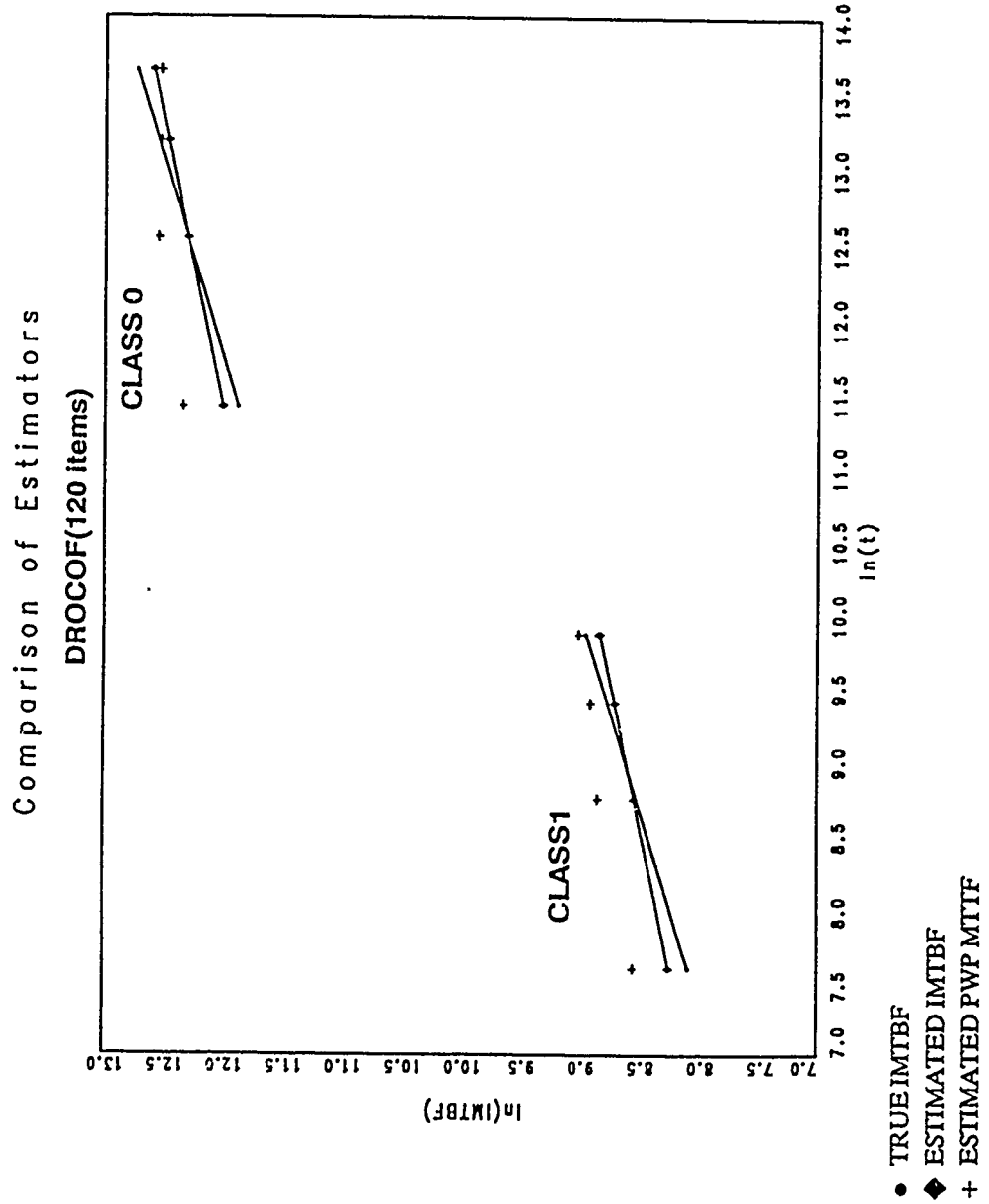


Figure 8. Comparison of estimators, DROCOF (120 items).

Four PWP estimates are plotted for each value of CLASS. Each point is an estimated mean time to failure (MTTF). The coordinates of a point are the expected time of nth failure versus the expected time to next failure (n + 1). For example, ir CLASS 1, based on the true underlying NHPP the expected time of the second failure is 78 hours. The PWP estimate of the mean time to third failure is 30 hours. Therefore, the plotting position of the point is (ln(78),ln(30)) = (4.36,3.40).

For CLASS 0, both the parametric and semi-parametric estimates appear close to each other and to the true values. There appears to be a tendency toward under-estimation, (negative relative bias). Table 2 supports this graphical evidence. The relative mean squared errors are small (1 to 2%) and the relative biases are negative. The semi-parametric PWP estimates exhibit more bias than do the parametric Lawless estimates. For CLASS 1 compared to CLASS 0, the PWP estimates exhibit less negative relative bias, but slightly higher relative error.

Figure 6 illustrates the results for the same true NHPP, but for larger sample size (120 versus 60). The results in Figure 6 are comparable to those shown in Figure 5, with slightly less relative error, as expected with increasing sample size (see Table 2).

TABLE 2

COMPARISON OF ESTIMATORS

RELATIVE BIAS AND ERRORS (%)

			---CLASS 0---			---CLASS 1---		
δ	I	METHOD	BIAS	MAD	MSE	BIAS	MAD	MSE
1.2324	60	LAWLESS	-5	10	1	2	8	1
		PWP	-10	10	2	-5	17	5
	120	LAWLESS	-1	9	1	1	9	2
		PWP	-5	8	1	-4	8	1
0.6162	60	LAWLESS	-5	10	1	2	8	1
		PWP	2	19	5	36	36	35
	120	LAWLESS	-1	9	1	1	9	2
		PWP	18	27	16	30	30	18

For the case of an IROCOF, the parametric Lawless estimators exhibit moderate under-estimation (negative bias) in CLASS 0 (higher reliabilities) and over-estimation (positive bias) in CLASS 1 (lower reliabilities). The semi-parametric PWP approach predominantly under-estimates the IMTBF of the process having IROCOF. The parametric estimates consistently outperform the PWP estimates in closeness (MAD and MSE) for IROCOF. However, Table 2 shows that the MSEs tend to be less than 5% and the MADs less than 10%, with one exception (17% for

CLASS 1 and sample size of 60).

Figures 7 and 8 relate to the case of decreasing rate of occurrence of failures (DROCOF). The parametric Lawless estimators perform identically for the DROCOF and the IROCOF. However, the PWP estimates consistently overestimate the expected time to next failure (IMTBF), particularly for early failures and for CLASS 1 (lower reliabilities). The PWP estimates tend to improve at higher reliabilities and with increase in the number of failures per unit. Table 2 summarizes the results for DROCOFs, showing that the PWP performs very well for a sample of 60 at CLASS 0. For the other cases investigated, the PWP biases are always positive, ranging from 18 to 36%.

The poor performance of the PWP for the DROCOF data suggests the need for further study. The results could be peculiar to the chosen values of the simulation parameters. Since the estimators improve with failure count, a start-up transient could be another explanation of the estimation problem. Table 3 provides additional insight. The results in Table 2 were obtained from a blocked analysis, where the covariate regression coefficients are assumed to be the same in all strata, permitting estimation of a single β. Table 3 gives the β estimates (point and 90% confidence interval) within strata, showing that for δ = 1.2324, the β's are similar in all strata, and indicating that blocking was valid. For δ = 0.6162 the β's in strata 1 and 5 differ substantially from each other

and from the other 3 strata. There also appears to be a trend in the β estimates across successive strata.

TABLE 3
STRATUM ESTIMATES OF BETA

Stratum	$\hat{R}$	STD. ERROR	90% Conf. INTERNAL	$\hat{\beta}$	STD. ERROR	90% Conf. INTERNAL
1	2.06	0.277	(1.60,2.52)	2.06	0.277	(1.60,2.52)
2	2.31	0.274	(1.86,2.76)	2.96	0.343	(2.40,3.52)
3	1.91	0.247	(1.50,2.32)	2.87	0.337	(2.32,3.42)
4	1.70	0.238	(1.31,2.09)	3.04	0.351	(2.46,3.62)
5	2.29	0.284	(1.82,3.91)	3.91	0.477	(3.12,4.69)

VI. RECOMMENDATIONS

Future research should expand the range of cases investigated under the alternate hypotheses of a parametric power law NHPP. The researchers plan to expand the study reported here to include the simulated data sets in Table 4. This research will address cases where the number of failures per unit is larger (10 versus 5) and cases of small sample size (5 or 10 individual units per class versus 30 or 60). The results of the expanded study will be published in a Master's thesis, and should be available prior to January 1990.

The researchers also recommend expansion of this robustness study to consider other simulated data sets with varied parameter values. Sample sizes other than those in Table 3 should be tested. This study has been limited to the case of failure truncation, where the last observed event for each unit

TABLE 4
SUMMARY OF SIMULATED DATA SETS
USING FAILURE TRUNCATION

IROCOF (δ = 1.2324)				DROCOF (δ = 0.6162)			
N = 5		N = 10		N = 5		N = 10	
I_0	I_1	I_0	I_1	I_0	I_1	I_0	I_1
60	60	30	30	60	60	30	30
30	30	5	5	30	30	5	5
10	10			10	10		

Note:

δ is the shape parameter of the NHPP, where 1.2324 represents an increasing rate of occurrence of failures (IROCOF) and 0.6162 represents a decreasing rate of occurrence of failures (DROCOF).

N is the number of failures per individual

I_0 is the number of individuals in CLASS 0

I_1 is the number of individuals in CLASS 1.

is a failure. There is a need to address the case of time truncation, where the observed lifetime is censored (truncated at a time other than a failure event). This research has considered the case where the alternate hypothesis is a power law NHPP. An extension of this work would be to consider the NHPP with log-linear intensity function (see Appendix, section 3.3.1) [13].

The Air Force should provide access to substantial actual data sets on repairable items, and support for applied research

using the estimators evaluated in this project.

The knowledge base for nonrepairable items is sufficiently defined to permit full-scale development of an expert system to aid engineers in analysis of nonrepairable failure data. This project has defined the knowledge base sufficiently to permit development of a prototype expert system to aid in analysis of repairable item failure data.

VII. REFERENCES

[1] Andersen P.K. and Gill R.D. (1982), "Cox's Regression Model For Counting Processes: A Lange Sample Study," The Annals of Statistics, 10,4, pp. 1100-1120.

[2] Ascher, H. (1983), "Analysis of Repairable Systems Reliability," NATO ASI SEMES, vol. F3, pp. 119-133.

[3] Ascher, H. (1983), "Regression Analysis of Repairable Systems Reliability," Electronic Systems Effectiveness and Life Cycle Costing (Skwirgznski, J.K., Ed.), pp. 118-133, Springer-Verlag.

[4] Ascher, Harold and Feingold, Harry, (1984). Repairable Systems Reliability, New York, Marcel Dekker, Inc.

[5] Baxter, J.J., Bendell, A. and Ryan, S.G. (1988), "Proportional Hazards Modeling of Transmission Equipment Failures," Reliability Engineering and System Safety, 21, pp. 129-144.

[6] Bendell, A. (1985), "Proportional Hazards Modeling in Reliability Assessment,", Reliability Engineering, 11, pp. 175-183.

[7] Bendell, A., Walley, M. Wightman, D.W., and Wood, L.M. (1986), Quality and Reliability Engineering International, 2, pp. 45-52.

[8] BMDP Statistical Software Manual (1985), Berkley, University of California Press.

[9] Blanks, Henry S. and Tordon, Michael J. (1979) "Laplace and Mann Aging Trend Test Effectiveness," Proc. Annual Reliability and Maintainability Symposium, Philadelphia, Pennsylvania, pp. 165-170.

[10] Booker, J., Campbell, K., Goldman, A.G. Johnson, M.E. and Bryson, M.C. (1981), "Application of Cox's Proportional Hazards Model to Light Water Reactor," Los Alamos Scientific Laboratory Report, LA-8834-SR.

[11] Cox, D.R. (1972), "Regression Models and Life Tables (with discussion)," Journal of the Royal Statistical Society, B34, pp. 187-220.

[12] Cox, D.R. (1975), "Partial Likelihood," Biometrika, 62, p. 269.

[13] Cox, D.R. and Lewis, P.A.W., (1968) The Statistical Analysis of Series of Events, London, Methuen and Co.

[14] Crow, L. H. (1974), "Reliability Analysis for Complex Repairable Systems," Reliability and Biometry, pp. 379-410.

[15] Dale, C. (1985), "Application of the Proportional Hazards Model in the Reliability Field," Reliability Engineering, 10, pp. 1-14.

[16] Davis, H.T., Campbell, K. and Schrader, R. M. (1980), "Improving the Analysis of LWR Component Failure Data," Los Alamos Scientific Laboratory Report, LA-UR 80-92.

[17] Engelhardt M. and Bain L.J. (1980), "Inferences on the Parameters and Current System Reliability for a Time Truncated Weibull Process," Technometrics, 22, 3, pp. 421-426.

[18] Engelhardt, Max and Bain, Lee J. (1987), "Statistical Analysis of a Compound Power-Law Model for Repairable Systems," IEEE Transactions on Reliability, R-36, 4, pp. 391-396.

[19] Gill, R. and Schumacher, J. (1987), "A Simple Test of the Proportional Hazards Assumption," Biometrika, 74, 2, pp. 289-300.

[20] Harmon, P., Maus R. and Morrissey W.(1988), Expert Systems Tools and Applications, John Wiley, New York.

[21] Jardine, A.K.S. (1983), "Component and System Replacement Decisions," Electronic Systems Effectiveness and Life Cycle Costing (Skwirzynski, J.K., Ed), pp. 647-654.

[22] Jardine, A.K.S. and Anderson, M. (1984), "Use of Concomitant Variables for Reliability Estimation and Setting Component Replacement Policies," 8th Advances in Reliability Technology Symposium, Bradford, B3, pp. 211-6.

[23] Kalbfleisch, J.D. (1978), "Nonparametric Bayesian Analysis of Survival Data," Journal of the Royal Statistical Society, B40, p. 214.

[24] Kalbfleisch, J.D. and Prentice, R.L. (1980), The Statistical Analysis of Failure Time Data, New York, Wiley.

[25] Kay R. (1977), "Proportional Hazard Regression Models and the Analysis of Censored Survival Data," Applied Statistics, 26, 3, pp. 227-237.

[26] Landers, T.L. and Kolarik, W.J. (1987), Life Cycle Life Testing: Using the Proportional Hazards Approach, DNA/SDIO Technical Report TTU-TR-3, Prepared for Headquarters Defense Nuclear Agency, Strategic Defense Initiative Organization.

[27] Landers, T.L. and Kolarik, W.J. (1987), "Proportional Hazards Analysis of Field Warranty Data," Reliability Engineering, 18, pp. 131-139.

[28] Lawless, J.F. (1987), "Regression Methods for Poisson Process Data," Journal of the American Statistical Association, 82, pp. 808-815.

[29] Lawless, J.F. (1982), Statistical Models and Methods for Lifetime Data, John Wiley, New York.

[30] Leemis L. and Schmeiser B. (1985), "Random Variate Generation for Monte Carlo Experiments," IEEE Transactions on Reliability, R-34, 1, pp. 81-85.

[31] Lewis, P.A.W. and Robinson D.W. (1974), "Testing For a Monotone Trend in a Modulated Renewal Process," Reliability and Biometry, pp. 163-182.

[32] Marek, Patrick (1989). Personal Correspondences regarding analysis and programs for the PWP model.

[33] Moller, S.K. (1976), "The Rasch-Weibull Process," Scandinavian Journal of Statistics, 3, pp. 107-115.

[34] Moreau, T., O'Quigley, J. and Lellouch, J. (1986), "On D. Schoenfeld's Approach for Testing the Proportional Hazards Assumption," Biometrica, 73, pp. 513-515.

[35] Nagel, P.M. and Skrivan, J.A. (1981), "Software Reliability: Repetitive Run Experimentation and Modeling," Boeing Computer Services Co. Report, BCS-40366, NASA Report CR-165836.

[36] Nagelkerke, N.J.D., Oosting J. and Hart, A.A.M. (1984), "A Simple Test of Goodness of Fit of Cox's Proportional Hazards Model,"Biometrics, 40, pp. 483-486.

[37] O'Quigley, J. and Moreau, T. (1986), "Cox's Regression Model: Computing a Goodness of Fit Statistic," Computer Methods and Programs in Biomedicine, 22, pp. 253-256.

[38] Prentice, R.L., Williams, B.J. and Peterson, A.V. (1981), "On the Regression Analysis of Multivariate Failure Time Data," Biometrika, 68, pp. 373-379.

[39] Proschan, F. (1963), "Theoretical Explanation of Observed Decreasing Failure Rate," Technometrics, 5, 3, pp. 392-396.

[40] SAS User's Guide: Statistics Version 5 Edition (1985), Cary, North Carolina, SAS Institute, Incorporated.

[41] SUGI Supplemental Library User's Guide, Version 5 Edition (1985), Cary, North Carolina, SAS Institute, Incorporated.

[42] Schoenfeld, D. (1980), "Chi-Squared Goodness-of-Fit Tests for the Proportional Hazards Regression Model,"Biometrica, 67, 1, pp. 143-153.

[43] Schoenfeld, D. (1982), "Partial Residuals for the Proportional Hazards Regression Model," Biometrika, 69, 1, pp. 239-241.

[44] Storer B.E. and Crowley J.(1983), "Comment," Journal of the American Statistical Association, 78, 382, pp. 277-281.

[45] Taha, Hamdy A., (1987), Operations Research: An Introduction, New York, Macmillan.

[46] Thompson, W.A., (1981), "On the Foundations of Reliability," Technometrics, 23, 1, pp. 1-13.

[47] Wightman, D.W. and Bendell, A. (1986), "The Practical Application of Proportional Hazards Modeling," Reliability Engineering, 15, pp. 29-53.

[48] Williams, Barbara J. (1980), Proportional Intensity Models for Multiple Event Times, Ph.D. Dissertation, University of Washington.

Appendices can be obtained from
Universal Energy Systems, Inc.

A MONTE CARLO COMPARISON

OF

VALIDITY GENERALIZATION PROCEDURES

By

Jorge L. Mendoza

and

Robert N. Reinhardt

Psychology Department
Texas A & M University
College Station, Texas 77843

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

Contract No. F49620-85-C-0013/SB5851-0360
Purchase Order No. S-760-6MG-136
Principal Investigator: Jorge L. Mendoza

PREFACE

The authors wish to express their appreciation to Dr. Malcolm Ree and Dr. Lonnie D. Valentine, Jr. of the Manpower Personnel Division of the Air Force Human Resources Laboratory. Their disscussions, suggestions and support were essential to this effort. The insights provided by Dr. Ree were especially helpful.

Summary

The Armed Services use tests for the selection and placement of military and civilian personnel. Numerous studies are conducted each year by the Armed Services to confirm assumptions about the validity of these tests. The cost is large and evidence that greater reliance could be made on past studies, or that more accurate procedures were available, would save both time and money. Demonstrating that a test can predict in a variety of locations without having to conduct a validity study in each location can yield significant savings. The Air Force would not always have to conduct a validity study before using a test if the validity of this had been demonstrated in a similar situation. A number of procedures have been proposed in the literature to determine when the validity of a test (or construct) generalizes across situations. These procedures are known as validity generalization procedures.

Investigators conducting validity generalization research have usually assumed a set of artifact distributions when estimating the mean and variance of the true validity distribution. The thinking has been that not enough information was available from the validation studies to allow for the estimation of the parameters of the artifact distributions. We have demonstrated, however, that under certain assumptions, validity generalization procedures can accurately estimate the mean and variance of true validity distributions using estimates of artifact parameters from the validation studies, without the assumed artifact distributions. Quality and not quantity was found to be the important factor in a validity generalization study. The results suggest that a

validity generalization study can be carried out effectively with few studies which provide accurate estimates of the artifact parameters even when only 50% of the studies contain artifact information. The paper outlines the statistical foundation for basing validity generalization studies on situational data, and compares six validity generalization studies using a Monte Carlo simulations. The procedures gave the most biased results when the selection ratio was small. Overall, the TSA1 procedure of Raju and Burke (1983) was the most accurate.

I. INTRODUCTION

Researchers have known for years that the validity of a test changes from one situation to another even when the same criterion is used across situations. This led Ghiselli (1966) to review hundreds of validity studies in an effort to categorize and identify situations in which tests are valid. He suggested that artifacts such as unreliability of the predictor and criterion, range restriction, and sampling error account for most of the variability of validity coefficients across situations. However, not until the seminal work of Schmidt and Hunter (1977) did researchers have a way to quantitatively estimate the magnitude of these effects.

Schmidt and Hunter's work stimulated other researchers and a number of validity generalization procedures appeared in the literature shortly. Callender and Osburn (1980) using a multiplicative model suggested a procedure somewhat different from Schmidt and Hunter's (1977). Building on this multiplicative structural model, Raju and Burke (1983) presented two validity generalization procedures requiring less assumptions than previous ones. Although the procedures developed by all these investigators (including those of Schmidt, Gast-Rosenberg & Hunter, 1980) are similar and generally yield equivalent results, they can yield different results under certain circumstances. A number of computer studies have investigated these similarities/differences and we will discuss them later in the paper (e.g., Callender, Osburn, Greener & Ashworth, 1982; Raju and Burke, 1983). Lately, a procedure which is different, both conceptually and computationally, has been suggested by Hedges (Note

1).

The primary goal of a validity generalization procedure is to estimate the mean and variance of "true" validity across situations, "true" validity being the validity of a test (or construct) when there is perfect reliability and no range restriction. Basically, after making a number of assumptions, the procedures provide a way to estimate the mean and variance of each artifact variable. Using these means and variances one adjusts the variance of the observed validity coefficients to obtain an estimate of the "true" validity variance. Since artifact information is frequently not available (Jones, 1950), Schmidt and Hunter (1977) proposed using hypothetical (assumed) distributions of artifacts. Rather than using statistics from each situation to estimate the artifact parameters, the parameters are obtained from the hypothetical distributions (Schmidt & Hunter, 1977; Schmidt, Hunter & Pearlman, 1982).

II. BACKGROUND

Callender and Osburn (1980) were the first to use a Monte Carlo simulation to evaluate the accuracy of validity generalization procedures. They created a distribution of true validities, attenuated this distribution with the assumed distributions of Schmidt and Hunter (1977), and applied the validity generalization procedures. Callender and Osburn (1980) tested the accuracy of procedures under three sets of assumptions (or cases). In Case 1, true validity was held constant while reliability and range restriction were varied; Case 2 was the opposite of Case 1; and in Case 3, all three variables were allowed to vary.

In an effort to lessen the independence restrictions of previous procedures Raju and Burke (1983) introduced two procedures (TSA1 and TSA2). The procedures require only that the covariances among the artifacts and "true" validity be zero instead of independent. Raju and Burke (1983) compared their procedures to the interactive and non-interactive procedures given in Schmidt et al. (1980) and to Callender and Obsurn's (1980) independent multiplicative procedure. Their study, a Monte Carlo simulation similar to that of Callender and Osburn (1980), compared the procedures in the absence of sampling error. The results from the five procedures for validity generalization were similar, with TSA1 in general slightly better than the others. For Case 1 (constant true validity) the non-interactive procedures tended to underestimate the variance while the other procedures generally overestimated the variance of true validity. In Case 2, the procedures generally overestimated the variance. Under Case 3 (everything varies), the procedures also overestimated true validity and variance, but only by small amounts. Later studies have found similar results (e.g., Kemery, Mossholder & Roth, 1987; Spector & Levine, 1987).

In the original procedure of Schmidt & Hunter (1977) the observed correlations, (rs), were transformed to Fisher's z prior to carrying out the calculations. This transformation was included to ensure that sampling error would not be correlated with true validity. Without the r to z transformation, the distribution of observed validities is usually negatively skewed (James, Demaree & Mulaik, 1986), and the estimation of "true" validity and error variance is biased.

Presently, however, the r to z transformation is seldom used by validity generalization researchers. Several researchers have investigated the effects of using untransformed rs (James, et al., 1986; Silver and Dunlop, 1987; Strube, 1988) and have found them inconsequential for large samples. James et al. (1986) provided evidence showing that the error variance is underestimated when the rs are not transformed. While Schmidt, Hunter and Raju (1988) agreed with these results they pointed out that for ns larger than 40, the bias will always be less than .005.

Validity generalization procedures appear to be reasonably accurate when the hypothetical distributions are representative of the data collected and the correlational assumptions are met. However, the procedures can be misleading when these hypothetical distributions are not representative of the data. The consequences of assuming unrepresentative hypothetical distributions have been investigated by Paese and Switzer (1988). This was a particularly important study, since previous Monte Carlo studies have assumed that the distributions were representative. In their study Paese and Switzer (1988) demonstrated that artifact variance can be overestimated or underestimated when the hypothetical distributions are not representative of the data, and they recommended caution in using hypothetical distributions. Hedges (Note 1) has also been critical of using assumed distributions and has proposed a method based on situational data. His method in essence is analogous to correcting the rs first and then computing the variance of the corrected rs.

OBJECTIVES

In our study we compared Hedges' (Note 1) validity generalization procedure with others when the procedures are based on situational data only, without the benefit of assumed distributions. The purpose of the study was to assess the accuracy of the validity generalization procedures when only sample data are utilized, and to present a statistical foundation for the application. While Callender and Osburn (1981) have considered the accuracy of validity generalization procedures with situational estimates, the focus of their study was not on the statistical properties of the procedures. We present an argument for, and explore some of the problems of, this approach. Then, we compare the validity generalization procedures using a Monte Carlo simulation. The procedures were compared under two sets of conditions: those with complete data, and those with missing data. The missing data case simulated a validity generalization study where some of the artifacts estimates were missing.

STATISTICAL FOUNDATION

In this section we review and integrate the research providing a statistical foundation for validity generalization procedures, then we show how the foundation can be expanded for situational estimates. Callender and Osburn (1980) presented a model for validity generalization based on the following decomposition of the observed correlation coefficient

$$r_i = \rho_i + e_i$$

where r_i is a sample correlation coefficient based on n_i subjects observed in the i^{th} situation ($i=1,...,k$), p_i is the population correlation in this situation, and e_i is the error of estimate. While Schmidt and Hunter (1977) did not explicitly specify a model in their seminal article, their validity generalization procedure in essence assumes the same model.

The number of situations k, in a validity generalization study, can be conceptualized either as a random sample from a large finite or infinite, number of situations (Hedges & Olkin, 1985, p.243; James, Demaree, Mulaik & Mumford, Note 2), or as a small number of situations which we can observe. This conceptualization of k is analogous to assuming, respectively, a random or fixed model in the analysis of variance. A number of validity generalization papers have not been clear about this point, but it appears that most of them implicitly have assumed the fixed model. The results given in this section hold whether the fixed model or random model is assumed. However, we will also assume that k is finite and that all situations are observed to simplify the statistical presentation. Further, for the present we will assume that the observed correlations have not been attenuated by unreliability of measurement, or restriction in range. As we develop the model we will relax these assumptions.

In their derivation, Callender and Osburn (1980) assumed e and p to be uncorrelated over independent situations in order to express the variance of the observed correlations as the sum

$$V(r) = V(p) + V(e) \qquad (1)$$

But, Hedges (Note 1) has shown, assuming normality for the X and Y scores being correlated, that e and p are correlated. Fortunately, this correlation is in general small and negative, especially when the n's are large. Specifically, Hedges showed that the covariance between e and p is given by

$$Cov(e,p) = [E(p^4) - E(p)E(p^3) - V(p)]/2(n-1)$$

where E denotes the expectation operator. In a later article Hedges (1989) showed that even when one uses an unbiased estimator of the correlation p, say R, (in the validity generalization calculations) and thus eliminate the covariance between p and e, the results obtained with R, in general, do not differ by much from those obtained with r. Hedges (1989), however, suggested using the unbiased estimate R in validity generalization when all of the n's are small.

Being able to express the variance of the correlation as a sum of two components (Equation 1), suggests that V(p) can be estimated with the difference between sample variances

$$\hat{V}(p) = \hat{V}(r) - \hat{V}(e)$$

Since error variance V(e) is the average of the situation error variances, the error variance can be written as

$$V(e) = E[V(e_i)]$$

When the number of situations k is finite, the error variance takes the form

$$V(e) = \frac{1}{k} \sum_{i=1}^{k} V(e_i) \qquad (2)$$

Under bivariate normality, $V(e_i)$ can be approximated by

$$V(e_i) = (1-\rho_i^2)^2 / (n_i-1)$$

Consequently, validity generalization researchers have suggested

$$\hat{V}(e) = \frac{1}{k} \sum_{i=1}^{k} [(1-r_i^2)^2/(n_i-1)]$$

as an estimator of V(e).

However, according to Thomas (1988), $\hat{V}(e)$ is a biased estimator of V(e), since

$$E[\hat{V}(e)] = \frac{1}{k} \sum_{i=1}^{k} [1 + E(e_i^4) - 2E(e_i^2)]/(n_i-1) \qquad (3)$$

(Looking at Equations 2 and 3, we can tell that V(e) is not $E[\hat{V}(e)]$.) The bias in $\hat{V}(e)$ while usually small, can occasionally lead to some under or over estimation of V(ρ), especially when the ns and k are small (Thomas, 1988).[1] V(r), of course is estimated by

$$\hat{V}(r) = \frac{1}{k} \sum_{i=1}^{k} (r_i - \bar{r})^2,$$

concluding the estimation process with,

$$\hat{V}(\rho) = \hat{V}(r) - \hat{V}(e).$$

The $\bar{r}$ is the mean of the observed rs. Raju and Burke (1983) suggest a mean weighted by sample size.

Callender and Osburn (1988) also examined the accuracy of estimating the error variance V(e). They varied both the number of rs (number of studies in a

[1] More disturbing is the fact that Thomas (1988) gave some examples (with small ns and k) where the probability of obtaining a small difference, $\hat{V}(r) - \hat{V}(e)$, is larger when the ρ's are unequal than when all are equal. In terms of hypothesis testing, these results imply that under certain conditions the test of hypothesis based on the difference $\hat{V}(r) - \hat{V}(e)$ is biased. (We use bias here not in the estimation sense, but in the testing sense; see Kendall and Stuart, 1960, p.209)

validity generalization study) as well as the ns in each r (number of subjects in each validity generalization study). Their results (reported in terms of standard deviations) showed that the estimate $\hat{V}(\rho)$ tends to be slightly smaller than the actual variance $V(\rho)$, "but the differences are in the third decimal place" (1988, p. 314). For correlations over .5 the differences averaged .002 while for those below .5 the differences averaged .001. In a similar fashion, Millsap (1988), assessed the sampling variance estimate, but, introduced attenuation due to criterion reliability into the Monte Carlo. He concluded that the slight negative bias did not have significant effect if $n > 60$.

Next, consider a validity generalization study where the k observed correlations have been attenuated by unreliability of measures and by direct restriction in the predictor's range. Raju and Burke (1983) following Callender and Osburn (1980) have suggested the following model for this condition

$$r_i = \rho^*_i + ei \qquad (4)$$

where

$$\rho^*_i = \rho_i \; a_i \; b_i \; c_i;$$

a_i is the square root of the criterion reliability, ρ_{xx}; b_i is the square root of the predictor reliability, ρ_{yy}; and c_i is a function of u_i, a_i and b_i, given by

$$c_i = \frac{u_i}{[1 + (u_i^2 - 1)\, \rho_i^2\, a_i^2\, b_i^2]^{1/2}}$$

The u_i is the ratio of the restricted standard deviation of the predictor to the unrestricted standard deviation of the predictor, and ρ_i is the unrestricted,

unattenuated, population correlation for the i^{th} situation. Note that all of the parameters given above are defined within a situation. Equation (4) can be derived by applying the usual psychometric corrections for attenuation and range restriction (see Raju & Burke, 1983; and Mendoza & Mumford, 1988).

Based on Hedges' (1989) results, we know that ρ^* and e in Equation 4 are generally uncorrelated, and that E(e) = 0, or that we can make it so using an unbiased estimator. Under these conditions, p, a, b, and, u are uncorrelated with e when n is large (see Appendix A). Consequently, Raju & Burke's (1983) Taylor expansion for V(p) holds; that is

$$\begin{aligned} V(\rho) \simeq\ & A^2V(\rho) + B^2V(a^2) + C^2V(b^2) + D^2V(u) + AB\ Cov(\rho,a^2) + \\ & AC\ Cov(\rho,b^2) + AD\ Cov(\rho,u) + BC\ Cov(a^2,b^2) + \\ & BD\ Cov(a^2,u) + CD\ Cov(b^2,u) + V(e) \end{aligned} \qquad (5)$$

where

$$A = .5 \times [\bar{r}/E(\rho) + (\bar{r})^3 w/E(\rho)E(u)^2]$$

$$B = .5 \times [\bar{r}/E(a^2) + (\bar{r})^3 w/E(a^2)E(u)^2]$$

$$C = .5 \times [\bar{r}/E(b^2) + (\bar{r})^3 w/E(b^2)E(u)^2]$$

$$D = \bar{r} - (\bar{r})^3/E(u), \quad W = 1 - E(u)^2,$$

$$E(\rho) = \frac{\bar{r}}{\{E(a^2)E(b^2)[E(u)^2 + (\bar{r})^2 w]\}^{1/2}}$$

and $\bar{r}$ is the mean of the observed correlations. Note that this is not the only possible representation of V(p); others are presented in Raju and Burke (1983) and Callender and Osburn (1980). Also note that the means, variances and covariances in Equation (5) involve the unobservable parameters a_i, b_i, c_i, and u_i. In most validity generalization studies, at best, what we have are the

estimates $\hat{a}_i$, $\hat{b}_i$, $\hat{c}_i$, and $\hat{u}_i$, or a set of assumed distributions for a, b, c, and u. We discuss next how one can use situation statistics to estimate the means, variances and covariances appearing in Equation (5).

To begin with, the three covariances involving p (the unrestricted, unattenuated population correlation) in Equation 5 are not estimable (James et al., Note 2), since p is not observable, and any estimate of p would necessarily involve the parameters a_i, b_i, and u_i (or the estimates $\hat{a}_i$, $\hat{b}_i$ and $\hat{u}_i$). Unlike the covariance between p and e whose effect on the estimation of V(p) depends on n, the effect of these covariances on the estimation of V(p) does not depend on n. Kemery et al. (1987) have shown how one can have a nontrivial covariation between p and the other parameters. Nevertheless, most validity generalization procedures have assumed that these covariances are zero.

On the other hand, if data are available for each situation, the covariances not involving p can be estimated using sample estimates. For instance, consider the covariance between the squared criterion and predictor reliabilities in Equation (5), $Cov(a^2,b^2)$. It is not difficult to show that

$$Cov(a^2,b^2) = Cov(\hat{a}^2,\hat{b}^2),$$

when the estimates of criterion and predictor reliabilities are unbiased in each situation (conditionally). Note that the estimates $\hat{a}_i$ and $\hat{b}_i$ must be from unrestricted samples (Mendoza & Mumford, 1988; Callender & Osburn, 1980). In the case of the criterion reliability, $\hat{a}_i$, this posses a special problem because unrestricted estimates are difficult to obtain. A criterion reliability estimate, corrected for restriction in range, could be used, but this estimate would likely

be biased. It is important to note that the estimate of the predictor reliability, $\hat{b}_i$, cannot be corrected for range restriction and must be from an unrestricted sample (Mendoza and Mumford, 1988). This caveat aside, it is clear that in certain cases we will be able to estimate the covariances between the artifacts.

The estimation of the means and variances in Equation (5) does not present a problem, if in each situation the estimates, $\hat{a}_i$, $\hat{b}_i$, and $\hat{u}_i$ are unbiased and consistent. For example, if $\hat{a}_i$ is an unbiased (even asymptotically) and consistent estimator of a_i, then for large n

$$E(\hat{a}) \approx E(a), \text{ and}$$

$$V(\hat{a}) \approx V(a).$$

Specifically, under the conditions specified above, it can be shown that

$$\lim_{n\to\infty} E(\hat{a}) = E(a), \text{ and}$$

$$\lim_{n\to\infty} V(\hat{a}) = V(a).$$

Hence, when n is sufficiently large, $V(\hat{a})$ can be used to estimate V(a), provided that $\hat{a}_i$ is available for all or most of the situations. Again, if a corrected estimator, say, $\hat{a}^*_i$ is used to estimate the criterion reliability, it is not clear whether V(a) would equal $V(\hat{a}^*_i)$. Even if the equality holds asymptotically, n may have to be very large before one can use $V(\hat{a}^*)$ to accurately estimate V(a). When parameter estimates are not available for each situation (which according to Jones, 1950; and Schmidt & Hunter, 1977 is the more typical case), the estimates are obtained from the assumed distributions. The accuracy of the assumed distributions, however, depends on the accuracy of the data used to create the assumed distributions. The estimates used to

create these distributions should be unbiased, consistent and from unrestricted samples. This issue has been largely overlooked.

It is beyond the scope of this paper to deal with the many issues involved in testing hypothesis in validity generalization studies. The interested reader is referred to James et al. (1988), Schmidt, Hunter, Pearlman & Hirsh (1988), Kemery et al. (1987), Sackett et al. (1986), and Spector and Levine (1987) for more details. Our concern in this paper is with the accuracy of estimation and not with hypothesis testing in general. We understand, however, their relationship.

It should be clear from the above presentation that if:

- n is sufficiently large
- X and Y are at least approximately normally distributed
- all covariances involving p are zero, and
- appropriate estimators of a_i, b_i and u_i are available

we can estimate V(p) from the mean and variances of the parameter estimates. However, two questions still remain: how accurate can we estimate V(p) and, how large does n have to be? These are the topics which follow.

III. METHOD

A FORTRAN program was written to simulate data under a number of conditions and to apply six validity generalization procedures. The program first independently sampled the parameters p, a, b, and u from a preselected set of distributions and then obtained a random sample conforming to these parameters. The parameter distributions used in the study are shown in Table

1 and some are depicted graphically in Figures 1, 2 and 3. Each condition tested in the study consisted of a combination of these distributions. The conditions are listed in Table 2.

Specifically, after p_i, a_i, b_i, and u_i were sampled the program used these parameters to create a multivariate normal distribution for $Z_i = (X_t\ Y_t\ X_e\ Y_e)$ with mean zero and variance

$$V = \begin{matrix} 1 & p_i & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1-b^2 & 0 \\ & & & 1-a^2 \end{matrix}$$

The elements X_t and Y_t in the vector represent the trait (true score) portion of each score, and X_e and Y_e represent the error portion. Using data from a random sample from this multivariate normal distribution we obtained the estimates $\hat{V}(X)$, $\hat{V}(Y)$, $\hat{V}(X_t)$, $\hat{V}(Y_t)$ and

$$\hat{a} = \frac{\hat{V}(Y_t)}{\hat{V}(Y)} \qquad \hat{b} = \frac{\hat{V}(X_t)}{\hat{V}(X)}$$

Next, the data were ranked on X ($X = X_t + X_e$) and the upper u% (selection ratio) of the sample were selected. From this restricted sample we obtained

$$\hat{V}(X'), \quad rxy, \quad \text{and} \quad \hat{u} = \frac{[\hat{V}(X')]^{1/2}}{[\hat{V}(X)]^{1/2}}$$

Note that the $\hat{u}$ is the ratio of the standard deviation of X in the restricted sample to the standard deviation of X in the unrestricted sample. The process

described above was then repeated k times to generate data for a validity generalization study.

The mean and variance of these estimates (as well as the estimates themselves) were then used in the six validity generalization procedures: the Non-interactive (NIA) and Interactive (INT) procedures (Schmidt & Hunter, 1977; Schmidt et al., 1980), the Independent Multiplicative (INM) procedure (Callender & Osburn, 1980); TSA1 and TSA2 (Raju & Burke, 1983); and a modified version of Hedges' (Note 1) procedure. The computational formulas for all of the procedures except for Hedges were taken from Raju and Burke (1983).[2]

Hedges (Note 1) originally recommended transforming to Fisher zs prior to estimating the true variance. Recently Hedges (1988) suggested using nontransformed coefficients. To incorporate his recent suggestion we dropped the z-transformation, and modified the conditional error variance $1/(n'_i-3)$ used in the EM algorithm of his procedure as follows

$$\frac{u^{-2}(1-r^2)^2}{n'_i \hat{\rho}_{xx} \hat{\rho}_{yy}(1+r^2(u^{-2}-1))} \qquad (6)$$

A "missing data" condition was simulated by deleting (losing) 50% of the available artifact estimates. Since data for validity generalization studies can be unavailable in a variety of ways, (eg., predictor reliability was reported, but, criterion reliability and selection ratio were not given; versus none reported) one must decide on a specific manner in which to delete the data. We used a simple procedure. The missing data condition represents a validity

[2] See pages 386-387 and 394-395 for the complete formulas and correction.

generalization study in which 50% of the situations contained no artifact information (all three estimates were missing) while the other 50% contained estimates of all three artifacts (Note that this is not exactly the same as having validity generalization studies with half of the rs.) This method assures that Hedges' procedure would be based on the same proportion of missing data as the other procedures. The results were accumulated over 250 replications.

Hedges' procedure for estimating true validity and variance utilizes the reliabilities and the range restriction from each situation. When these are available an estimate of the true validity and variance are produced via a maximum likelihood iterative approach. This procedure results in an estimate referred to as Hedges' "T". However, when either of the reliabilities or the range restriction is missing from a situation, T can not be calculated. Hedges (Note 1) gave a method for replacing the missing Ts with predicted $\hat{T}$s based on a regression equation between the rs (observed validities) and the Ts. For situations in which artifact information is missing the conditional error variance, as given by Equation 6, can not be computed and thus must also be estimated. The conditional error variance was estimated using a regression equation between the rs and the conditional error variances calculated when the Ts were available.

A difference between the current simulation and previous ones is in the generation of the simulated data. In most previous simulations the mean and variance of the assumed distributions (representing the true validity, criterion and predictor reliability and restriction in range) were used in the calculations of

the validity generalization procedures (Callender & Osburn, 1980; Raju & Burke, 1983; Paese & Switzer, 1988). This results in variance estimates for artifacts which do not change across validity generalization studies. While the current study also uses theoretical population distributions to generate data, we did not use the mean and variance of these distributions [E(a), E(b), E(u), V(a), V(b), V(u)] in the calculations of the validity generalization procedures, instead we used the estimates $E(\hat{a})$, $E(\hat{b})$, $E(\hat{u})$, $V(\hat{a})$, $V(\hat{b})$, & $V(\hat{u})$. A similar approach was used in the later simulations of Callender (Callender & Osburn, 1981; Callender et al., 1982).

CONDITIONS SIMULATED

To establish a connection with previous research the first application of our simulation was to replicate the conditions reported by Raju & Burke (1983). We refer to these as calibration runs, since their main purpose was to confirm the accuracy of our simulation. These runs were based on conditions 1-10 from Table 2. Results for Conditions 1-4, 5-9, and 10, were compared with the results reported by Raju and Burke (1983) for "Case 1", "Case 2", and "Case 3", respectively. The result of these comparisons is presented later.

For Case 1, true validity was held constant in each run, while allowing the reliabilities and range restriction to vary. The means and variances of the artifact populations used in these calibration runs were those suggested by Schmidt & Hunter (1977) and used by Raju & Burke (1983) (See the distributions labeled CR060L, PR080L & SR065L in Table 1.) While Raju and

Burke (1983) used these artifact distributions to attenuate nine populations of true validity with means of (.1)(.1)(.9) and zero variance, we used .0, .1, .5, and .9 (Conditions 1-4 in Table 2). Each Case 1 calibration run used samples of size 500, and 100 situations (rs).

In Case 2, (conditions 5-9) only the true validity varied. However, in spite of the artifact population variances being zero, the estimates of the population artifacts ($\hat{a}_i$, $\hat{b}_i$, and $\hat{u}_i$) contained sampling error variance, and therefore varied across situations.[3] Our calibration runs used the true validity distribution labeled TV050H in Table 1. This distribution has a mean equal to .5 and a variance of .082. This distribution is virtually identical to the distribution used by Raju and Burke (M=.5,V=.083). As with the previous conditions, these were examined with 100 rs in each study. However, to investigate the effect of sample size, each r was based on an n of 250 instead of 500. We limited our calibration here to five of the 27 conditions studied by Raju and Burke (1983), conditions 5-9 in Table 2.

Condition 10 duplicated the condition reported by Raju and Burke (1983) for Case 3. However, their results represents the average of nine trials (validity generalization studies) and ours is the average over 250 studies. This condition was run with sample size of 500 and number of situations (rs) of 100, 25 and 10.

[3] The only exception is when the variance of the artifacts is zero and the mean is one. In this case the artifact is estimated without sampling error; that is, the estimate is set to 1.0.

After establishing the accuracy of our simulation, we turned or attention to examining the accuracy of the six procedures under a variety of other conditions (See Table 2, Conditions 11-24) and the effects of missing data. This set of conditions represents a wide range of means for the population distributions of both true validity (0.0, 0.35, 0.5, 0.6, 0.9) and the artifacts (0.1, 0.6, 1.0). While all possible combinations of these values were not run, conditions 11-24 represent a set of runs which both cover the range of values which might be expected and the various combinations described by the "Case 1, 2 & 3" approach.

As noted previously, our simulation was based on parameter estimates containing sampling error variance. Thus, when we sampled from an artifact distribution with zero variance, there was still variance in the estimate of the artifacts. Therefore, our Cases 1, 2, and 3 are not exactly like Cases 1, 2, and 3 as described by Raju and Burke (1983). We will refer to our cases as Case 1′, Case 2′ and Case 3′. An exception to this is found in conditions 21-24 where the reliability estimates were set equal to 1 resulting in no sampling variance for the reliability estimates. The only variance in these conditions was due to the sampling variance in r.

The calibration runs were based on 250 validity generalization studies, each with 100 situations (rs) per study. However, in conditions 11-24 we used only 32 situations per study. Few investigators have 100 situations (rs) to include in their validity generalization study. We think that 32 is a more realistic number. Furthermore, Paese and Switzer (1988) compared the interactive and non-

interactive procedures using 32 situations per validity generalization study, and in a study utilizing actual data from the petroleum industry Callender and Osburn (1981) indicated that 32 was the average number of rs (situations) per study when performance criteria were used. Settling on the 32 situations per validity generalization study allowed us to observe the validity generalization procedures under 32 studies in the complete data case and 16 studies in the missing data case.

IV. RESULTS

The calibration runs replicated accurately results from previous research. As noted earlier, our simulation contains sampling error not included by previous simulations. Therefore, to establish the accuracy of our simulation, rather than comparing means, we computed correlations between the results reported by Raju and Burke (1983) and those from our simulation. The correlation measured the relationship between our results and Raju and Burkes' (1983) results for a validity generalization procedure over situations. In other words, we computed a correlation for each procedure over the situations.

Table 3 reports the bias obtained in estimating the means for the six procedures in Case 1′. The correlations between these data and those reported by Raju and Burke (1983) were above .999 for all of the procedures. Similarly, Table 4 gives the bias for Case 1′ when the variance of true validity was estimated. These variance results correlated above .995 with Raju & Burkes' results. In addition to the correlation results, the estimates produced

by our simulation resemble closely those obtained by other researchers. The Non-Interactive Additive (NIA) produced variance estimates which decreased in accuracy as p increased. TSA1 produced the most accurate estimates under Case 1′ but underestimated variance when true validity was .5 or above.

For Case 2′ Raju and Burke reported that the five validity generalization procedures produced the same results for mean and variance when the selection ratio was 100%. This made calculating correlations impossible. However, Raju a[illegible]e showed that all the procedures estimated the parameters exactly when the selection ratio was 100% (and the artifacts did not vary, Case 2), the amount of bias in our results was also very small, verifying the accuracy of our simulation. The top 1/3 of Tables 5 and 6 give the bias for the Case 2′ conditions when the selection ratio was 100%. The average bias for these portions of Tables 5 and 6 are -.0015 and -.0022 respectively. This amount of bias is virtually the same as found by Raju and Burke (1983), however, it seems that the validity generalization procedures tend to underestimate the variance of true validity when there is but sampling error in the artifacts.

The results for the remainder of the Case 2′ (selection ratio < 100%) are given in the lower 1/3 of Tables 5 and 6. These results, when correlated with those of Raju and Burke (1983) yielded correlations above .999 for mean validity estimates. For estimates of variance where the selection ratio was 50% (middle 1/3 of Table 6), all the correlations were above .97. However, for estimates of variance where the selection ratio was 10% (bottom 1/3 of Table 6), the

correlations were .86, .82, .84. -.39 and -.40 for NIA, INT, INM, TSA1, and TSA2 respectively. In our simulation, estimates of variance decreased in accuracy as reliabilities decreased, rather than increasing in accuracy as reliabilities decreased, as reported in Raju and Burke (1983).

For Case 3, Raju and Burke (1983) examined only one set of conditions. Tables 7 and 8 give respectively the bias in mean validity and variance under Case 3′. The average difference between our results and those reported by Raju and Burke (1983) for this condition was .004. The correlations between our results and those of Raju and Burke were .97 and .39 for mean and variance results, respectively.[4]

In addition to establishing that our simulation accurately replicates previous research, the calibration runs also provided information about the relative accuracy of Hedges procedure. For Case 1′, Hedges' procedure was more accurate in estimating the mean validity than the other procedures. Under this condition variance was slightly overestimated by Hedges' as opposed to being (in general) slightly underestimated by the other procedures. Hedges' estimates of both mean validity and variance tended to maintain more accuracy as p increased than the other procedures.

For Case 2′, Hedges' procedure again tended to be as accurate as the other procedure, but only at selection ratios of .5 or greater. At a selection ratio of 50%, Hedges' procedure overestimated mean validity by an average of .016 while the other procedures overestimated by .019 on average. In a similar

[4] The correlations calculated here were across procedures, since only one condition was examined..

fashion, the variance estimates of Hedges' averaged .002 under the .0816 actual variance while the other procedures over estimated by an average of .013 . When the selection ratio was decreased to 10% all the procedures became more inaccurate. However, Hedges' procedure lost more of its accuracy than the others. The average bias in the mean validity estimate from the other procedures was .031 while Hedges' overestimated by .055. Even more striking were the overestimates of variance produced by Hedges' procedure. For conditions where Hedges' procedure overestimated the variance (.0816) by .162 on average, the other procedures overestimated the variance by .067.

For the Case 3 condition replicated, Hedges' procedure was generally more accurate in estimating both mean validity and variance. For estimates of mean validity, accuracy was generally maintained as the number of situations (rs) was decreased from 100 to 25 to 10. However, when the number of situations was decreased, Hedges' lost more accuracy than the other procedures when estimating variance. Interestingly some estimates of variance from the other procedures became more accurate as the number of studies decreased while Hedges' became less accurate.

The remaining results are organized in terms of selection ratio and conditions and represent conditions wwith only 32 situations (rs). Table 10 presents the bias of the validity generalization procedures in estimating the variance in true validity when the predictor and criterion contained error of measurement and the selection ratio was set to .1. Similarly, Table 12 presents

bias results for estimating variance with the validity generalization procedures when the predictor and criterion were error free and the selection ratio was .1. The bias in estimating the means under these conditions are given in Tables 9 and 11. Table 9 reports bias for unreliable measures, whereas Table 11 reports bias for the error free measures. Tables 13 and 14 give the mean and variance results, respectively, when the selection ratio was .6 and the measures contained error. In addition, a number of conditions that were run with a selection ratio of 1 are not presented in the tables, since the bias was virtually zero for all of the validity generalization procedures when the sample size was 125 regardless of condition or selection ratio. Two general results can be seen from these tables. First, the bias in all of the validity generalization procedures decreased as (applicant) sample size increased. Second, the validity generalization procedures were more accurate in estimating the mean than in estimating the variance of the true validity distribution.

The choice of parameter distributions in general did not moderate the bias in the procedures. A look over Tables 9 through 14 reveals that there was not much variability in bias over the conditions tested. Bias was much more dependent on the selection ratio and sample size than on the choice of distribution. Generally, in conditions where the predictor and criterion were error free (Tables 11 & 12) the bias tended to be smaller than in those conditions were there was unreliability (Tables 9 & 10). However, for the error free conditions where the variability of true validity was high (Condition 18), the bias was greatest and as large as the bias found in the conditions with

unreliability (Conditions 11-16). The reader should note that in conditions where there is no artifactual variability, the NIA and INT procedures converge to a common value as do TSA1 and TSA2 (Tables 11 & 12).

The selection ratio was the most pernicious of all of the factors investigated, as the selection ratio decreased the bias in the validity generalization procedures increased. When the selection ratio is .1, one needs at least 250 applicants to obtain accurate estimates. Figure 4 depicts the average accuracy of the validity generalization procedures in estimating true validity variance when the measures contain error and the selection ratio is .1. The figure was obtained by averaging over Table 10. Figure 4 clearly illustrates the effect of sample size on the estimation of variance. The bias in the validity generalization procedures was in the neighborhood of .10 when the (applicant) sample size was 125 and the selection ratio was .1, except for the Hedges procedure which had a larger bias. This is not surprising considering that Hedges' procedure is based on single observations and not aggregated variances.

The effect of sample size on variance estimation can also be seen in Figure 5. This figure, as the previous one, presents average bias results. Specifically, the results given in Figure 5 are the average of those found in Table 12. The validity generalization procedures were more accurate when the predictor and criterion were error free. For most of the validity generalization procedures the effect of error free artifacts was to cut the bias in half. When the sample size was 125 bias was .05, half of the bias in Figure 4.

The inverse relationship between selection ratio and bias, which can be seen by comparing results across the tables, does not seem to have been caused by sample size alone. If the relationship between bias and selection ratio was just due to sample size, one would expect the bias in a condition with 250 applicants and a selection ratio of .1 to be equivalent to a condition with 25 applicants and a selection ratio of 1. However, when we duplicated the conditions given in Table 10 (the results are not presented in the tables) with a sample size of 25 and a selection ratio of 1, the largest biases encountered in a condition were -.0101 for NIA, -.0094 for INT, -.0110 for INM, - .0110 for TSA1, -.0041 for TSA2, and .0185 for Hedges. These values are lower than any found in a row of Table 10 for a sample size of 250 (and selection ratio =.1). In spite of having effectively equal sample sizes, the condition with the smallest selection ratio always had the larger bias. Similar results were found for the mean. Since the reliability estimates in fact were based on a larger sample size (250) when the selection ratio was .1 than when the selection ratio was 1 (25), the increased in bias in the small selection ratio condition suggest a decreased accuracy in the estimation of the sampling variability of r and $\hat{u}$.

Results from the incomplete data conditions paralleled those from the complete data conditions. Surprisingly, the validity generalization procedures estimated as well with missing data as with complete data (Tables 9-14). The exception was the Hedges procedure which was not effective in the missing data condition when the selection ratio was .1. Although the bias results were similar under the complete and incomplete data cases, the standard errors (not

reported here) were not. The standard errors tended to be slightly larger (in the magnitude of .005) in the missing data conditions. Figure 6, gives the average bias in estimating variance with missing data for TSA1. The points in Figure 6 represent the bias obtained with missing data when averaged over Tables 10 and 12. As we can see from this figure, the TSA1 procedure with a sample size of 250 and missing data, contained (on the average) a bias of .02 which compare favorably with the bias for complete data (Figures 4 & 5). Overall, the TSA1 procedure was the most accurate of the validity generalization procedures.

V. DISCUSSION

Most of the validity generalization procedures accurately estimated the mean and variance of the true validity distribution even when some artifactual data were missing. For the most part, the validity generalization procedures were accurate when applied without assumed artifactual distributions. Although studies have shown that the validity generalization procedures in general are accurate when used with suitable artifactual distributions, if the artifactual distributions are not representative, the procedures can be in substantial error.

The accuracy of the validity generalization procedures depended mainly on the selection ratio. Selection ratios of 1 generally provided estimates with small biases even when the sample size was as low as 25 (applicants). In contrast, selection ratios of .1 yielded the largest bias. When the (applicant) sample size

was 125 or over and the selection ratio was greater than .1 the bias in the validity generalization procedures was generally small. The bias was larger when estimating the variance. Although the reliabilities of the measures did not affect bias as much as the selection ratio, there was a tendency for the bias to increase as the measures became more unreliable. Most of the time the procedures were positively biased.

TSA1 and TSA2 were generally superior to the other validity generalization procedures, TSA1 being slightly better than TSA2. For example, when the selection ratio distribution had a mean of .1 (and no variance) the TSA1 procedure needed 250 applicants per study for accurate estimation, whereas Hedges' procedure required at least 500. In addition, TSA1 did not require many studies to accurately estimate mean and variance. In cases where both sample size and selection ratio were moderate or sample size was large and the selection ratio was small, TSA1 retained accuracy with as few as 12 studies. This suggest that the quality of data obtained for a validity generalization study are much more important than the quantity. It is better to have a few good validation studies even if some of the artifactual information is missing than many studies with questionable data. This supports previous research where k (the number of studies) has been found to be less important than the sample size per study.

VI. CONCLUSIONS

Our results and those of Callender and Osburn (1981) justify the use of validity generalization procedures with situational data. We have seen that the procedures do not require a great number of studies but a few quality ones. An issue remaining for the applied research is that of obtaining an accurate estimate of predictor and criterion reliability. Obviously, if one were certain that a set of assumed distributions were representative, one should opt for the assumed distributions. But the representativeness of the assumed distributions has not being adequately addressed in the literature. Presently, the researcher is on his own without a guideline for acceptance, rejection, or construction of the assumed distributions. Consequently, we feel that in most cases, it is safer and more effective to conduct validity generalization studies with situational data. At any rate, it is important that we continue to investigate how the procedures perform with situational data and/or suggest ways in which to establish the suitability of assumed distributions. Also, we should examine the effect of different methods of estimating the artifact parameters on the validity generalization procedures.

Before closing we should say a few words regarding the accuracy of validity generalization procedures in light of their assumptions. The validity generalization procedures studied, with the exception of Hedges', rely on zero covariances between the artifacts and true validity for accuracy. As previously mentioned these covariances can not be estimated, but can effect the accuracy of the procedures. The simulation included neither covariances between the artifacts and true validity, nor correlation between the artifacts. However, the

correlations between artifacts could be estimated and incorporated in Equation 5.) In the current study, we assumed a set of reasonable, but favorable conditions, for the procedures in the simulation. This was done because most previous simulations with assumed distribution utilized similar conditions. We wanted to be able to compare our results with those of other investigators. However, before fully endorsing validity generalization procedures, more research is needed to explore the consequences, if any, of violating these assumptions, and those regarding estimates of artifacts.

We have demonstrated the accuracy of the validity generalization procedures with situational data when their assumptions are met. It is not clear, however, at this point whether these assumptions are usually met in applied settings such as the Armed Services. Until we know more about the likelihood of meeting these assumptions, applied validity generalization studies should be interpreted with caution. Essential to this interpretation is that we have well-behaved estimators (estimators which are consistent and estimate the unrestricted reliability of the measure) of reliability in our study. As mentioned earlier, obtaining well-behaved estimates especially of the criterion reliability could in some situations prove difficult. Furthermore, more research is needed to determine the robustness of the validity generalization procedures to the correlation of the artifacts with true validity. Even though this is not an assumption of the Hedges' procedure, it appears that his procedure needs some refinements especially in the case of missing data before it can be used freely.

VII. RECOMMENDATIONS

We feel that validity generalization studies conducted by the Air Force should not be based on assumed distributions but on situational data. If assumed distributions are used particular care should be put in the derivation of this distributions. It is important to know what type and kind of information went into the construction of these distributions. When using situational data, the emphasis should be in the quality of the studies used in the validity generalization study and not so much on quantity. However, the investigator should be careful not to introduced a sampling bias with this sampling plan. Until more research is conducted, it appears that the TSA1 procedure is the most robust of the validity generalization procedures available and we recommend its use. Hedges' procedures shows promise but needs further development.

We recommend that the Air Force continue its research efforts in validity generalization procedures. It is clear that these procedures could be useful but it also clear that more research is needed before they can be fully endorsed, or investigators can generate new procedures. Additional research should pay particular attention to the criterion reliability problem, the correlation between the artifacts and true validity, and to the possibility of improving Hedges' procedure.

Reference Notes

1. Hedges, L. V. (1988). The meta-analysis of test validity: Some new approaches. Paper presented at the Conference on Test Validity for the 1990's and Beyond. Princeton, New Jersey.
2. James, L. R., Demaree, R. G., Mulaik, S. A., & Mumford, M. D. (1988). Validity generalization: 10 years of suspended judgement. Unpublished manuscript, Georgia Institute of Technology.

References

Callender, J. C., & Osburn, H. G. (1980). Development and test of a new model for validity generalization. Journal of Applied Psychology, 65, 543-58.

Callender, J. C. & Osburn, H. G. (1981). Testing the constancy of validity with computer-generated sampling distibutions of the multiplicative model variance estimate: Results for petroleum industry validation research. Journal of Applied Psychology, 66, 274-281.

Callender, J. C., & Osburn, H. G. (1988). Unbised estiamtion of sampling variance of correlations. Journal of Applied Psychology, 73, 312-315.

Callender, J. C., Osburn, H. G., Greener, J. M., & Ashworth, S. (1982). Multiplicative validity generalization model: Accuracy of estimates as a functin of sample size and mean, variance, and shape of the distribution of the true validities. Journal of Applied Psychology, 67, 859-867.

Ghiselli, E. E. (1966). The Validity of Occupational Aptitude Tests. New York: Wiley.

Hedge, L. V. (1988). An unbiased correction for sampling error in validity generalization studies. Journal of Applied Psychology. (in press).

Hedges, L. V., & Olkin, I., (1985). Statistical Methods for Meta-Analysis. Orlando, Florida: Academic Press.

James, L. R., Demaree, R. G., & Mulaik, S. A. (1986). A note on validity generalization procedures. Journal of Applied Psychology, 71, 440-450.

Jones, M. H. (1950). The adequacy of employee selection reports. Journal of Applied Psychology, 34, 219-224.

Kemery, E. R., Mossholder, K. W., & Roth, L. (1987). The power of the Schmidt and Hunter additive model of validity generalization. Journal of Applied Psychology, 72, 30-37.

Kendall, M. G., & Stuart, A. (1973). The Advanced Theory of Statistics. New York: Hofner.

Millsap, R. E. (1988). Sampling variance in attenuated correlation coefficients: A Monte Carlo study. Journal of Applied Psychology, 73, 316-319.

Mendoza, J. L., & Mumford, M. (1987). Corrections for attenuation and range restriction on the predictor. Journal of Educational Statistics, 12, 282-293.

Paese, P. W., & Switzer, F. S. (1988). Validity generalization and hypothetical reliability distributions: A test of the Schmidt-Hunter procedure. Journal of Applied Psychology, 73, 267-274.

Raju, N. S., & Burke, M. J. (1983). Two new procedures for studying validity generalization. Journal of Applied Psychology, 68, 382-395.

Sackett, P. R., Harris, M. M., & Orr, J. M. (1986). On seeking moderator variables in the meta-analysis of correlational data: A Mote Carlo investigation of statistical power and resistance to Type I error. Journal of Applied Psychology, 71, 302-310.

Schmidt, F. L., Gast-Rosenberg, I., & Hunter, J. E. (1980). Validity generalization results for computer programmers. Journal of Applied Psychology, 65, 643-661.

Schmidt, F. L., Hunter, J. E., & Pearlman, K. (1982). Progess in validity generalization: Comments on Callender and Osborn and further developments. Journal of Applied Psychology, 67, 835-845.

Schmidt, F. L., Hunter, J. E., Pearlman, K. & Hirsh, H. R. (1985). Forty questions about validity generalization and meta-analysis. Personnel Psychology, 38, 679-798.

Schmidt, F. L., Hunter, J. E., & Raju, N. S. (1988). Validity generalization and situational specificty: A second look at the 75% rule and Fisher's z transformation. Journal of Applied Psychology, 73, 665-672.

Silver, N. C., & Dunlop, W. P. (1987). Averaging correlation coefficients: Should Fisher's z transformation be used? Journal of Applied Psychology, 72, 146-148.

Spector, P. E., & Levine, E. L. (1987). Meta-analysis for integrating study outcomes: A Monte Carlo study of its susceptibility to Type I and Type II errors. Journal of Applied Psychology, 72, 3-9.

Strube, M. J., (1988). Averaging correlation coefficients: Influence of heterogeneity and set size. Journal of Applied Psychology, 73, 559-568.

Thomas, H. (1988). What is the interpretation of the validity generalization estimate $S^2_p = S^2_r - S^2_e$? Journal of Applied Psychology, 73, 679-682.

Appendices can be obtained from
Universal Energy Systems, Inc.

FINAL REPORT NUMBER 116
A NETWORK TUTOR BASED ON THE HEURISTIC OF POLYA
PENDING
Dr. Philip D. Olivier
760-6MG-032

GRAPHICAL PROGRAMMING OF SIMULATION MODELS IN AN OBJECT-ORIENTED ENVIRONMENT

Research Initiation Program
conducted for

Universal Energy Systems
and
Air Force Office of Scientific Research

by

Mufit H. Ozden

Department of Systems Analysis

Miami University

Oxford, Ohio 45056

January 15, 1990

SUMMARY OF RESEARCH WORK

Title: GRAPHICAL PROGRAMMING OF SIMULATION MODELS IN AN OBJECT-ORIENTED ENVIRONMENT

Principal Investigator: Mufit H. Ozden
Dept. of Systems Analysis
MIAMI UNIVERSITY
Oxford, OH 45056

Objective: The purpose of this research is to develop a graphical programming methodology for simulation in an object-oriented environment.

Research Period: 12 months in 1989

Funding: Total funding requested from the AFOSR RIP program is $ 20,000.

ACKNOWLEDGEMANTS

I wish to thank the Air Force Systems Command and Air Force Office of Scientific Research for sponsorship of this research and Universal Energy Systems for their administrative work.

Capt. Douglas Popken played an important role in orientation of this research and in keeping my enthusiasm high during the period of this research. I would like to express my sincere gratitude to him and wish the best success in the PRISM project.

GRAPHICAL PROGRAMMING OF SIMULATION MODELS IN AN OBJECT-ORIENTED ENVIRONMENT

by

Mufit H. Ozden

Department of Systems Analysis
Miami University
Oxford, OH 45056

ABSTRACT

Graphical programming, which is used here to mean creation of simulation models graphically, has been used in conjunction with conventional simulation languages via block diagrams or activity networks. Its beneficial effects on model development in simulation have been generally accepted. However, none of these conventional simulation languages has reached a level of graphical programming that would impact the user's programming task substantially. Today, this is possible with the current software and hardware technology. An interactive incremental programming environment supported by a good graphical programming system that helps automatic model development, specification and verification would be greatly appreciated in modeling in general, but especially in simulation of complex real-life systems. Such a visual system could essentially be a conceptual framework for analysis of the problem at hand and become a means of communication among the people who are involved in development and management of systems. It can actually be the programming facility forming a friendly interface between the modeler and the computer.

In this research paper, a prototype graphical programming methodology for modeling and automatic interpretation of simulation problems is developed in the object-oriented environment of the Smalltalk-80 language. The modeler uses a high-level graphical representation formalism based on the activity-cycle diagrams to define simulation problems. The activity-cycle diagrams constructed in a highly interactive mode are then interpreted into the underlying programming language and executed automatically. Thus, the modeler does not have to know the underlying Smalltalk-80 language in order to use our prototype simulation system. The future expansions to this system will include a model editor and output display systems.

I. BACKGROUND AND MOTIVATION

Logistics and Human Factors Division of the Air Force Human Resources Laboratory (AFHRL/LRL) is currently undertaking a study that will expand the capabilities of the Air Force in analyzing logistics support systems. As a part of the Productivity Improvements in Simulation Modeling (PRISM) project (Popken,1988), the system currently under study is an Integrated Model Development Environment (IMDE) which will create a state-of-art development and test system for the various simulation models of capability assessment in an object-oriented environment. The IMDE will consist of an integrated set of hardware and software tools which support model specification, development, and verification as well as specific functions such as data retrieval and update. The user-friendliness and programming efficiency will be the most important feature of such a development environment.

Today, it is a well known fact that , the software component usually forms the bottleneck in design and management of complex, computer-based systems and the old paradigms of software technology have not been able to offer a comprehensive remedy to this bottleneck problem. Graphical programming of simulation models in object-oriented environment mainly addresses some aspects of this extremely important issue in simulation and therefore is future oriented and experimental in nature. Due to the new and evolving nature of the technology and graphical programming concepts, the graphical programming methodology composed of existing and novel ideas will be created as a

prototype system before a full scale production system can be attempted. Because of the prevalence of this critical issue, this research has many fruits to bear for the systems under development as well as for the other software systems.

In my initial exploratory research at the Wright Patterson AFB during the summer of 1988, the types and benefits of graphical programming and the capabilities of several languages were studied for object-oriented simulation . The findings and recommendations of this research work have been compiled as a report submitted to AFHRL /LRL and UES (Ozden, 1988). As a result of this research, the future research needs of graphical programming of simulation models were identified for the PRISM project. As follow-up research, this graphical programming methodology has been developed as a prototype system for the AF logistics support systems. In the future, this methodology may be modified sightly for efficiency and should be expanded to include more high-level modeling concepts in order to serve as a basis for broader model development environment.

II. SIMULATION IN AN OBJECT-ORIENTED ENVIRONMENT

Although the object-oriented paradigm is a relatively new popular concept in software engineering, the idea of programming based on objects was first developed in Simula (Dahl and Nygaard, 1966), which is a simulation extension to the Algol-60 language. The basic idea is to modularize the programming tasks on the basis of abstract or physical objects of the system. The data structures and methods associated with an object are encapsulated within the object so that the only way its data can be accessed

or changed, or one of its methods can be invoked, is by sending an appropriate message to the object. Programming in this paradigm involves creating a set of objects with the proper methods that will be invoked at the appropriate time through message passing among these objects. An object-oriented language comes with its own abstract classes of objects which form together a programming environment. An object can acquire the data structures and methods that it does not contain specifically from its "superior" classes. This is called "inheritance" of data and methods, and it is one of the most important characteristics of the object-oriented environments. The inheritance provides a flexible programming environment that is organized in a hierarchical structure of object classes with reusable programs.

In the object-oriented paradigm, objects of the simulated world can interact with one another closest to their behavior patterns in the natural setting. These objects can be categorized into different kinds of classes. Objects created from each class will be similiar but not necessarily identical. This is actually a higher level of abstraction and more natural way of programming than it is possible with the procedure-oriented simulation languages, (Shannon, 1987).

Object-oriented simulation programs make excellent use of modern software engineering concepts, such as modularization, extensibility, inceremental and exploratory style of programming (Stairnmand and Kreutzer, 1988). This style of programming will be one of the essential characteristics of a rapid model development environment for complex simulation problems. The

availability of the computer languages using these advanced software concepts and the rapid developments in hardware technology enable us to employ graphical programming methodology for simulation models on work stations and some high-end personal computers. It has been suggested that with the advent of parallel computers, future simulation environments can be built on the object-oriented paradigm in which concurrency will be a natural extension, (Jefferson, 1984 and Unger, 1987), increasing the efficiency of such programs many fold.

Smalltalk-80 (Smalltalk) is one of the major object-oriented programming languages, (Goldberg and Robson, 1983). Everything in Smalltalk is an object organized as an instance of a class. Classes are arranged in a tree structure with each class having exactly one parent class. The root class of the tree structure is "Object". A subclass inherits all the variables and methods of parent class recursively. Simulation in Smalltalk is facilitated with the use of a set of abstract object classes. The prorammer uses some of these classes directly or may extend them with more specialized features through creating their subclasses. In a simulation study, a set of instances of these object classes are created to act according to the behavior patterns ascribed to the objects in the particular simulation situation. Smalltalk provides excellent support for discrete event simulation with its reusable general and simulation related classes, and with very powerful coding and debugging tools leading to high productivity in writing and modifying existing simulation applications. This same conclusion has also been reached by several other researchers, (Knapp, 1987; Bezivin, 1987; Ulgen and Thomasma, 1986).

Here, the Smalltalk language is selected as the computer language for the development of protype graphical programming methodology for the IMDE.

III. GRAPHICAL PROGRAMMING OF SIMULATION MODELS

Visual interactive simulation (VIS) is a term that has been used in connection with a simulation program which has features for graphical creation of simulation models (Graphical programming), dynamic display of the simulated system (Visual Output) , and the user interaction with the running program (User interaction), (O'Keefe, 1987; Hurrion, 1986). See Figure 1. In a graphical programming facility, a simulation model is created visually on the screen in an interactive and exploratory style. The Visual Display facility portrays the dynamic behavior of the system components on the screen usually as animated graphs. The User Interaction facility allows the user to interact with the running program. Interaction can be such that the simulation halts and requests information from the user, or the user stops the simulation at will and interacts with the running program. Recent research attributes various observed benefits to VIS, (Hurrion,1986; O'Keefe, 1987; Sargent, 1986; Ozden, 1988; Browne et al. 1986).

Simulation modeling is a complex task demanding both the creativity of the modeler and good support tools of the development environment. Briefly, it involves translation of some abstract problem view as conceptualized by the modeler into a computer program which is executable by the computer. In term of the simulation methodology that views the world as entities and

resources necessary for their activities, the modeler first needs to identify and form a conceptual system as a collection of simulation objects with their data structures and methods. The entities, called "simulation objects" here, are defined as any object that has a process to execute in the simulation model. During execution of activities of a process, resources may be consumed or simply needed by some activities, and/or resources may be produced at the end of some activities. Therefore, the resources in a simulation model serve to restrict and synchronize the activities of the simulation objects as closely to their natural setting as possible. The objects and resources may have real counter parts in the simulation problem or they may be abstract or imaginary to fashion the desired effects.

In this object-oriented simulation environment where the domain independent and domain dependent object classes exist together, simulation modeling amounts to creating the application-specific classes and the instances of all relevant object classes at the appropriate simulation time. The behavior pattern (process) of a simulation object is defined in terms of the methods which reside within the object and send proper messages to the other simulation objects. Although existence of abstract simulation object classes with the generic data structures and methods is very convenient for modeling, it still entails a considerable programming job to define the application specific classes and objects with correct processes in terms of the underlying programming language, i.e., in our case, a good deal of working knowledge with the underlying Smalltalk language is necessary.

With the graphical programming facility, the user would deal with the underlying language indirectly in an easier and more natural form. By means of graphs, icons, menus, windows and forms, a graphical programming interface guides the user in model specification in a structured and interactive form minimizing programming errors. The user can explore additioal features of the problem easily and increase the complexity of the model incrementally at will. When desired at any level of the model development, the abstract representation of the problem at hand can be interpreted and executed automatically.

The graphical programming approaches currently used by simulation languages can be classified in three groups:i) Network and block diagrams; ii) Icons, menus, forms and windows; iii) Dialogs and tree structured menus. These systems are described in the following sections.

Network and block diagrams have been used as a modeling and communication tool in conventional simulation languages, such as GPSS, SLAM, and SIMAN etc. Here, the activities of the simulation entities are described by a sequence of blocks or a network of nodes. But these are language-dependent representations and usually the number of blocks or nodes may be quite large, (e.g., over 60 for GPSS). So the modeler's job is to find the right sequence of these macro elements with the correct parameter assignments. This is far from being a straight forward task for many real-life simulation problems. Some problems may even require some external subroutines to be written in another programming language due to the restrictive programming features of these simulation languages.

The programming language interface with icons, menus, forms and windows has originated from artificial intelligence research because of its crucial need for friendlier programming environment for the type of problems studied. Some object-oriented languages extensively employ this convenient form of interface in which flexible, and reusable codes must be browsed and modified frequently in an interactive mode. In simulation, this type of programming style has recently been used in specific application areas of queueing networks, such as computer performance evaluation or manufacturing, (e.g., Melamed and Morris, 1985; Browne, et al. 1986; Sinclair et al. 1985; Duersch and Laymon, 1985; Stanwood et al. 1986).

Dialog-based programming is new in simulation. It has been developed as a part of a simulation environment, (Unger et al. 1984; Birtwistle and Luker, 1984). It originated from the idea that all simulation programs can have a structured model specification regardless of the application area. Therefore, a generic structured dialog with the user can be prepared beforehand to obtain the necessary information for any simulation model. A different dialog style programming was also developed for simulation in a restricted area by Ingalls, 1986. Here, the dialog is based on a set of menus structured in the form of a tree. The user chooses a path for model specification starting from the root of the tree toward the lower branches by picking up his choices from menus.

IV. THE PROPOSED GRAPHICAL PROGRAMMING METHODOLOGY

The role of the graphical programming in a simulation study is depicted in Figure 1. Ideally, a good graphical programming methodology is expected to meet the following criteria when used in a simulation environment:

a) It should facilitate easy use of the simulation environment;
b) The graphical programming scheme should itself be easy to use;
c) The modeler's productivity should be increased;
d) It should minimize programming errors;
e) It should enhance easy visualization of the conceptual problem;
f) It should serve as a good coomunication medium for people.

These criteria are certainly not in conflict with one another, and an improvement in one may very well mean betterment of others.

Specifically, the earlier exploratory research at AFHRL/LRL has identified the following facilities necessary for the graphical programming component of the IMDE :

a) A graphical programming editor that will create new object classes and graphical elements (icons, menus and forms etc.) to be stored in the simulation "library" and edit the old objects from the library and the simulation applications saved in the form of graphical models. It will have a "dictionary" access to the library of objects. The dictionary could be for most part iconic and organized in some hierarchical fashion for easy access.

b) A graphical model builder and interpreter that will be used to graphically build simulation models with the existing classes of simulation objects in the environment and automatically translate graphical models into computer executable programs.

c) View builder. The style of model development in this environment will be mostly exploratory and incremental. The objects and their interactions as created in part (a) should be able to be viewed graphically in a static and dynamic manner (e.g., as activity cycle diagrams). When the programs are run, the simulation with different display views should be able to be observed to facilitate verification of the model created so far. The display views may be created on the process of a selected simulation object with some kind of indication for the currently active activity (for example, reversed video), and at a desired speed the progression of its process can be observed dynamically. Also, display views may be created for the selected resources in order to observe the queues of simulation objects in front of them dynamically.

V. GRAPHICAL MODEL BUILDER AND INTERPRETER

In the analysis phase of a simulation study, graphical representation of a simulation problem is a very useful conceptualization and communication medium between humans. It can also become an important communication framework to specify simulation models to the computer. It is desirable that the graphical representation formalism be a simple, high-level system abstraction so that it enhances understanding and

conceptualization of simulation problems. But at the same time, it should have powerful conceptual features to be able to model all possible simulation situations. That is, on one hand it should contain all the concepts that are essential to represent all the logical relations and behavioral aspects of simulation objects that may exist in a simulation problem and on the other hand, the number of graphical tools representing these concepts should be small and high-level enough for easy human comprehension and use. The graphical representation scheme should not directly deal with the low-level concepts related to the simulation world views or underlying computer language. Once the main logical and structural representation of a simulation problem is accomplished, other necessary information, such as the input data related to a particular experimental run, can be obtained automatically from the user in some interactive form of communication.

In a simulation model, we can define mainly two types of simulation objects (SO) : permanent simulation objects (PSO) and temporary simulation objects (TSO). The SOs undertake a collection of activities as they play the roles of their dynamic behavior in simulation model. An activity is defined as any operation of a SO that takes some simulated time and during this time, it may require cooperation of other SOs, and may utilize and/or produce some amounts of certain resources. A SO may be either in an idle state waiting for other SOs for activity synchoronization, or for the needed resources to become available, or in a busy state, possibly tying up the resources and other SOs for the duration of an activity. A PSO repeats its

process to the end of some prespecified time (usually to the end of simulation), after entering the simulation environment at a specified time. There is usually a fixed number of the same kind of PSO in a simulation model. A TSO, however, arrives at the simulation environment randomly and executes its process and leaves the environment. Therefore, the number of TSOs in a simulation environment at a given time is a random variable.

Activity-Cycle Diagrams

In this research, we propose a modified version Activity-Cycle diagrams (ACD) as the graphical representation formalism. Basically, ACDs are directed networks of activities which have special features to represent essential simulation concepts, such as activity synchorization, resource utilization and production. ACDs are language independent, high-level graphical abstractions of stochastic systems that turn out to be extremely powerful graphical representation for simulation models, as also found to be the case by other researchers (e.g., Rodrigues, 1988; Birtwistle, 1979; Pidd, 1988).

In an ACD representing the behavior of a SO, the connections between activities may be a deterministic direct connection, or a conditional or probabilistic branching type. Also an activity may be initiated at completion of any of several activities independently. Figure 2 shows the primary elements of the ACDs adopted here.

An activity cannot start before the simulation object acquires the necessary resources and/or coopts other needed simulation objects that are indicated by the entering arrows at

the top of the activity box. When an activity is completed, some resources may be produced and/or some simulation objects may be released. These are represented by the arrows leading out of the bottom of the activity box.

For a PSO, the process cycle is repeated continuously as represented by the closed loop of activity flows in its ACD. But for a TSO, the process is concluded at the last activity and therefore the activity flow is sequentially.

In the following section, we will describe a logistics support problem that is consequently represented as ACDs.

A Logistics Support Problem

This problem is a modified version of a simulation problem taken from Prisker, Sigal and Hammesfahr, 1989. It is a support process for an hypothetical air transportation system. Airplanes, choppers and a maintenance facility are the main features of this system. Airplanes use sophisticated equipment (boxes) that are necessary for their missions and these boxes have random service lives at the end of which airplanes are forced to end their flights and go under a repair process at the maintenance facility. Choppers that develop troubles frequently also use the facility for emergency service, and can be repaired in a relatively short period of time by a repairman. For the maintenance of airplanes the same repairman diagnoses the problem to a faulty box (discovery period) then he removes the box from the airplane. If a spare box is available the repairman can install the good box onto the airplane right away and the airplane can return to flight when a pilot becomes available. The

faulty box is repaired in the maintenance facility using a sophisticated testing machine and finally delivered to the spares room for the future use.

Many decision parameters exist in this simple logistics support problem that can be studied using a simulation model. Operational readiness of the airplanes and choppers is affected by the performance of the maintenance facility, by the number of spares, the fleet sizes of airplanes and chopers, and many other parameters which relate to how fast certain activities can be accomplished. Modeling of this simulation problem requires first analysis of the problem to define the necessary simulation objects and their processes for the object-oriented simulation. and then representation of stochastic behaviors of these relevent elements into a correct computer program that will simulate the system to a desired level of details. In the following section, the graphical representation scheme (ACD) will be presented.

Representation of the Logistics Support Problem as ACDs

As it is the case with almost all kinds of models, there may be several ways that one can visualize a simulation problem in an ACD representation depending on the problem conditions, goals, and the level of details desired. The desirable form should be the one that represents the problem the most natural way in its physical setting. The first question to be settled in drawing ACD diagrams of a simulation problem is which objects will be defined as simulation objects and which as resources. We can offer some guide-lines to help distinguish between formulation of simulation objects and resources. Admittedly, sometimes these are not as

definitive as to become rules and modeling feat is still mostly an art form that heavily relies on the modeler's imagination, analysis and design ability as well as the conditions and goals of the simulation problem at hand. In general terms, it is important to note the following characteristics associated distinctly with simulation objects and resources. A simulation object is an active participant (actor) that has a certain process to carry out in the simulation scene. On the other hand, a resource is passive (accessory) in that it is needed in certain quantity in execution of some activities, but it is not possible or desirable to give any structural order to these activities in the form of a process. For example, the repairman in the above problem takes place in several activities of several simulation objects in a random order on the basis of the first-come-first-serve principle. Therefore, resources restrict and regulate activities of simulation objects.

In the logistics support problem, one can clearly identify the airplanes as PSOs, the choppers as TSOs, and the repairman, the pilots and the testing machine as different types of resources. But, how are we going to treat the boxes ? They could be treated as a PSO type that is coopted by the airplanes during their flights, discovery and removal activities and then they go under repair and delivery operations individually. In this form, each box could be allowed to have more individualistic behavior that might be useful if there were several types of boxes. Thus, each type of box might, for example, mean a different type of flight mission for the airplanes and possibly a different type of

repair process for the maintenance facility.

As the additional information on the logistics support problem introduced above, we imagine that there is a physical maintenance facility into which airplanes have to be moved in order to diagnose their faulty boxes and due to various reasons such as magnetic disturbances, the maintenance facilitiy is available either for an airplane for diagnosis and removal of the box, or for repair of a faulty box, but not for both simulataneously. Since we are dealing with some generic type of "box" here and not so much concerned with the identity of an individual box, we will simply treat the box object as a type of consumable resource and the maintenance facility as another PSO. As such, one unit of box will be "consummed" in the beginning of the installation activity of the airplane and one unite of box will be "produced" at the end of the delivery activity of the maintenance facility.

The ACDs of the logistics support problem are given in Figure 3. An airplane object repeats its process which is composed of "flying" and "installing" activities during the entire simulation period. Note that for easy identification, the activity names are chosen as the verbs with the "ing" endings to describe the action. For flight of an airplane one pilot is needed as resource. At the end of flight because of a failed box, the airplane releases the pilot and is coopted by the maintenace facility for discovery and removal of the faulty box. These two activities will be carried out in a coopted manner between these two simulation objects and they will need one repairman. After the removal of the faulty box, the airplane ruturns to its own

activity of box installation. When one good box and one repairman are available, the airplane object starts its installation activity and releases the repairman at the end of this activity. If a good box is not immediately available it has to wait for one produced by the delivery activity of the maintenance facility. After the faulty box is removed from the airplane it is repaired in the maintenance facility using the testing machine. Then, "delivering" activity produces one good box and the process of the maintenance facility starts all over again.

Note that the synchronized activities, "discovering and removing" beetween the airplanes and the maintenance facility take place only in one place: in our case, as the part of the maintenance facility ACD. This means that only one airplane can be served at a time and the entire facility will be closed to new airplanes until a good box is produced by the last delivery activity since the maintenance facility will be able to attend to one activity at a time.

Choppers have a very simple single-activity process for which the repairman must be captured. The repairman is released when the "fixing" activity is completed and becomes available for other simulation objects.

The ACDs shown in Figure 3 form the main activity structures of the SOs and their logical relations. Additional necessary information (e.g., about the entrance times of the SOs, the amount of resources available, etc.) is obtained from the modeler in a dialog style during the constructions of these ACDs.

As mentioned earlier, it is extremely easy for the human

beings to understand and analyze simulation problems in the simple ACD formalism. And at the same time, it is also quite straight forward to build ACDs interactively and to have them interpreted into computer programs automatically. Another benefit of ACDs comes at the stage of verification and validation of simulation models when the events can be observed in these same ACD display in an interactive mode.

In ACD representation, the process of a simulation object is composed of a special directed network of activities and each activity is described with a fixed set of attributes: resource needs, coopted simulation objects, duration of the activity, the resources and other simulation objects to be released at the end of the activity, and finally the name of the next activity to be activated. In addition, some information related statistical data collection may be requested. All this information can be requested from the user automatically in a cyclic manner one activity after another. The regularity in definitions of the activities which basically form the activity-cycle diagrams occurs across all problem instances and has been exploited successfully to interprete general types of simulation problems into computer programs.

As defined earlier, a SO is either a permanent or a temporary object in the simulation model. A PSO repeats its process continuously during simulation, and a temporary SO leaves the simulation model when its process is completed. The only additional information that needs to be requested for a TSO is related to the time period spent outside of the model, i.e., the interarrival time. For permanent SOs, the number of same type of

objects to be engaged in the simulation model must also be supplied as a model parameter.

Building Activity-cycle Diagrams Interactively

Here, we will describe an efficient interactive way of building ACDs using our protype program called GraphSIM (Graphical Simulation Modeling) in the Smalltalk-80 environment. Smalltalk is a graphically oriented language which frequently utilizes the three-button mouse for programming. In Smalltalk terminology, the leftmost, middle and rightmost buttons of the mouse are referred to as the red (RB), yellow (YB) and blue (BB) buttons, respectively. When these buttons are clicked in certain "hot" regions (windows) of the screen, the Smalltalk usually answers with pop-up menus and by selecting the choices the user invokes certain operations, such as creating a system browser or executing a selected statement, etc. In the GraphSIM, the same mode of communication is used extensively.

In order to create the GraphSIM environment, one needs press the BB of the mouse in the clean part of the screen and select the choice "run GraphSIM" from the menu of a pop-up window. As a result of this action, the GraphSIM window will be constructed on the terminal screen by the computer as shown in Figure 4. The GraphSIM window is a composite window which is composed of mainly three areas: the top part of button windows is reserved for execution control of an existing simulation model; the middle section is for interactive construction of ACDs and is composed of four subwindows for simulation objects, resources, activities and activity elements, respectively; the bottom section is only

for pictorial display of ACDs. The top and middle sections of the GraphSIM window are hot areas in that they respond to the RB, YB and BB buttons of the mouse. The bottom section is passive in this respect. As any other standard system windows in Smalltalk, the GraphSIM window can be closed, framed , moved or collapsed at any time by the blue-button commends. The input-windows of the GraphSIM respond to the YB by creating different pop-up windows in the different regions, some of these pop-up windows are shown in Figure 5.

The graphical construction of ACDs for a simulation problem in the GraphSIM window is carried out in the follwing steps:

1) When the YB is pressed in an empty or unhighlighted SO-window a pop-up menu presents two choices as to adding a new SO or inspecting the simulation model. The second choice takes the user to the underlying simulation environment in Smalltalk to inspect the instance of the object called SimulationModel and it is therefore for a sophisticated user only. But when the first choice is picked, the user is asked about the name, the entering time of the new SO in a special pop-up, fill-in-the-blanks window. Consequently, the user is asked whether the SO is a permanent or temporary SO. If the user says it is permanent, the number of such objects to be introduced into the simulation model is requested. Otherwise, the interarrival time distribution for the temporary SO will be requested.

This same process of questioning will be repeated automatically for the new simulation objects until the user responds with a blank entry as the name of the SO.

2) Next, when the YB is clicked in the resource window, the user may choose to add a new resource or do nothing as the menu of a pop-up window. For addition of a new resource, the user is questioned about the name and quantity of the resource in two new pop-up, fill-in-the-blanks windows. The name of new resource is entered into the resource window at this point after typing the required information. Information for additional new resources will be requested automatically until the user responds with a blank resource name. After completing resource definitions the GraphSIM is ready to extract informations related to the process definitions of SOs.

3) In the SO-window, a SO name is highlighted when the RB is clicked on a SO objec If the YB is pressed when a SO is highlighted the user is asked whether activity descriptions are desired for the selected SO. This same point of entry can also be reached by selecting the appropriate one among many other choices in a menu that pops up when a SO name is highlighted in the SO window and the YB is pressed in the activity window.

Subsequently, the user will be promted to name the activities desired to be entered for the process of the selected SO automatically using unique names one after another. Again, the recursive nature of questioning will be broken when the user enters a blank response as an activity name. At this point, the user is asked whether further description of each activity entered hence is desired. If the response is positive the communication continues as in the following step.

4) Coming from the previous step directly, or selecting the appropriate entry in the YB pop-up menu of the activity window when the SO name is highlighted, the GraphSIM repeats a sequence of questions requesting information about the attributes of each activity that has been entered as the process of the selected SO.

The first question asked about an activity is related to whether recording a time period starting in the beginning of the activity for the purpose of output data collection is desired: if positive, the user is then asked at the end of which activity of the selected SO this time period will end and what the step size for the histogram representation will be. Next, the user is asked repeatedly whether the activity needs any resources until the response is "no"; if positive, the needed resource is identified from menu of the existing resources known by the GraphSIM and the quantity of the resources to be acquired in the beginning of the activity is requested.

After the resource information, the user is then asked via a pop-up menu whether coopting another SO is necessary for execution of this activity. If coopting is not necessary for the current activity, the user picks up "skip" from the menu. If coopting is the case, the coopted SO is picked up from a menu of the existing SOs known by the GraphSIM. Then, the user is asked to identify the activity of the coopted SO, at the end of which the SO will be coopted. Subsequently, he will be asked to identify the activity of the coopting SO at the end which coopting will end and the coopted SO will return to its own process. One important point in the definition of activities of SOs is that the activities of the coopted SOs have to be defined

before the related coopting SO's. Thus, all the information about the coopting operation is requested at one point, that is, in describing the coopting SOs.

The last piece of information that needs to be supplied is about the probability distribution for the activity duration. The user is asked to pick up the appropriate distribution from a pop-up menu of distributions available in the simulation environment and is then asked to fill in the blanks in a pop-up window for parameters of the selected distribution.

5) After defining the attributes of all the activities of a SO, what is left to be done in completing the process network is to identify the connections among these activities. This also can be thought of another attribute of an activity which points to the next activity to be started upon its completion. Currently, only sequential connections are allowed in the GraphSIM. Probabilistic or state dependent branching at the end of activity is left for the future development of the GraphSIM. For the purpose of defining these sequential connections, if the YB is pressed in the activities window when a SO is highlighted in the SO window, a pop-up menu with many choices appear. The choice "next accivity automatically" picked up from this menu will do the job of connecting all the activities sequentially.

6) For processes of TSOs, we don't need to do anything else since these objects will leave the simulation model upon completing their last activities. However, for the permanent SOs we have to identify to which activity the process will return

after the last activity. For this purpose, after highlighting the SO and its last activity in the SO and activity-windows respectively, the choice "next action" is highlighted in the activity-elements-window. Then, when the YB is pressed in the activity-elements-window and subsequently "add" is selected from the resultant pop-up menu, another menu of all activities of the highlighted SO will be displayed. The appropriate next activity should now be selected from this menu.

Note that now the ACD diagram of the selected SO also appears in the graphical display-window whenever an SO is highlighted in the SO window. Figure 6, 7, and 8 show the GraphSim window for the three SOs of the logistics problem.

7) We can run the simulation model directly from the GraphSIM window. If the RB is clicked on the Start-Simulation window the user is promted for the simulation period, and consequently the simulation starts. When the simulation ends, the hot windows start responding to the mouse actions. The output data will be printed on the System Transcript window when the RB is clicked on the Finish-Up window. The output histogram for airplane flight times for the logistics problem is shown in Figure 9.

VI. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

The GraphSIM meets all of the requirements laid out for a good graphical programming facility in Section 4. The ACD formalism is extremely easy and intuitive for simulation problems and takes little time to learn. Therefore, it also proves to be beneficial as a communication medium for the modeler, user and managers who are all interested in the different aspects of

simulation models. It is clear is that "seeing" and easy comprehension of the simulation models make believers out of everyone involved in the use of the software.

At the present stage of this research, the prototype system developed for the graphical model builder and interpereter clearly demonstrates that automatical programming of simulation problem is possible once a concise high-level conceptual representation of the problem is at hand. However, this prototype system is not complete in the sense that the branching and confluence type of connections among the activites are not allowed. At present, only the sequential type of connections are possible and this limits the modeling power of the GraphSIM. For this purpose, a probabilistic or state dependent branching must be included as another feature of the activity objects. Further, the multiple copies of the PSOs are not currently permi ed even if the request does not create a terminating error. On the other hand, some other type of errors that may result from, for example, trying to edit already existing elements of activities, cannot be handled easily by the novice user. Therefore, the GraphSIM is not considered a robust programming environment, yet. But all these can be remedied easily in piece-meal with additional programming development, or all at once with addition of a graphical programming editor to the GraphSIM. We currently opted for the second option and in the near future, the GraphSIM will be expanded to include the facilities for a graphical programming (model) editor and view builder as described in Section 4.

Another line of research that will be followed is to enrich the modeling capability of the GraphSIM with additional high-level concepts for complex real-life systems that demand more expressiveness on the part of the modeling system. Currently, the GraphSIM is able to handle resource utilization/production and activity synchronization for the simulation objects. This is sufficient for modeling most common simulation situations. But, for the problems that require complex interactions among the simulation objects, other high-level concepts are needed as modeling tools to represent the possible complex situations easily and explicitly.

For example, a simulation object may carry out concurrent activities as a part of its process (i.e., and-branches or concurrence concept); or a team of simulation objects may be needed to undertake a certain type of mission as a group under certain circumsantances and they carry out their usual processes individually at other times (coherence concept); or, disruptions may occur causing a certain simulation object to be permanently or temporarily disabled at any of its activities due to a random or state-dependent events (distruption concept); and eventually, the concept of intelligent simulation objects that are capable of making their own decisions should be developed to deal with abundant decision situations that arise frequently in simulation problems. (e.g., see (Ozden, 1989) for the intelligent simulation objects).

Another research study should be directed toward the computational efficiency of simulation programs in the object-oriented environments. While we were developing the GraphSIM

system, our main concern was not to create an efficient programming environment. So a lot of improvements in object class definitions can still be incorporated into this prototype system to increase the execution speed significantly. But major gains in this respect can be expected from the object-oriented languages that take advantage of the RISC and parallel processing technologies in the future. The earlier research efforts in the literature show a good prospect for the object-oriented languages.

REFERENCES

Bezivin, J., "Timelock: A Concurrent Simulation Technique and its Description in Smalltalk-80,Conf. of 1987 W. Simul.Conf.,1987.

Birtwistle, G.M.,Discrete Event Modelling on Simula, MacMillan, 1979.

Birtwistle, G. M. and P. Luker, "Dialogs for Simulation," Proc. of Conf. on Simul. in S. Typed Lang., San Diego, 1984.

Birtwistle, G. M., B. Wyvill,D. Levinson and R. Neal, "Visualising a Simulation Using Animated Pictures," Proc. of Conf. on Simul. in S. Typed Lang., SanDiego, 1984.

Browne, J. C., J. E. Dutton and D. M. Neuse, "Introduction to Graphical programming of Simulation Models," Proc. of 1986 S. Simul. Comp. Conf., 1986.

Dahl, O. J. and K. Nygaard, "SIMULA: An Algol Based Simulation Language," Comm. of ACM, Vol. 9, 1966.

Duersch, R. R. and M. A. Laymon, "Programming-free Graphic Factory Simulation with GEFMS/PC, "Proc. of 1985 W. Simul. Conf. 1985.

Goldberg, A and D. Robson, Smalltalk-80: The Language and Its Implementation, Addison Wesley, Reading , Mass, 1983.

Hurrion, R. D., "Visual Interactive Modelling," EJOR,Vol.23,1986.

Ingalls, R. G., "Automatic Model Generation," Proc. of 1986 W. Simul. Conf., 1986.

Jefferson, D., "Future Directions in Simulation at the Conference on Simulation in Strongly Typed Languages (Panel)," Proc. of Conf. on Simul. in S. Typed Lang., SanDiego, 1984.

Knapp, V.,"The Smalltalk Simulation Environment," Proc. of 1987 W. Simul. Conf., 1987.

Melamed, B. and R. J. T. Morris, "Visual Simulation: The Performance Analysis Workstation,"Computer, IEEE, Vol.18, 1985.

O'Keefe, R. M., "What is Visual Interactive Simulation ? " Proc. of 1987 W. Simul. Conf., 1987.

Ozden, M., "A Simulation Study of Multi-load-carrying Automated Guided Vehicles in a Flexible Manufacturing System," Int. J. Prod. Res., Vol.26, 1988.

Ozden, M., "Simulation of Intelligent Objects," CORS/TIMS/ORSA meeting at Vancouver, CAN, 1989.

Popken, D. A., "The Productivity Improvements in Simulation Modelling (PRISM) Project: Concepts and Motivations," AFHRL Technical Paper, 88-31, 1988.

Sargent, R. G., "The Use of Graphical Models in Model Validation," Proc. of 1986 W. Simul. Conf., 1986.

Shannon, R. E., "Models and Artificial Intelligence," Proc. of 1987 Winter Simul. Conf. , 1987.

Sinclair, J. B., K. A. Doshi and S. Madala, "Computer Performance Evaluation with GIST: a Tool for Specifying Extended Queueing Network Models," Proc. of 1985 W. Simul. Conf., 1985.

Stairmand, M. C. and W. Kreutzer, "POSE: a Process-oriented Simulation Environment Embedded in Scheme," Simulation, Vol. 50, 1988.

Stanwood, K. L., L. N. Waller and G. C. Marr, "System Iconic Modeling Facility," Proc. of 1986 W. Simul. Conf. 1986.

Thomasma, T. and O. M. Ulgen, "Modeling of a Manufacturing Cell Using a Graphical Simulation System Based on Smalltalk-80, Proc. of 1987 W. Simul. Conf., 1987.

Ulgen, O. M. and T. Thomasma, "Simulation Modelling in an Object-oriented Environment Using Smalltalk-80," Proc. of 1986 W. Simul. Conf., 1986.

Unger, B., G. Birtwistle, J. Cleary, D. Hill, G. Lomow, R Neal, M. Peterson, I. Witten, and B. Wyvill, " JADE: A Simulation and Software Prototyping Environment," Proc. of Conf. on Sim. in S. Typed Lang., SanDiego, 1984.

Unger, B. W., "Simulation Environments & Parallelism," in Simulation Environments of the 1990's (Panel), Proc. of 1987 W. Simul. Conf. , 1987.

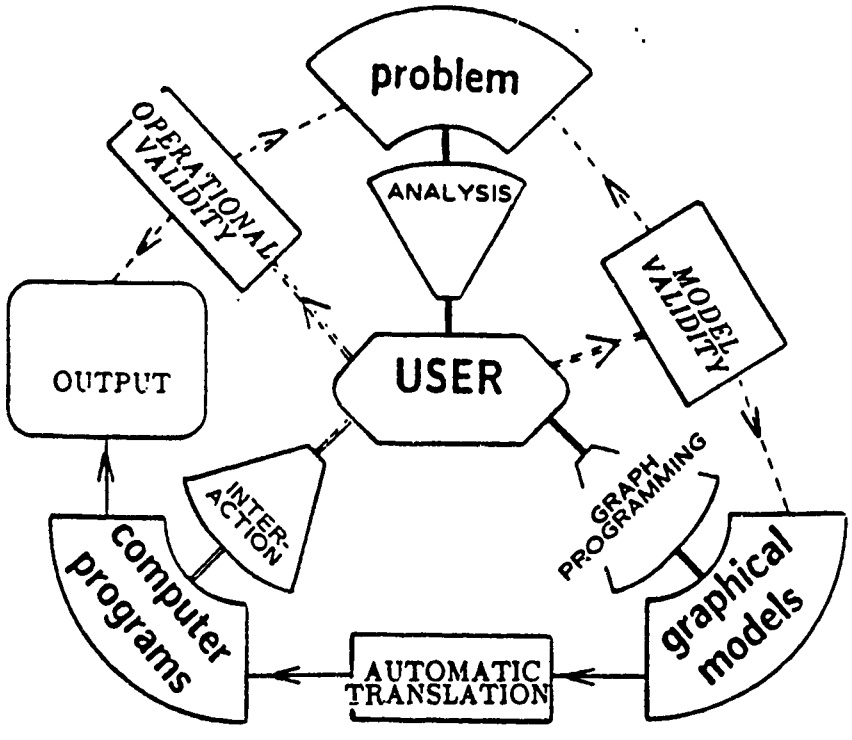


Figure 1: Main elements of visual interactive programming.

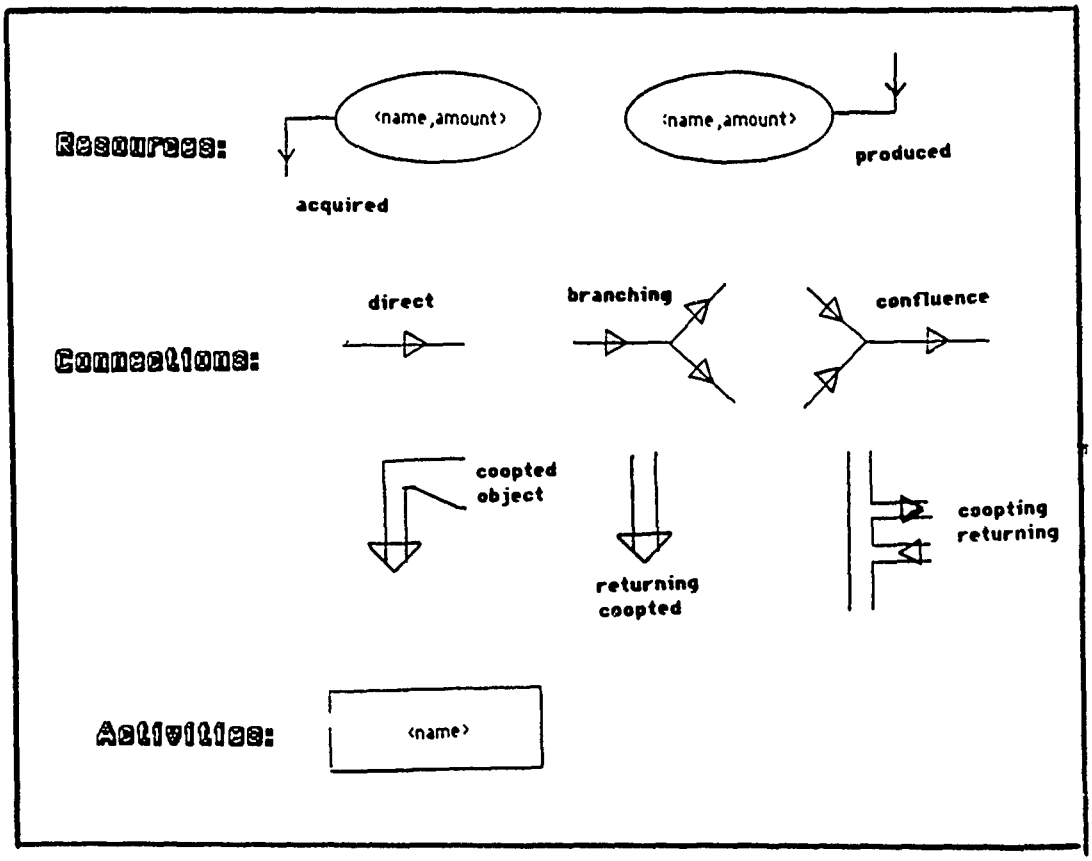


Figure 2: The primary elements of activity-cycle diagram

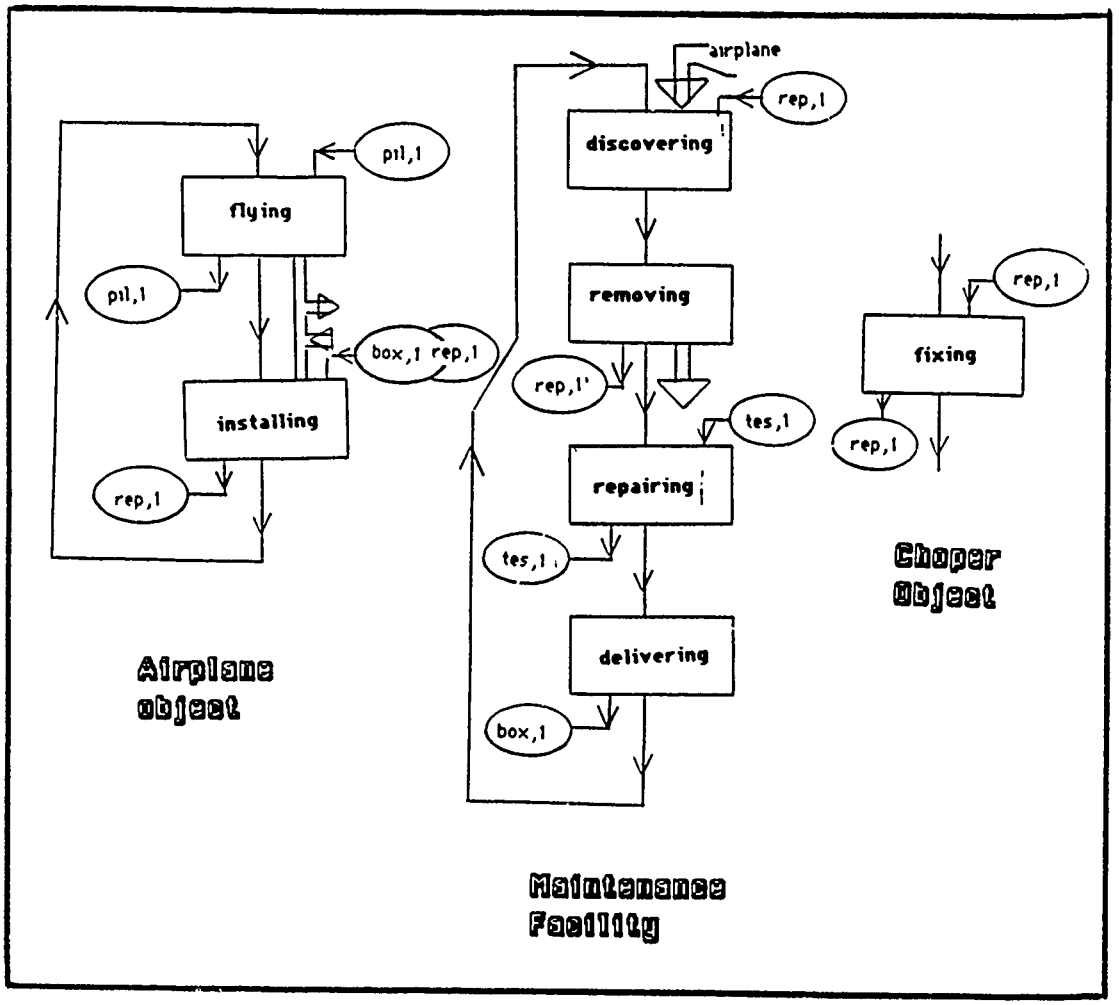


Figure 3 : Activity-cycle diagrams of the logistics problem

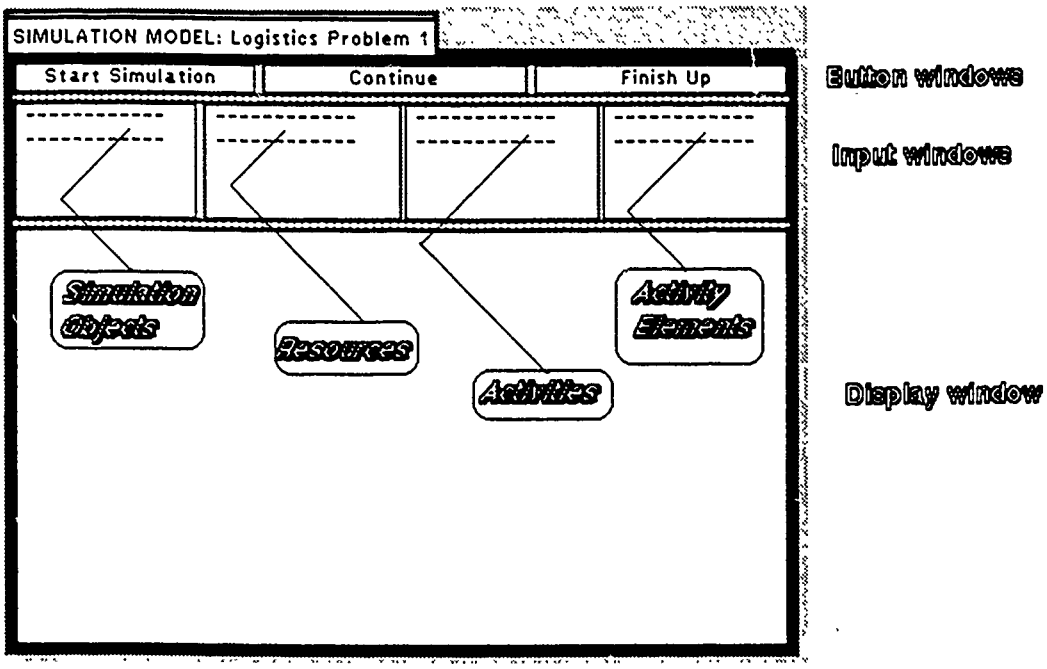


Figure 4 : GraphSIM windows

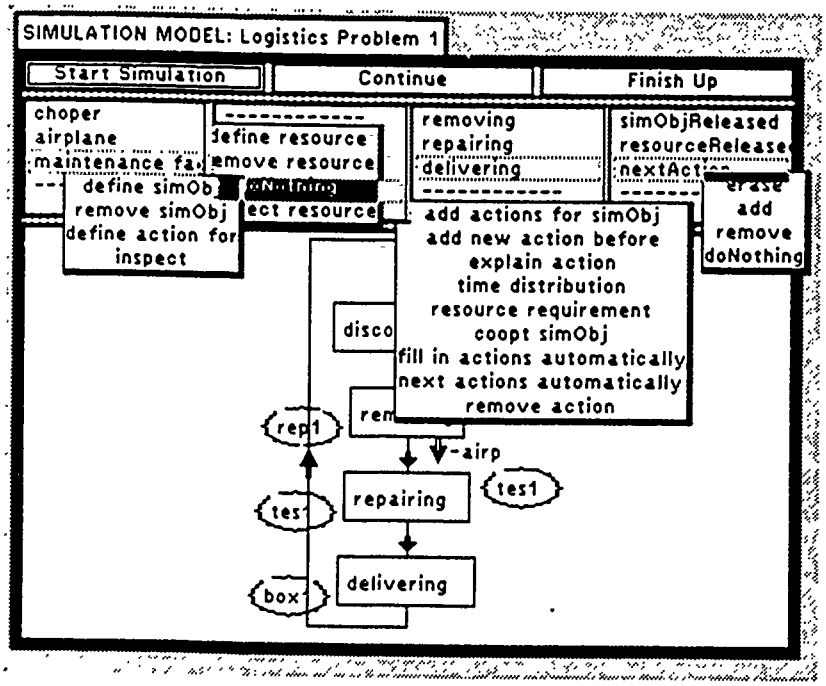


Figure 5 : GraphSim window with its pop-up subwindows

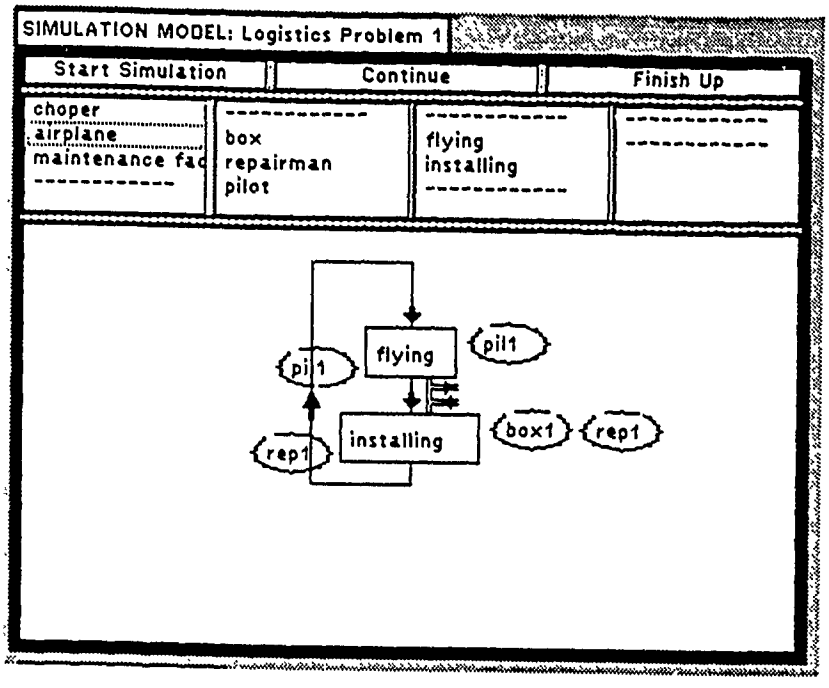


Figure 6 : Airplane Simulation Object as entered in GraphSIM

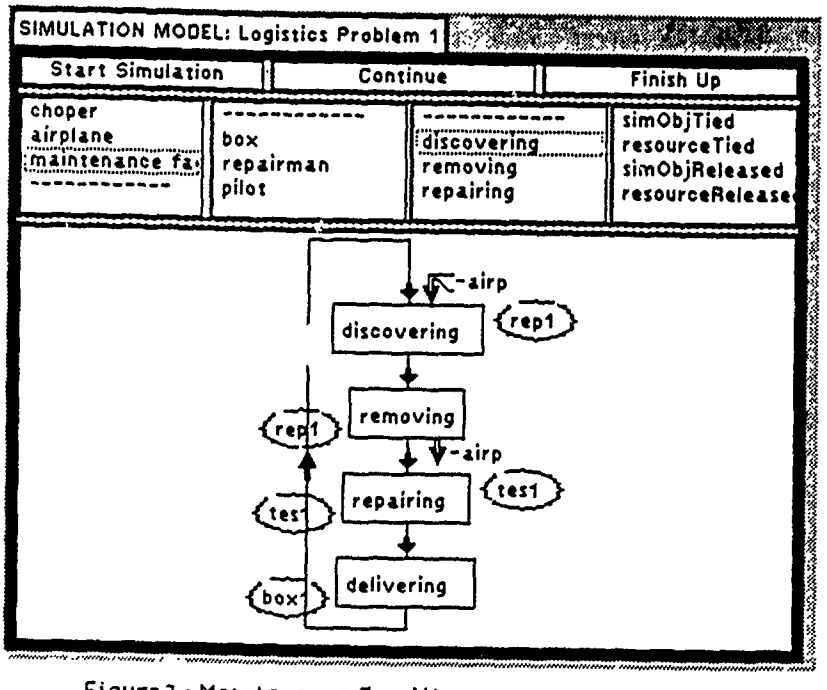


Figure 7 : Maintenance Facility as entered in GraphSIM.

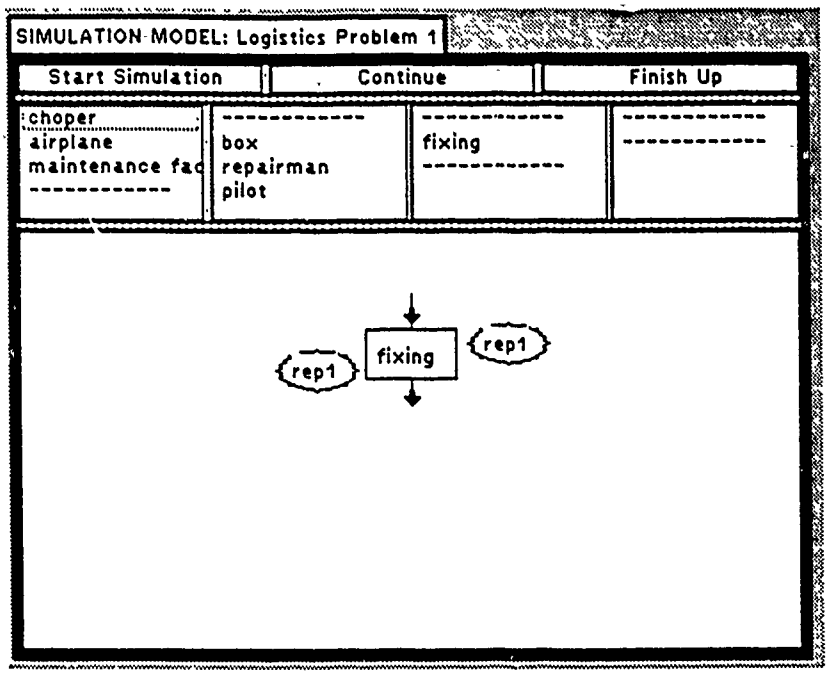


Figure8 ; Choper entered as a temporary simulation object in GraphSIM.

System Transcript

AIRPLANE(FLYING-FLYING)

Sample Size	Minimum Value	Maximum Value	Average Value
54	0.110992	81.9915	9.92394

Entry	Number of Objects	Frequency
0-5	21	0.388889 |
5-10	13	0.240741 |
10-15	9	0.166667 |
15-20	6	0.111111 |
20-25	3	0.0555556 |
25-30	1	0.0185185 |
30-35	0	0.0 |
35-40	0	0.0 |
40-45	0	0.0 |
45-50	0	0.0 |

'Entries outside of the range = 1

Figure 9 : Output histogram on flight times of airplane during 500 time unites of simulation

FINAL REPORT NUMBER 118
REPORT NOT AVAILABLE AT THIS TIME
Dr. Dharam S. Rana
210-9MG-043

1989 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

REFINEMENT CONSIDERATIONS FOR AN ADVANCED INSTRUCTIONAL DESIGN ADVISOR

Prepared by:	Jonathan Michael Spector, Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Math, Computing & Info Sciences Jacksonville State University Jacksonville, AL 36265
Research Location:	AFHRL/IDC Brooks AFB San Antonio, TX 78235
USAF Researcher:	Joseph Scott Newcomb, Ph.D.
Date:	August 22, 1989
Contract Number:	F49620-88-C-0053/SB5881-0378

REFINEMENT CONSIDERATIONS FOR AN ADVANCED INSTRUCTIONAL DESIGN ADVISOR

by

Jonathan Michael Spector

ABSTRACT

The Advanced Instructional Design Advisor (AIDA) is a system currently being designed by AFHRL/IDC to provide automated and intelligent assistance to Air Force instructional designers. AIDA is intended to alleviate the difficulties and expenses of designing effective instructional materials given the complexities of advanced hardware and software technologies, the variety of instructional settings, and the relative naivete of instructional designers. In order to design an effective set of tools, several critical issues need to be resolved, including the identification of useful instructional models and taxonomies of knowledge, a determination of the role for artificial intelligence in AIDA, and an analysis of instructional design requirements. This research effort represents a first attempt to address such critical design issues for AIDA.

ACKNOWLEDGMENTS

I would like to thank the United Stated Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Jacksonville State University was most helpful in providing additional support funds for travel and necessary supplies. Universal Energy Systems was also very helpful in providing administrative support.

I am grateful to the Air Force Human Resources Laboratory for supporting continuing research and development on the Advanced Instructional Design Advisor. I especially appreciate the thoughtful guidance of Dr. Scott Newcomb, AFHRL/IDC Branch Chief, and Dr. Dan Muraida, AIDA Project Manager. Their probing ideas and suggestions have proved most beneficial. In addition, I appreciate the help and support of the other personnel in the Training Technologies Branch (AFHRL/IDC), including Major Bill Wimpee, Mr. Dick Vigue (retired), Mrs. Barbara Eaton, Capt. Bill Dimitroff, Mr. Bill Hawks, and Lt. Teresa Rushano. Mrs. Sherri Barnes provided administrative assistance well above and beyond the call of duty. Members of AFHRL/IDI, especially Lt. Col. Hugh Burns (retired) and Dr. Kurt Steuck, offered many useful ideas. I also appreciate the encouragement offered by Dr. Henk Ruck and Lt. Col. Roger Ballentine.

I. INTRODUCTION

The instructional process involves three major phases: 1) front-end analysis (FEA), 2) design, development, and delivery (DDD), and 3) rear-end analysis (REA). Instructional Systems Design (ISD) has been used by the Air Force to guide progress in all three phases, although its usefulness is somewhat limited since it provides little specific or detailed guidance at any phase. ISD does provide a general guideline of steps to follow in designing and developing instruction, and this general guideline will most probably be retained in some form in whatever advanced systems are developed.

Phase 1, FEA, includes an analysis of instructional requirements and the design requirements for the instructional solutions, including requirements for various courses. Phase 2, DDD, is where the primary course-level and lesson-level planning, development, and implementation occur. The work done in DDD is intended to comply with the instructional requirements of FEA. Phase 3, REA, is basically an analysis of how well the results of DDD met the requirements of FEA. The boundaries between these phases are not always sharply drawn, and the phases do not always occur in a linear or chronological order. Sequences are usually cyclic and somewhat irregular.

These three phases roughly correspond to the six steps in the classical waterfall life-cycle model of software development: analysis, design, implementation, test, and maintenance (Fairley, 1985). As is the case with many software development efforts, the first and third phases (FEA and REA) receive minimal attention in the instructional design process. It is an all too common problem to become absorbed with the details of DDD and to short-circuit both FEA and REA. A rational and balanced approach distributes the effort among all three phases. One of the goals for AIDA is to provide a full range of automated and intelligent tools to assist in all three phases of ISD.

As advances in both hardware and software technologies have occurred, courseware designers have naturally wanted to incorporate a wider range of materials, including graphics, video, and sound, in a much more complex computer environment. As a consequence, many persons with many different talents and skills are involved in the ISD process, especially in the DDD phase. For example, if we imagine an instructional design effort for a computer-based instructional (CBI) system, we immediately see a need for subject matter experts (SMEs), instructional designers, production specialists (e.g. video specialists), and system specialists (i.e. those knowledgeable about the particular computer hardware and software systems being used). As the

number of specialists increases, it is reasonable to expect course development time, costs, and quality to increase. As the complexity of the development and delivery systems increases, we again have similar expectations about development time, costs, and quality. A major challenge in designing AIDA is to design a system that reduces course development time and costs and yet contributes to course quality and effectiveness.

In the summer of 1988, the Air Training Command issued a Manpower, Personnel, and Training Need (MPTN) citing a lack of effective CBI deployment and proposing the development of specifications and guidelines for effective authoring and presentation complete with instructional strategies. AIDA is being designed by AFHRL/IDC in response to this MPTN. The current design effort is being done under AFHRL/IDC Task Order Number 0006.

The preliminary conceptual work was done as part of my Air Force Summer Faculty Research at AFHRL/IDC in 1988 (Contract Number F49620-87-R-0004). The expectation is that a prototype system will be built (possibly in FY '90) according to the design requirements currently being developed.

II. OBJECTIVES OF THE RESEARCH EFFORT

The basic objective of this research effort was to assist AFHRL/IDC in the refinement of the design for AIDA. The original proposal included as specific objectives the following: 1) an evaluation of SOCRATES, a system developed and used at the Air Univeristy's Air Command & Staff College, 2) a description of useful instructional models, 3) a description of an instructional model selection algorithm to match learning objectives, knowledge types, instructional settings, and instructional strategies, 4) a modular description of AIDA, 5) participation in the formal review of the design requirements for AIDA, and 6) an evaluation of other research and development efforts in this area.

Those specific objectives were modified to comply with current efforts in AFHRL/IDC and in accordance with guidance from Dr. Scott Newcomb. During the period of this research effort, AFHRL/IDC's emphasis shifted from the developmental to the theoretical (it has recently shifted back to the developmental). As a consequence, this effort focused on three specific objectives: 1) an analysis of the current state of the science of instructional design, including on-going research and development efforts, 2) an analysis of the role for artificial intelligence in AIDA, and 3) participation in the design requirements kick-off meetings.

III. PROBLEM AREAS EXPLORED

In order to achieve the revised research objectives and to support AFHRL/IDC's efforts to refine the design of AIDA, several different problems were evaluated. In determining the current state of instructional design science, two problem areas were explored: 1) Can theories of knowledge, learning, and instructional design be meaningfully integrated?, and 2) What theories are existing instructional design systems incorporating?

In order to determine the role for artificial intelligence in AIDA, again two problem areas were explored: 1) Is there a role for expert systems in the ISD process?, and 2) Is there a role for neural networks in the ISD process?

In order to provide support for AFHRL/IDC's current efforts under Task Order Number 0006, two additional problems were explored: 1) What kind of consultants should be involved in the design effort?, and 2) What specific tasks should be assigned these consultants?

IV. RESEARCH ACTIVITIES

With regard to the two problems in the area of determining the state of instructional design science, there were four research activities: 1) a review of current and classical

attempts to integrate theories of knowledge, learning, and instruction, 2) participation in the 1989 ITS Conference hosted jointly in San Antonio by AFHRL/IDI and SouthWest Research Institute, 3) participation in the ITS Workshop held as part of the Software Engineering Institute's 1989 summer conference, and 4) participation in the Utah State University Annual Summer Instructional Technology Institute on "Using Artificial Intelligence in Education: Computer Based Tools for Instructional Design."

With regard to the second set of problems in the area of determining the role of artificial intelligence in AIDA, there were two research activities: 1) a review of current literature in the area of artificial intelligence and instructional design, and 2) a presentation of "The Theoretical Limitations of Neurocomputing" at the 1989 Jacksonville State University Neural Networks Lecture Series.

With regard to the third set of problems in the area of supporting AFHRL/IDC organize the kick-off meetings associated with Task Order Number 0006, there were three research activities: 1) pre-planning and coordination with AFHRL/IDC in drafting documents pertinent to the meetings, 2) attendance and participation in the meetings, and 3) presenting an analysis of the meetings to AFHRL/IDC.

V. RESEARCH FINDINGS

The first set of research activities were an attempt to determine the state of instructional design science. Basically the activities involved a review of the relevant literature and participation in a number of workshops and seminars about state-of-the-art instructional design systems.

In reviewing the literature an attempt was made to identify explicit attempts to integrate theories of knowledge, learning, and instruction. The only clearly integrated and synthesized theory was found in the classical literature of Plato's dialogues, especially the Meno and the Theatetus (Jowett, 1871).

Plato's theory of learning is that learning is a process of remembering what the soul has learned in previous incarnations. Instruction then becomes a two step process. First, cause the learner to become aware that he has forgotten something important, such as the meaning of virtue. Then, by a series of analogies and reminders, draw out of the learner what he has forgotten. One can see a certain similarity between these steps as practiced by Socrates and Robert Gagne's nine events of instruction, which include gaining the learner's attention, informing him

of the objective, stimulating recall, etc. (Gagne, 1985).

Plato's theory of knowledge involves two key aspects: 1) knowledge involves certainty, and 2) behavior is a measure of knowledge. Making the correct claims about virtue is not sufficient to establish that one knows what virtue is. One must also behave in such a way that it is obvious that one understands the meaning of virtue, and, of course, Socrates had very high standards. What is most interesting is to see the behavioral component in this early view.

While I feel certain that most modern theorists will reject certain aspects of this integrated theory (e.g. immortal souls or infinite reincarnations), I also feel that the integrated view is important. In the course of elaborating advanced learning theories, we will need to take into account a complete account of a person. It will not be sufficient to examine just the brain or just behavior or just verbal responses. One modern version of such an attempt to provide a holistic and integrated account of learning can be found in Michael Arbib's <u>In Search of the Person: Philosophical Explorations in Cognitive Science</u> (Arbib, 1985).

In examining current instructional design systems and the development of CBI, several noteworthy items were

discovered. First, the history of CBI reveals a bias towards the DDD phase of instructional design. Little emphasis has been given to FEA and REA. Within the DDD phase, the historical trend has been to first emphasize development and delivery and, almost as an afterthought, to emphasize design.

Moreover, advances in technology have made possible the incorporation of sophisticated graphics, video, and sound. These added complexities have aggravated the ease-of-use problems with CBI development systems. The design phase has yet to receive adequate attention in terms of automated tools and techniques (Montague, 1983). There are mouse-driven graphics editors and the like, but there are very few aids designed specifically to advise authors how best to organize and deliver course materials in a computer-based environment.

Roughly stated, what needs to occur with regard to courseware authoring systems is to provide additional tools to support the design phase of DDD and the other two phases (FEA and REA), and to integrate all these tools in an easy to use, intuitive system.

In response to the two questions posed in this area (Can theories be integrated?, and What theories are being used?),

it is possible to provide an integrated account of learning, instruction and knowledge (Plato has done so), and modern theorists typically do not make explicit the connections between and among knowing, learning, and instructing. Those theories currently being used include Gagne's Nine Events of Instruction and Merrill's Component Display Theory.

The second set of research activities involved an exploration of the role for artificial intelligence in instructional design. This exploration followed two paths: 1) neural networks, and 2) expert systems, including intelligent tutors. These are the two major areas of artificial intelligence applications, so this was a natural course to pursue.

While there is a great deal of current activity in the area of neural networks, there does not appear to be any direct or immediate application in the area of instructional design. Neural nets are basically learning automata, insprired by the workings of the human brain. Areas of application include pattern recognition, speech recognition, and robotics (Soucek, 1988).

The story with regard to expert systems is different, however. Expert systems have become commerically successful within the last ten years in a variety of domains. It is

possible to divide expert systems into two categories: 1) Diagnostic systems that proceed from symptoms backwards toward a relatively small set of causes, and 2) Planning systems that proceed forward from a given set of restrictions to a projected set of outcomes. In both types of systems there is a similar overall architecture involving a set of rules, a set of facts, and an inference mechanism (Luger, 1989).

Early expert systems were written in LISP or PROLOG or another high level computer language. The challenge for computer science was to design effective inference mechanisms. As algorithms for the match and apply phase of the inference engines became more complex, higher level tools were developed. For example, the RETE pattern matching algorithm developed by Charles Forgy at Carnegie Mellon University to match facts with the IF-components of rules was incorporated in the OPS series of production system development tools (Giarratano, 1989). There now exist a number of expert system shell development languages which hide the complexities of the inference mechanisms from the user. The idea behind these tools is to make the technology of expert systems available to domain experts and require minimal knowledge of artificial intelligence and programming. Examples of such systems include ART, EXSYS, GURU, KEE, and Personal Consultant (Firebaugh, 1988).

What should occur with courseware authoring systems is exactly what has been happening with regard to expert system development tools. That is to say, easy to use tools for domain experts need to be developed.

In addition, incorporating expert systems into the instructional design process appears both possible and worthwhile, especially in a typical military setting where the instruction designer has had little formal training in the instructional design process. Diagnositic expert systems have been successfully incorporated in the form of intelligent tutors, which analyze student responses and adjust courseware delivery in accordance with a diagnosis of a student's (mis)understanding based on comparison with a model of an expert's understanding. Intelligent tutoring systems have been shown to be effective in specific and restricted domains, but they are expensive to develop and to implement (Wenger, 1987).

Another challenge is to develop expert system planners to aid courseware designers develop consistently effective course materials in a cost-effective and systematic manner for a variety of subject matter domains and knowledge types (e.g. declarative, procedural, and causal). There are a few intelligent systems to aid course authors design course materials for declarative or factual knowledge types.

For example, the Air University's Air Command and Staff College at Maxwell AFB has developed a system called SOCRATES, which is intended to assist subject matter experts organize course material into effective lesson materials. SOCRATES is based on Robert Gagne's nine events of instruction and David Merrill's Component Display Theory (Doucet, 1988). The Army Research Institute has been funding a system developed by David Merrill at Utah State University called ID Expert, which is an expert consultation system for the design and development of instruction. Kent Gustafson at the University of Georgia has developed a system for Apple Training Systems called IDioM, which is intended to provide support for the entire instructional design process. IDioM is written in HyperCard and is probably the friendliest and easiest to use system of those mentioned. It is also one of the very few which acknowledges the FEA phase of instructional design. IDioM's intelligence is primarily restricted to the selection of appropriate templates for instruction based on input parameters, but this is one useful form of machine intelligence, and it does automate the ISD process.

There are a few other systems under development, such as the Alberta Research Council's Expert CML and the Navy's Authoring Instructional Materials (AIM). The effectiveness

of all of these systems has yet to be fully evaluated. Furthermore, no system exists as yet to advise course authors how best to design materials that involve multiple knowledge types for a variety of subject matter domains and instructional settings. Part of the challenge for AIDA is to accomplish this and also to provide a full range of automated tools to support all three phases of the instructional design process in an easy to use, flexible, intuitive system.

One of the most fundamental issues yet to be resolved is how to represent (if it possible) the knowledge of experts, specifically instructional design experts (Winograd, 1986). Because expert systems are intended to emulate human experts in the sense of producing results consistently similar to those produced by human experts, it will be necessary to identify human experts in the area of instructional design and to measure the effectivenss of automated design tools against the effectivenss of those experts.

My answers to the two questions about the role for aritifical intelligence are obvious: 1) There is a role for expert planning systems in the instructional design process, and possibly a role for an intelligent tutor to teach instructional design, and 2) There is no immediate role for neural networks in the instructional design process.

The last set of research activities involved supporting AFHRL/IDC plan and coordinate the kick-off meetings for Task Order Number 0006. For a complete analysis of the results of this process see Appendix A.

VI. RECOMMENDATIONS

My basic recommendation to AFHRL/IDC is that it is definitely worth continuing with the design and development of a prototype AIDA system. In addition, I agree with the analysis of the consultants recruited by Mei Associates for Task Order Number 0006 that it is now time to shift the emphasis back to the less theoretical and more practical design and development issues. The need for a guiding and coherent theory is obvious, but the design and development of AIDA need not await the validation of such a theory. Indeed, AIDA is intended partly as a tool to use in refining and integrating theories of knowledge, learning, and instruction.

I also recommend that AFHRL/IDC not abandon attempts to make AIDA an intelligent system. There appears to be no immediate role for neural networks in AIDA. However, there are potential roles for a number of expert system planners throughout the ISD process.

With regard to prototyping (not immediately pressing, but soon to be of concern), I recommend that consideration be given to developing the prototype using an object-oriented design/hyper-media development system such as HyperCard on the Apple MacIntosh or NextStep on the NeXT computer system. The prototype system need not use the same hardware or software as the final target system. The critical choices for a prototype development system are cost and ease of development effort. The point of the prototype is to illustrate the functionality of the proposed system in an effort to prove the concept is workable. It should be noted that adoption of this prototyping recommendation involves abandoning the Instructional Support System (ISS) as a test-bed delivery system for the prototype.

APPENDIX A: AIDA PROGRESS REVIEW

DATE: 10 AUGUST 1989

CONTRACT: F33615-99-C-0003 (SBA 18810059)

TASK NUMBER: TASK ORDER 0006

TASK TITLE: SPECIFICATIONS FOR AN ADVANCED INSTRUCTIONAL DESIGN ADVISOR (AIDA) FOR CBT

1.0 INTRODUCTION

The purpose of this report is to review progress to date on the Advanced Instructional Design Advisor (AIDA) per Task Order 0006. The scope of Task Order 0006 is to refine and document the concept and functional specifications for AIDA as the basis for a subsequent prototyping task. The opinions expressed herein are solely those of the author; while every effort has been made to be accurate in reporting the substance of the meetings, it was necessary on occasion to draw conclusions and make inferences. Please advise the author of any errors, oversights, or confusions, so that appropriate corrections can be made in subsequent reports.

The primary contractor for this task is Mei Associates, Inc. with Dr. Albert Hickey serving as project manager. Mei Associates has identified and recruited the consultants who will perform the primary research associated with the task.

The general approach proposed by Mei Associates is to assign selected consultants specific research objectives and to bring the group together periodically to review, refine, and integrate the findings.

The consultants recruited by Mei Associates for this task are Dr. Robert Gagne, Dr. Henry Halff, Dr. David Merrill, Dr. Harold O'Neil, Dr. Martha Polson, Dr. Charles Reigeluth, and Dr. Robert Tennyson. Drs. O'Neil and Reigeluth will serve as reviewers and critiquers, while the other consultants will be responsible for analyzing the results of research that is believed relevant to the design and development of AIDA.

The kick-off meeting was divided into three parts due to prior commitments and conflicts in schedules of the participants. This report is a review and analysis of the three part kick-off meeting.

2.0 KICK-OFF (PART I)

Part I of the kick-off meeting was held at the Mei Associates Corporate Offices in Waltham, MA on 6 July 1989. In attendance were Dr. Peng-Siu Mei (President of Mei Associates), Dr. Albert Hickey (Mei Associates Project Manager), Mr. Leonid Altschul (Mei Associates Software Engineer), Dr. Robert Gagne (Consultant), Dr. Scott Newcomb (AFHRL/IDC Branch Chief), Dr. Daniel Muraida (AFHRL/IDC AIDA

Project Manager), and Dr. Michael Spector (AFHRL/IDC Visiting Research Associate).

The meeting focused primarily on three activities: 1) reviewing the history and status of the AIDA project, 2) reviewing the statement of work developed by Mei Associates, and 3) developing an agenda for Part II of the kick-off meeting. AFHRL/IDC was primarily responsible for the first activity, which was basically a review of the project since its inception during the summer of 1988 with an emphasis on its current direction and status.

The second activity was a detailed review of the statement of work associated with this task. Several modifications were made to the initial statement of work. It was decided that there was a need for only two reviewers/critiquers; Drs. O'Neil and Reigeluth were selected to serve in this role. It was also decided that this task was best divided into two parallel cycles. Each cycle would involve specific research/review assignments. Each cycle would consist of two meetings: 1) a design meeting to discuss the current status and to make specific assignments, 2) and a review meeting to discuss findings and to revise the status. It was also agreed that prior to each meeting all attendees should receive all new material that had been developed. Dr. Hickey agreed to coordinate the mailing of appropriate materials. Mei Associates agreed that the the statement of

work modifications did not substantially change the scope of the task and would not affect the project's budget.

The third activity was to develop an agenda for Part II of the kick-off meeting. It was readily apparent that those consultants attending the next meeting would also require a review of the project's history and status as well as a review of the revised statement of work. Dr. Gagne suggested that what was needed to make the first cycle effective was specific research assignments for the consultants. Dr. Gagne proposed three such assignments and suggested that each consultant should be assigned two so as to insure depth and comprehensiveness of coverage. The suggestion was to have those consultants in attendance at Part II modify the tentative assignments as they deemed appropriate.

Dr. Gagne's tentative task assignments were keyed to an AFHRL/IDC diagram suggesting relationships in the areas of epistemology, theories of learning, and theories of instruction. AFHRL/IDC was emphasizing theoretical foundations for the AIDA project. The reasons for this emphasis were twofold: 1) the project was being funded with 6.2 exploratory development funds, and 2) many researchers had noted a lack of an integrated theory to guide the development of instructional design theory.

Dr. Gagne's three proposed tasks have been detailed in Appendix B to Mei Associates' Project AIDA Report: Part II of the Kick-off Meeting. The first task focused on major findings in cognitive learning research that would have definite consequences for instructional design. The second task focused on analyzing the contents of cognition from the perspectives of both cognitive psychology and epistemology, again trying to formulate clear implications for instructional design. The third task focused on analyzing current principles of instructional theory in an effort to develop common ideas and terminology.

Part I of the kick-off meeting was adjourned with all parties in attendance happy with the minor changes to the statement of work, satisfied with the emphasis on theoretical foundations, and pleased with the specific assignments proposed for Part II of the kick-off meeting.

3.0 KICK-OFF (PART II)

Part II of the kick-off meeting was held at the Air Force Human Resources Laboratory at Brooks AFB, TX on 18 - 19 July 1989. In attendance were Dr. Scott Newcomb (AFHRL/IDC), Dr. Dan Muraida (AFHRL/IDC), Dr. Michael Spector (AFHRL/IDC), Dr. Barbara Sorenson (AFHRL/MO), Dr. Albert Hickey (Mei Associates), Dr. Henry Halff (Consultant), Dr. David Merrill (Consultant), Dr. Robert Tennyson (Consultant), and Dr. John Ellis (NPRDC guest observer).

This two day meeting consisted primarily of three activities: 1) a review of the history and status of AIDA, 2) a review and analysis of the modified statement of work associated with the project, and 3) a thorough discussion and modification of the specific assignments for cycle 1. The first activity was the responsibility of the AFHRL/IDC personnel. A presentation similar to that made at the previous meeting was followed by many probing questions from the participating consultants. The questions basically centered around two concerns. First, what specific details about the AIDA concept could AFHRL/IDC provide? Second, was the primary focus on theoretical foundations wise and justified at this stage of AIDA?

AFHRL was initially reluctant to provide additional details about AIDA because it was then conceived to be part of the consultants' task to provide this detailed theoretically-based conception of AIDA. Dr. Merrill insisted on knowing how AFHRL was thinking about AIDA, and Dr. Newcomb did provide a broader and fuller view of AIDA, including a review of the long-range vision of an integrated set of intelligent, computer-based tools to assist throughout the ISD process. Dr. Spector also made available a complete copy of his original white paper on AIDA.

Dr. Merrill thought that what was needed was not a re-examination of the related theories, but specific details about the nature, functionality, and assumptions of AIDA that could serve to guide the development of a prototype. In short, Dr. Merrill was looking for a requirements analysis and functional specifications to guide a software engineering process. Specifically, Dr. Merrill emphasized the need to perform a knowledge engineering analysis and address issues of knowledge representation in such a complex system.

Dr. Halff also believed the orientation should be more pragmatic and oriented toward a prototype. Dr. Halff suggested a "Wizard of Oz" paper prototype might serve to help elaborate the conception of AIDA and identify potential areas of difficulty. More specifically, Dr. Halff believed that there would turn out to be very few useful general principles of instructional design that could be incorporated into AIDA. He argued that the specific subject matter, student characteristics, and environment would dictate most of the significant components of the system.

Dr. Tennyson suggested that there was a need for an underlying philosophy and set of assumptions to guide the design and development of AIDA. Dr. Tennyson felt it important to make this framework explicit before continuing. He also felt a prototyping task was a reasonable goal to

guide the work, and an important outcome of this effort would be the identification of priority cycles for a series of evolving prototypes as AIDA became more fully elaborated in later years.

In summary, the consultants generally felt that the emphasis should shift to pragmatic concerns such as the nature of the knowledge bases needed for an intelligent instructional design advisor, the minimal set of modules/functions to be included in a first prototype, and an explicit statement of assumptions and principles. While the consultants recognized the important role for an integrated instructional design theory, they thought the theory development and refinement would evolve from the process of building, testing, and refining prototypes.

With regard to the second activity, reviewing the revised statement of work, the consultants thought the general scheme of two parallel cycles was acceptable. The goal of cycle 1 would be to clarify the overall concept of AIDA, identify underlying theoretical assumptions, and establish a suitable framework for elaboration of the AIDA conceptual scheme; this would be a first cut at the functional specifications and preliminary design. The goal of cycle 2 would be to make the design specifications more detailed and to select and prioritize functions for prototyping. The results of the second cycle should be specific enough to

guide prototype development. In addition, those present felt that the two critiquers should serve a dual role of synthesizing the findings and reports of the five consultants and identifying weak areas and oversights. An essential part of the synthesis would be to develop a common terminology. It was again stressed that distributing materials in advance of meetings would facilitate discussion at the meetings.

The third activity involved a revision of Dr. Gagne's suggested assignments. The revision reflected the shift in emphasis from theory development to design of a prototype to aid theoretical development and also aid the practical development of useful automated tools for instructional design. Dr. Halff's revised assignment is to provide a concept of how AIDA would support the development of two broad instructional paradigms, identifying the relevant principles of learning and instruction. Dr. Merrill's task is to identify the general concept of AIDA by indicating the functions that AIDA should perform, including knowledge acquisition and strategy analysis; Dr. Merrill will also identify the underlying principles of epistemology, learning theory, and instructional theory. Dr. Tennyson's assignment is to update the ISD model based on advances in cognitive science and educational technology. It was suggested that Dr. Gagne might respond to his first two tasks, as originally planned. In any event, Drs. Gagne's and Polson's

tasks will be clarified at the next meeting.

It should be obvious that this second kick-off meeting involved some serious re-thinking of the project's primary focus and orientation. While the importance of establishing theoretical foundations was acknowledged, it was generally felt that the specification of functions should not be forced to await a well developed and integrated theory of instructional design. As a consequence, there was a return to the emphasis on developing functional specifications that could guide development of a prototype as stated in Mei Associates original statement of work and in Dr. Spector's original white paper.

Many useful conclusions were reached at this meeting. A variety of practical issues were raised, ranging from identifying and involving users to specifying the details of the AIDA EXEC core. Everyone seems to be in agreement that the design of a prototype is a desirable goal and that this project should be laying the detailed groundwork for that effort.

The nature of the prototype to be built following this project may be worth considering. Dr. Halff suggested developing a paper prototype ("Wizard of Oz" technique) and then role play through the system guided by the principles committed to paper. The advantage of this approach is

cost. Dr. Merrill suggested that we consider adopting an object-oriented design approach and to develop a prototype using a computer-based hyper-media, such as HyperCard stacks on an Apple MacIntosh. The advantages of this approach are fidelity to an actual design environment, ease of rapid prototyping, and the effectiveness of presenting an actual working model. Since the development of a prototype is beyond the scope of this project, this issue need not be resolved now.

There may be significant disagreement about two essential matters, however. First, some participants may be viewing this project as a software engineering task involving the development of a design specification for a product (a prototype) based on a requirements analysis statement, while some may view the development of a prototype as a practical way to elaborate and evaluate an integrated theory of instructional design. AFHRL/IDC does not have a specific requirements analysis document and does not have committed users. AFHRL/IDC does want to involve potential users and eventually design/develop a useful set of tools to make instructional design more cost effective; but there is still a definite interest in developing an integrated theory and not merely adopting a design philosophy. Because this project falls into the arena of exploratory development, the contractor and consultants are encouraged to be creative and

inventive; as yet, there is no specific product to be built.

Second, some participants may believe that many of the important principles in all three areas (knowledge, learning, instruction) have been discovered; what is then needed is a common terminology and flexible knowledge base to accomodate what is already known about instructional design. On this view, it is worth considering who is on the leading edge of the curve in the area of intelligent instructional design and to consider adopting the principles and ideas of those leaders. AFHRL/IDC does not have a comprehensive list of acknowledged leaders, although in recommending potential consultants to Mei Associates, AFHRL/IDC obviously was aware of important developers and designers in this area. AFHRL/IDC is more likely to take the view that this cycle of the project is an attempt to specify the geometry of the curve of advanced instructional design. AFHRL/IDC does not now place itself on whatever curve is sketched, although the hope is that an analysis of current efforts and trends will help locate AIDA on next year's leading edge of instructional design. The significance of evaluating the current state of instructional science is to build on the best knowledge now available, not to be confined by current biases and practices.

4.0 KICK-OFF (PART III)

Part III of the kick-off meeting was held at the Air Force Human Resources Laboratory at Brooks AFB, TX on 8 - 9 August 1989. In attendance were Dr. Scott Newcomb (AFHRL/IDC), Major Bill Wimpee (AFHRL/IDC), Dr. Dan Muraida (AFHRL/IDC), Lt. Teresa Rushano (AFHRL/IDC), Dr. Michael Spector (AFHRL/IDC), Dr. Albert Hickey (Mei Associates), Dr. Robert Gagne, (Consultant), and Dr. Martha Polson (Consultant).

This last part of the kick-off meeting consisted primarily of two activities: 1) reviewing the results of the previous sessions, and 2) assigning cycle 1 tasks to Drs. Gagne and Polson. The meeting began with a short review of the history and status of AIDA to bring Dr. Polson up to date on the background for the project.

In reviewing the results of the previous meetings, several concerns emerged. Dr. Gagne was concerned that the consultants might not produce any results that could be synthesized since there had been significant departure from the original three tasks. Dr. Polson suggested that all consultants needed to respond to some common task to guard against this possibility. Since the other 3 consultants had already adopted their tasks, it was decided that it was too late to change their assignments. Those present did feel, however, that activities could be kept on course by asking all consultants to make a list identifying the assumptions

and underlying cognitive principles relevant to their assignments. Dr. Hickey agreed to send a letter to each consultant requesting that this list of assumptions and principles be provided with the task paper.

Dr. Gagne wanted AFHRL to roughly identify the product for this task. Dr. Newcomb indicated that the product was to be a description of the content of a potential software prototype of AIDA. The essential point was that this task did not involve any actual software production, although the description should be both specific and functional enough to guide a future software development.

Dr. Polson indicated there were at least 3 different schemes for this endeavor: 1) a tutoring system, 2) an advising system, and 3) a critiquing system. Dr. Newcomb responded that the emphasis would probably fall on the second scheme. Dr. Spector indicated that the 3 schemes were not mutually exclusive and that aspects of the other 2 schemes could possibly be included.

In discussing how AFHRL/IDC was thinking about AIDA, it became clear that Drs. Muraida and Spector thought that AIDA would incorporate some aspects of intelligent systems (such as specialized expert systems for particular aspects of instructional design), but Dr. Newcomb felt that AIDA might not be an intelligent system, but more of a sophisticated

tool to automate the instructional design process. Dr. Polson thought that AFHRL/IDC did not have the staff expertise to become involved with intelligent systems. Dr. Muraida indicated, however, that there was some existing expertise in this area (Dr. Spector has taught graduate courses in artificial intelligence and built small expert systems using both EXSYS and PROLOG) and there was a commitment to acquire additional expertise.

With regard to the theoretical concern for an integrated theory to guide the development of AIDA, Dr. Polson claimed that theories of knowledge, learning, and instruction could not be integrated, but that it would be possible to draw from an agreed base of principles. Dr. Gagne felt that we should not abandon theoretical concerns and suggested that we all re-read the 4th edition of Conditions of Learning.

Dr. Polson suggested that what was missing from the discussion was principles for the use of graphics and illustrations, a description of how artificial intelligence would be used, and a set of requirements for ease of use. All present agreed that these concerns would indeed be important to cycle 2 of this task.

The second major activity involved task assignments for Drs. Gagne and Polson. Dr. Gagne was concerned that no one had taken on his original task 3, which focused specifically on principles of instructional theory; he accepted this task as

his assignment for cycle 1. Dr. Polson felt most competent to tackle Dr. Gagne's original task 2, which focused on cognitive theory; she accepted this task with minor modifications to include both forms and processes of cognition and de-emphasize the contributions of epistemologists.

A suggestion was made at the previous meeting that AFHRL should be duplicating the efforts of the consultants. Dr. Merrill had also suggested that AFHRL should draft 3 documents: 1) a needs assessment, 2) a requirements analysis, and 3) a detailed system specification. Dr. Spector thought it was too early to attempt to do the third. Drs. Muraida and Spector agreed to tackle the first 2 documents as their tasks for cycle 1, so long as it was understood that their efforts would be somewhat hypothetical and speculative, since there is not a customer in any traditional sense.

5.0 SUMMARY

By way of summary, there are at least 3 significant issues that emerged from this kick-off series of meetings: 1) AFHRL/IDC needs to provide the group with a more detailed vision of AIDA, 2) the group needs to form a more definite view of how intelligent AIDA should be, including the forms this intelligence should take, and 3) the group needs to

develop a consensus with regard to underlying assumptions and principles in the areas of learning and instruction. Associated with this last issue is the need to develop a common terminology before proceeding with cycle 2. Since there may be a divergence of AFHRL/IDC visions of AIDA, the consultants will need to respond to what they believe most achievable in a follow-on prototyping task.

There is certainly a wealth of expertise represented in the group gathered together for this task. The consultants have already presented many provocative ideas and asked many probing questions. Task assignments for cycle 1 have been clarified and promise to produce results that can be built upon in designing AIDA. We all look forward to the cycle 1 review meeting in October.

P.S. There is an Octoberfest celebration in Texas!

REFERENCES

Arbib, M. A. In Search of the Person: Philosophical Explorations in Cognitive Science. Amherst, MA: The University of Massachusetts Press, 1985.

Doucet, R. C. & Ranker, R. A. "SOCRATES: An Instructional Design Advisor Program." Proceedings: 1988 Conference on Technology in Training and Education. Biloxi, MS: Air Training Command, 1988.

Fairley, R. Software Engineering Concepts. New York: McGraw-Hill Book Company, 1985.

Firebaugh, M. W. Artificial Intelligence: A Knowledge-Based Approach. Boston: Boyd & Fraser Publishing Company, 1988.

Gagne, R. M. The Conditions of Learning and Theory of Instruction. New York: Holt, Rinehart and Winston, 1985.

Giarratano, J. & Riley, G. Expert Systems: Principles and Programming. Boston: PWS Kent Publishing Company, 1989.

Jowett, B. (tr). The Dialogues of Plato. Oxford: The Clarendon Press, 1871.

Luger, G. F. & Stubblefield, W. A. Artificial Intelligence and the Design of Expert Systems. Redwood City, CA: Benjamin Cummins Publishing Company, 1989.

Montague, W. E.; Wulfeck, W. H.; & Ellis, J. A. "Quality CBI Depends on Quality Instructional Design and Quality Implementation." Journal of Computer-Based Instruction 10.3 (1983): 90 - 93.

Soucek, B. & Soucek M. Neural and Massively Parallel Computers: The Sixth Generation. New York: John Wiley & Sons, 1988.

Wenger, E. Artificial Intelligence and Intelligent Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge. Los Altos, CA: Morgan Kaufmann Publishers, 1987.

Winograd, T. & Flores, F. Understanding Computers and Cognition: A New Foundation for Design. Norwood, NJ: Ablex Publishing, 1986.

ENGINEERING DESIGN WITH DECISION SUPPORT:

AN `ATION OF GOAL DECOMPOSITION

Charles E. Wells

University of Dayton

Research sponsored by
Air Force Office of Scientific Research,
Bolling AFB, DC.

ABSTRACT

The engineering design problem of large systems is characterized by a hierarchial decomposition because of the complexity of the design process. This paper examines a modeling technique called *Goal Decomposition* that accommodates a hierarchial decomposition, and is therefore useful in the design of large systems. Two applications of goal decomposition are included to investigate the strengths and weaknesses of this technique.

ENGINEERING DESIGN WITH DECISION SUPPORT: AN APPLICATION OF GOAL DECOMPOSITION

1. INTRODUCTION

Mathematical modeling is becoming a vital tool in engineering design [12]. However, for the design of large systems, these models quickly become too large and cumbersome to be of significant value. As a result, research has progressed into the decomposition of the design problem with the intent of breaking the main design into a series of smaller design problems that can be solved and then recombined in the hope of generating a good global design.

Figure 1 illustrates a network model of the design decomposition problem. The node at the top level (level 0) of the network represents the system which is to be designed. As we pass downward through the network, the parts of the system are exploded into greater detail, until we reach the bottom of the network where no further decomposition is possible.

The network shown in Figure 1 can be thought of as actually containing two different types of optimization problems: (1) A high level optimization, and (2) Several lower level optimizations. The high level optimization is equivalent to answering the question, "What type of system do we want and what do we want it to be able to do?" This optimization problem is typically characterized by multiple objectives. An example application of this type of optimization can be found in [5]. In the case that several designs have been already established or

"frozen," Pareto optimal designs can be determined by the method given in [11] to reduce to list of candidate designs to some managable number. The design decisions that are made at this level are then passed downward to the next level (level 1) in the form of system specifications.

At the middle and lowest levels in the network, the objective is not to optimize the design of the subcomponents over some set of objectives, but instead to meet some stated level of attributes that has been provided by a parent optimization. In some cases it may be suboptimal for a subcomponent to over-perform as well as under-perform. For example, in designing an aircraft, we would like to achieve a specified level of speed (an attribute measuring performance); however, optimizing the wing design for speed would clearly be suboptimal if the landing gear of the aircraft is not designed to accommodate the greater speeds.

Of course, other objectives besides performance may be of interest. Reliability, cost, supportability and producibility are some objectives that frequently must be addressed in designing systems and subsystems. In each case, a stated level of achievement is what is required at the subsystem level.

The need for communication between levels in the hierarchy of Figure 1 is currently satisfied by multiple design iterations in real world design processes. In fact, it must be emphasized that a design of a large, complex system is characterized by many design changes through iterations. As a result, decision support for design decomposition should accommodate these design

alterations brought about by new information passed between levels of the network.

1.1 The Rationale Behind Decision Support

Decision support for the design process has several appealing functions besides potentially improving the ultimate system design. A decision support system (DSS) should provide greater managerial control of the design process. For example, a DSS should be a mechanism for coordinating subsystem designs. This would allow the designer to recognize when a particular subsystem design is insufficient to meet system needs, and what the best alternative designs are. This feature would generate critical information early in the planning phase, and therefore direct management's attention to the important design issues.

In addition, a decision support system for design would allow the designer to assess the degree of "criterion build-up" that results from the inability to meet system specifications. For example, using a support system would allow the designer to determine the loss of system performance due to not meeting a specific system specification.

The purpose of this paper is to describe a goal decomposition approach that can be useful in a DSS for the planning phase of engineering design and to examine two example applications. In the following section we describe some of the approaches that currently exist for decision support of the design process. In section 3, a goal decomposition method is examined and related to the network in Figure 1. In addition, some refinements to the goal decomposition algorithm are

suggested. Section 4 contains two example applications. Section 5 contains the concluding remarks.

2. DECISION SUPPORT FOR THE DESIGN PROCESS

The engineering design problem is typically expressed as:

$$
\begin{aligned}
(P)' \qquad & \min \quad f(\underline{x},\underline{p}) \\
\text{subject to:} \quad & g_i(\underline{x},\underline{p}) \leq 0, \quad i = 1,\ldots,I, \\
& h_j(\underline{x},\underline{p}) = 0, \quad j = 1,\ldots,J.
\end{aligned}
$$

In this model, the vector $\underline{x}$ is a vector of design variables; that is, these are the variables that represent the freedoms of choice that the design engineer has available. The vector $\underline{p}$ is a vector of model parameters.

Johnson and Benson [8], [9], have decomposed the design problem by treating a subset of variables in the design problem as parameters ($\underline{p}$) and optimizing over the remaining variables. The original subset of "parameters" are then released and optimized.

Sobieszczanski-Sobieski, et.al., [13], [14], have used "linear decomposition" in an attempt to simplify large nonlinear models that frequently represent a design problem. This decomposition is accomplished by computing the sensitivity of the global solution to changes in subproblem solutions [15], and representing these changes by first order approximations. The ultimate advantages of this approach result from the explicit inclusion of subproblem constraints in the global optimization model and a reduction in the number of variables in the global

model. In recent work [6], Haftka has attempted to improve upon the linear decomposition method by removing the difficulties encountered due to discontinuities in the derivatives that migrate upwards through the hierarchy.

Several authors have suggest the use of knowledge-based systems to aid in the design construction. These approaches can be dichotomized by the type of knowledge required in the system and the use of that knowledge. In one case, the knowledge is used for optimization [1], [2]. Systems of this type can be thought of as addressing the more general issue of how knowledge can be used to aid in the search of optimal solutions for mathematical models that are characterized by nonlinearities. Surprisingly, the use of knowledge specific to the design domain has not been used to full advantage in these systems.

The other type of knowledge-based system contains knowledge specific to design, but does not emphasize optimization [3], [4]. While these systems can produce quality designs, the designs are based on heuristics and are therefore best suited for producing designs that meet stated specifications.

3. GOAL DECOMPOSITION

In this paper we will examine a method of decomposition that makes use of goal programming. One of the primary motivating factors for this approach is that it is felt that the employment of goals instead of "pure constraints" conforms closely to the type of decision making currently used in the initial design process. Moreover, the decision support architecture which is

examined may take advantage of a variety of other modeling techniques currently available; i.e., the Sobieszczanski-Sobieski approach, knowledge-based models, etc.

In what follows, we will use the term *objective* to mean a state that a decision maker has identified as desirable to attain. The term *attribute* is used to describe a measure which can be used to determine the degree to which the desirable state has been achieved. We will assume throughout this paper that with sufficient perseverance, a set of attributes may be determined that will measure the objective to the satisfaction of the decision maker.

3.1 Decision Variables

One of the fundamental characteristics that differentiates the goal decomposition approach from the methods of decomposition that have been found in the mechanical/aerospace engineering literature is the treatment of design variables. Consider once again the design hierarchy depicted in Figure 1. Imagine that an optimization has been performed at level 1. Information has been passed down from level 1 to level 2 via the vector $\underline{u}$. This vector contains the attributes of the subcomponent being designed at level 2 that are desired from the perspective of the component being designed at level 1. As a result, the vector $\underline{u}$ describes what it is that we want built (the subcomponent at level 2).

The optimization at level 2 in turn will describe how we will build that subcomponent by further refinement of the subcomponent into its major subcomponents. The design variables

at level 2 are a vector $\underline{w}$ that contain the desired attributes of the subcomponent at the 3rd level of the hierarchy. These values are determined from an optimization that will choose $\underline{w}$ so that the attributes of level 1 ($\underline{u}$) are met as closely as possible while not violating any constraints that are placed on $\underline{w}$.

The goal programming formulation at level i, i = 1,...,I-1, may be described as

$$\text{(P2)} \qquad \min \Sigma \; b_j^+ d_j^+ + b_j^- d_j^- \tag{1}$$

such that:

$$f_k(\underline{w}) - d_k^+ + d_k^- = u_k, \quad k = 1,\ldots,K, \tag{2}$$

$$\underline{w} \in \Omega. \tag{3}$$

As described above, the value of u_k in (2) was determined from a goal optimization at level (i-1), i = 2,...,I-1. In the case that the value of u_k was generated at the 0th level (the top level), some other multiobjective technique might have been used to determine the desired values of the attributes of the top level.

The values b_j^+ and b_j^- are weights assigned to the deviations d_i^+ and d_i^-, respectively, which are the amount by which we have exceeded or fallen short of the ith goal. The function $f_k(\underline{w})$ is the link between the subcomponent attributes and the component attribute u_k.

The actual value of the kth attribute that can be achieved in (P2) is the value of $f_k(\underline{w})$ in equation (2) (recall that the actual value that we would like to achieve is u_k*). The value of the kth attribute achieved, $f_k(\underline{w}^*)$, where $\underline{w}^*$ is the optimal value

of $\underline{w}$ in (P2), may not equal u_k^* for two reasons: (1) The goal set designated by equations (2) may be overly demanding and therefore all of the goal may not be achievable; (2) The space Ω in equation (3) may constrain the problem from achieving all of the goals.

The goal decomposition algorithm can be stated as follows:

A1. Optimize the attributes of level i, i = 1,...,I-1, by mathematically relating those attributes to attributes of the subcomponents (level (i-1)).

A2. Using the values of the level i attributes obtained in the optimization at level (i-1) as goals, optimize the attributes of level (i+1), i = 1,...,I-1.

A3. If necessary, perform a reoptimization at level i by passing constraints on the attributes from level (i+1) to level i.

Several iterations between levels (i+1) and i may be necessary to obtain a satisfactory design. The constraints that are added to the level i model in step A3 are needed to cause a reallocation of resources. An example of the addition of these constraints is shown in the first application of Section 4.

Since all of the goals in (P2) may not be met, the weighting b_i^+ and b_i^- in the objective function will tend to dictate which goals are satisfied. Next we will show how these weighting may be selected so that the pursuit of the goals at level i is consistent with the desired design attributes at level (i-1).

Theorem 1: Let (P2) describe the optimization at level i and let the following model describe the descendant optimization at level (i+1):

(P3) $$\min \Sigma\ c_i^+ e_i^+ + c_i^- e_i^- \tag{4}$$

such that:

$$g_j(\underline{x}) - e_j^+ + e_j^- = w_j^*, \quad j = 1,\ldots,J, \tag{5}$$

$$\underline{x} \in X. \tag{6}$$

Then, to first order approximation, the values of c_i^+ and c_i^- should be chosen according to the following algorithm:

(B1) For each $f_k(\underline{w}^*)$ in equation (2), use the following rules to compute new constants a_k^+ and a_k^-:

(B1.1) If $f_k(\underline{w}^*) = u_k$, let $a_k^+ = b_k^+$, $a_k^- = b_k^-$;

(B1.2) If $f_k(\underline{w}^*) > u_k$, let $a_k^+ = b_k^+$, $a_k^- = -b_k^+$;

(B1.3) If $f_k(\underline{w}^*) < u_k$, let $a_k^+ = -b_k^-$, $a_k^- = b_k^-$.

(B2) Form the values c_j^+ and c_j^- in equation (4) by

$$c_j^+ = \sum_k a_k^+ \, \partial f_k/\partial w_j \, (\underline{w}^*), \tag{7}$$

$$c_j^- = \sum_k a_k^- \, \partial f_k/\partial w_j \, (\underline{w}^*). \tag{8}$$

Proof:

Consider the first order Taylor's series expansion of $f_k(\underline{w})$ around $\underline{w}^*$:

$$f_k(\underline{w}) - f_k(\underline{w}^*) \doteq \sum \partial f_k/\partial w_j \, (\underline{w}^*) \, (w_j - w_j^*). \tag{9}$$

The term on the left hand side of (9) is the amount that the optimal value $f_k(\underline{w}^*)$ has been missed by the selection of $\underline{w}$. The right hand side of (9) contains the term $(w_j - w_j^*)$, which is the amount by which the level (i+1) optimization was unable to meet the goal w_j^* that was passed down from the level i optimization. Thus equation (9) depicts the influence that the

failure to meet goal w_j^* at level (i+1) has on the optimization at level i. If, for example, $f_k(\underline{w}^*) > u_k^*$, the optimal value of $\underline{w}^*$ caused the attribute represented by $f_k(\underline{w}^*)$ at level i to be exceeded, and $d_k^+ > 0$, $d_k^- = 0$ in equation (2). If the value of $\underline{w}$ realized from

$$w_j = g_j(\underline{x}), \quad j = 1,\ldots,J,$$

does not equal w_j^* (the goal) but instead exceeds w_j^*, then the value of the objective function at level i will change by

$$b_k^+ [f_k(\underline{w}) - f_k(\underline{w}^*)],$$

(here we assume that $\underline{w}$ is near $\underline{w}^*$ so that $f_k(\underline{w}) > u_k^*$).
But this is just the right hand side of (9) multiplied by the appropriate weight from the level i optimization. Multiplying the right hand side of (9) by b_k^+ yields the result for the special case that $f_k(\underline{w}^*) > u_k^*$ (number 2, part A of the algorithm). The other parts of the algorithm can be determined from similar arguments.

3.2 Refinements to the Goal Decomposition Algorithm

(1) If there are no constraints in equation (3) that connect w_j's from different subproblems, then a feasible design at level i may be assured easily at level (i+1) by adding the constraint

$$g_j(\underline{x}) \in \Omega$$

to problem (P3).

(2) If the optimization at level (i+1) does not meet the goals $\underline{w}^*$, then a reoptimization at level i may generate a better (or feasible) solution when taking into account the best values of the attributes that could be achieved at level (i+1). In this

case, the following cuts may be added to equation (3):

(A) If $g_j(\underline{x}^*) = w_j^*$, then add no cuts;

(B) If $g_j(\underline{x}^*) > w_j^*$, then add the cut

$$w_j > g_j(\underline{x}^*);$$

(C) If $g_j(\underline{x}^*) < w_j^*$, then add the cut

$$w_j < g_j(\underline{x}^*).$$

(3) The problem that a multipurpose component (a component with multiple parent nodes in the decomposition network) causes can be coordinated in this process by adding separate goals from each parent optimization. For example, if a multipurpose component is passed down the goal $w_1^* = 1$ from one parent optimization and $w_1^* = 2$ from the other parent optimization, then the following two goals should be added to the subproblem:

$$g_1(\underline{x}) + d_1^- - d_1^+ = 1,$$

$$g_1(\underline{x}) + d_2^- - d_2^+ = 2,$$

where $w_1 = g_1(\underline{x})$. The coefficiens in the objective function will help determine which of these contradicting goals should dominate.

4. APPLICATIONS

4.1 A Conceptual Application

The first system design that will be optimized using goal decomposition is depicted in Figure 2a. We begin with this conceptual system to illustrate the techniques that may be employed to model system cost and reliability. Similar models may be built to include attributes measuring system performance,

supportability, and producibility.

The network decomposition of this system is shown in Figure 2b. At level 1, the system is viewed as consisting of two components in series, as indicated by the dotted line in Figure 2a. We will assume that Component 1 may be viewed as consisting of two identical components in parallel, while component 2 consists of two components in series.

We shall also assume that an optimization has been performed at level 0 that specified the target values of the reliability and cost of the entire system to be $R^* = .995$ and $C^* = \$2000$. Let:

R_i = Reliability of component i, i = 1,2,
= Probability that the component operates without failure over a specified time interval,

C_i = Cost of component i, i = 1,2,

R_{ij} = Reliability of subcomponent (i,j), i = 1,2, j = 1,2,

C_{ij} = Cost of subcomponent (i,j), i = 1,2, j = 1,2.

It is also assumed that R_1 and R_2 must be at least .95 and .92 respectively, and that the total budget for the system is \$2600.

In order to build a mathematical model of the system, it is first necessary to determine how reliability influences cost. Suppose that by experience it is known that the reliability of the subcomponents is related to the cost of the subcomponents in the following way:

$$C_{11} = -200 \ln\{1 - R_{11}(2 - R_{11})\}, \quad (10)$$

$$C_{12} = C_{11}, \quad (11)$$

$$C_{21} = 1.2/(1 - R_{21}), \quad (12)$$

$$C_{22} = 1/(1 - R_{22}). \quad (13)$$

Notice that in each of these cases the cost of the subcomponent increases rapidly as the reliability approaches 1.

The development of the cost functions C_1 and C_2 for an optimization at level 1 in Figure 2b presents two significantly different problems. Component 1 consists of two identical components in parallel; and, therefore, the cost function may be written simply as

$$\begin{aligned} C_1 &= C_{11} + C_{12} \\ &= -400 \ln\{1 - R_1\}, \end{aligned} \quad (14)$$

since

$$R_1 = R_{11}(2 - R_{11}). \quad (15)$$

On the other hand,

$$C_2 = C_{21} + C_{22} \quad (16)$$

cannot be written as a simple function of

$$R_2 = R_{21}R_{22}. \quad (17)$$

Consequently, the developer of a model has two alternative approaches: (1) Include in the level 1 optimization model mathematical expressions that involve R_{21} and R_{22}, which are level 2 attributes; or (2) Estimate the cost function C_2 as a function of R_2. The former alternative regresses toward the problem that prompted decomposition in the first place: the model begins to grow larger because we have decided to include too much detail in the model at the higher levels in the network. The second alternative risks the loss of some information in the estimation process but remains in the spirit of decomposition.

For example, using equations (12) and (13), the following table of costs for alternative designs may be easily constructed:

R_{2i} (i = 1,2)	C_{21}	C_{22}
.9	$ 12	$ 10
.925	16	13.3
.95	24	20
.975	48	40
.99	120	100
.995	240	200
.999	1200	1000

Table 1--C_{21} and C_{22}

Using equation (17) and different combinations of R_{21} and R_{22}, values of C_2 may be computed from Table 1. Next, a least square estimation technique may be used to develop an estimate of C_2 as a function of R_2. Such a model was developed for this example from 15 combinations of R_{21} and R_{22} which yielded the cost function:

$$C_2 = 154.2 + 4.8(1 - R_2)^{-1}. \qquad (18)$$

The optimization model at level 1 may be written as:

(M1) $\quad \min z = 5000d_1^- + d_2^+$

such that

$$R_1R_2 - d_1^+ + d_1^- = .995$$

$$C_1 + C_2 - d_2^+ + d_2^- = 2000$$

$$C_1 + C_2 \leq 2600$$

$$d_1^+d_1^- = 0$$

$$d_2^+d_2^- = 0$$

$$.95 \leq R_1 \leq 1$$

$$.92 \leq R_2 \leq 1$$

as well as equations (14) and (18) are satisfied and all of the variables in the model are non-negative.

The coefficients in the objective function of (M1) reflect the trade-offs that were judged reasonable by the decision makers at level 0 and have been assumed for this example. The optimal solution to (M1) generated by GINO [10] is:

$z = 158.187$	$d_2^- = 0$	$C_2 = 455.48$
$d_1^- = .032$	$d_2^+ = 0$	$R_1 = .97896$
$d_1^+ = 0$	$C_1 = 1544.52$	$R_2 = .98407$

4.1.1 Level 2 Optimization Models

To generate a model that optimizes the design of the subcomponents, we need to use the values of C_i and R_i, $i = 1,2$, that were computed in the optimal solution to (M1). In addition, the results of Theorem 1 are needed to search for solutions that are consistent with the level 1 optimization. The models (M11) and (M12) for components 1 and 2 respectively and their solutions may be found in appendix A.

(M11) and (M12) contain constraints that are local to the subcomponent design as depicted by equation (6). For example, in (M11), $R_{11} \leq .8$. The solutions for C_i and R_i in (M1) are used as the goals in (M11) and (M12). Also notice that the constraints $0.95 \leq R_1 \leq 1$ have been included in (M11) as suggested in (1) of section 3.2.

Table 2 compares the values of R_i and C_i, $i = 1,2$, that were generated in the solutions of (M1) with (M11) and (M12). It is clear from this table that the largest value of R_1 that can

possibly be generated is .96, and that \$1287.55 is the largest budget needed; otherwise, additional dollars would have caused the solution of R_1 in (M11) to be higher.

Variable	Derived from Level 1	Achieved at Level 2
R_1	.97896	.96
C_1	1544.52	1287.55
R_2	.98407	.97752
C_2	455.48	455.48

Table 2--Attribute Goals and Values Realized in (M11) and (M12)

If the constraints $R_1 \leq .96$ and $C_1 \leq 1287.55$ are added to (M1), the new solution to (M1) becomes

: = 216.272	$d_2^- = 0$	$C_2 = 712.45$
$d_1^- = .043$	$d_2^+ = 0$	$R_1 = .96$
$d_1^+ = 0$	$C_1 = 1287.55$	$R_2 = .9914$

This solution is, in fact, the globally optimal design; i.e., the design that would be generated if we had built one model that had contained all of the constraints on the component and subcomponent designs.

In summary, in order to generate the globally optimal design, it was necessary to pass information downward through the decomposition network that described the desired attributes of the system. Information in the form of constraints was passed upward through the network that indicated the best achievable

values of the attributes. A reoptimization was performed that reallocated the resources (in this case budget dollars) to improve the design. Further optimization iterations, in general, would continue to improve the design.

4.2 A Second Application

Consider the problem of designing a water pumping system and a structural support for the pumping system. As shown in Figure 3a, it is assumed that the water pumping system consists of two parallel pumps that draw water from a lower reservoir and pump the water into a reservoir that is 150 feet above the lower reservoir [16]. The pumping system is to be positioned on a 30 foot reinforced concrete beam of depth 30 inches as discussed in [9]. The decomposition network for the pump/beam system is shown in Figure 3b.

Assume that a budget of $950 may be spent on the project but that the desired budget is $900 (as determined in an optimization at level 0). A goal of 200 feet for the head loss has been established to take into account the loss of energy due to friction in the pipes. This head loss must be, at a minimum, 160 feet. In addition, a goal of 6480 in k of flexural strength is desired in the beam. Define:

C_1 = Cost of the pump subsystem,

C_2 = Cost of the beam subsystem,

H = Head loss,

S = Flexural strength of the beam.

Prior experience has shown that C_1 and C_2 increase with H

and S in the following way:

$$C_1 = .01125H^2 + .875H, \quad (19)$$

$$C_2 = .05787S. \quad (20)$$

Assuming once again that the optimization at level 0 has supplied the coefficients in the objective function that reflect the trade-offs that would be considered appropriate at the top level, the pump/beam optimization model can be written as:

(M2) min $z = d_1^+ + 5d_2^- + .0926d_3^-$

such that

$$\begin{aligned} C_1 + C_2 - d_1^+ + d_1^- &= 900 \\ H - d_2^+ + d_2^- &= 200 \\ S - d_3^+ + d_3^- &= 6480 \\ C_1 + C_2 &\leq 950 \\ H &\geq 160 \\ d_i^+ d_i^- &= 0, \quad i = 1,2,3 \end{aligned}$$

equations (19) and (20) hold and all variables in the model are nonnegative. The optimal solution to (M2) is:

$z = 96.875$	$d_2^+ = 0$	$C_2 = 375$
$d_1^- = 0$	$d_3^- = 0$	$H = 183.34$
$d_1^+ = 13.59$	$d_3^+ = 0$	$S = 6480$
$d_2^- = 16.66$	$C_1 = 538.59$	

4.2.1 Level 2 Models

Next consider the optimization of the pump subject to local constraints. In this example, assume that the headloss may be expressed in terms of the flow rates (cubic feet/second), F_i, through pump i, i = 1,2:

$$H = A + 30F_1 - 6F_1^2 \quad (21)$$

$$H = B + 20F_2 - 12F_2^2 \quad (22)$$

In these expressions, A and B are parameters which specify attributes of the pumps. The selection of the pumps is restricted to

$$50 \leq A \leq 130 \quad (23)$$

$$200 \leq B \leq 220. \quad (24)$$

The amount of flow in each pump is restricted to be at least 2 cubic feet/second.

Let C_{1j} be the cost of pump j, j = 1,2. The cost of the pumps is related to the flow rate via

$$C_{11} = 100\ F_1 \quad (25)$$

$$C_{12} = 120\ F_2. \quad (26)$$

The total budget for the pump subsystem will be at least $525. The optimization model (M21) for the pump subsystem and the optimal solution is given in appendix B.

The beam subsystem has design variables W and V, which represent the width of the beam and the cross sectional area of the reinforcing steel. For this example, it is assumed that the reinforcing steel has standard sizes that come in .25 inch increments. The flexural strength, S, of the beam is related to W and V by

$$V - .2458\ V^2/W \geq S/1080. \quad (27)$$

The unit cost of the concrete is $.02/square inch/lineal foot and the unit cost of steel is $1.00/square inch/lineal foot. The length of the reinforcing steel is slightly less than the length of the beam (29.4 feet), so that the cost of the steel is given by 29.4V. The cost of the concrete is given by

.02 x (cross-sectional area of the beam) x 30

= .02 x (W x 30) x 30

= 18 W.

The width of the beam must be at least 8 inches to provide sufficient support to the pump subsystem. Once again using Theorem 1, the optimization model (M22) for the beam subsystem may be formulated and solved as shown in appendix C.

The value of V in the solution of (M22) is not a integral multiple of .25; as a result, this solution is not truly feasible given the problem constraints. In order to achieve the optimal feasible solution, two new models (M23) and (M24) may be formed by adding the constraints $V \leq 7.75$ in the first case and $V \geq 8$ in the second case to (M22). Model (M24) has no feasible solution, while the optimal solution to (M23) is given by:

$z = 12.69$	$e_2^- = 3.15$	$S = 6376.32$
$e_1^- = 103.05$	$e_2^+ = 0$	$W = 8$
$e_1^+ = 0$	$C_2 = 371.85$	$V = 7.75$

This solution is therefore the optimal attributes for the design of the beam subsystem. The solution to (M21) and (M23) may be used to generate constraints in order to further constrain the solution to (M2). This would begin a new iteration of the goal decomposition algorithm. However, a designer may decide that the design achieved by combining the solution to (M21) and (M23) to be satisfactory.

5. CONCLUDING REMARKS

This paper has attempted to develop a broader perspective

of the engineering design problem by distinguishing the type of optimization problems that occur at the top and lower levels of the design hierarchy. One key feature of middle and lower level optimizations is that subsystem optimization does not correspond directly to performance, cost, etc. optimization. A goal decomposition method was described and illustrated for subcomponent optimization. This decomposition method had a natural objective function that tended to cause the solution of the subproblem to drive the solution of the optimization one level higher to its best possible feasible solution.

The architecture of the decision support system described using goal decomposition readily enhanced the iterative nature of the design process. In the first example shown in this paper, a single reoptimization found the best global solution to the design problem once additional information was supplied to the level 1 model by subsystem optimizations.

6. REFERENCES

[1] Arora, Jasbir, and G. Baenziger, "Uses of Artificial Intelligence in Design Optimization," *Computer Methods in Applied Mechanics and Engineering*, 1986, vol. 54, pp. 303-323.

[2] Agogino, Alice, "AI in Design: Qualitative Reasoning and Symbolic Computation," preliminary proceedings for the *Study of the Design Process: A Workshop*, Oakland, California, 1987, pp. 263-294.

[3] Brown, David, and B. Chandrasekaran, "An Approach to Expert Systems for Mechanical Design," in *Proceedings of the IEEE Computer Society, Trends and Applications Conference*, 1983.

[4] Brown, David, and B. Chandrasekaran, "Knowledge and Control for a Mechanical Design Expert System," *Computer*, 1986, pp. 92-100.

[5] Clark, Dennis, and Aaron DesWispelare, "Applications of Multiple Objective Optimization Theory to a Preliminary Missile Design Study," *Design Studies*, 1985, vol.6, pp. 83-90.

[6] Haftka, Raphael, "An Improved Computational Approach for Multilevel Optimal Design," *J. Struct. Mech.*, 1984, vol. 12, pp.245-261.

[7] Johnson, R., and R. Benson, "A Basic Two-Stage Decomposition Strategy for Design Optimization," *Journal of Mechanisms, Transmissions, and Automation in Design*, 1984, vol. 106, pp.380-386.

[8] Johnson, R., and R. Benson, "A Multistage Decomposition Strategy for Design Optimization," *Journal of Mechanisms, Transmissions, and Automation in Design*, 1984, vol. 106, pp. 387-393.

[9] Liebman, J., N. Khachaturian, and V. Chanaratna, "Discrete Structural Optimization," *Journal of Structural Division, ASCE*, 1981, vol. 107, no.st11, Proceeding Paper 16643, pp. 2177-2197.

[10] Liebman, J., L. Lasdon, L. Schrage, and A. Waren, *Modeling and Optimization with GINO*, The Scientific Press, Palo Alto, Ca., 1986.

[11] Lin, Jiguan, "Multiple-Objective Problems: Pareto-Optimal Solutions by Method of Proper Equality Constraints," *IEEE Transactions on Automatic Control*, 1976, vol. AC-21, pp.641-650.

[12] Siddall, James, *Optimal Engineering Design*, New York, Marcel Dekker, Inc. 1982.

[13] Sobieszczanski-Sobieski, Jaroslaw, "A Linear Decomposition Method for Large Optimization Problems - A Blueprint for Development," *NASA TM83248*, 1982.

[14] Sobieszczanski-Sobieski, Jaroslaw, Jean-Francois Barthelemy, and Gary Giles, "Aerospace Engineering Design by Systematic Decomposition and Multilevel Optimization," *NASA TM85823*, 1984.

[15] Sobieszczanski-Sobieski, Jaroslaw, Jean-Francois Barthelemy, and Kathleen Riley, "Sensitivity of Optimum Solutions of Problem Parameters," *AIAA Journal*, 1982, vol. 20, pp.1291-1299.

[16] Stoecker, W., *Design of Thermal Systems*, McGraw-Hill Book Co., New York, 1971.

APPENDIX A

(M11) min $z = -4920.34e_1^+ + 4920.34e_1^- + e_2^+$

such that

$$
\begin{aligned}
R_1 - e_1^+ + e_1^- &= .97896 \\
C_1 - e_2^+ + e_2^- &= 1544.52 \\
R_1 &= R_{11}(2 - R_{11}) \\
R_1 &\geq .95 \\
R_1 &\leq 1 \\
R_{11} &\leq .8 \\
e_i^+ e_i^- &= 0 \quad i = 1,2
\end{aligned}
$$

as well as equations (10), (11), and (14) and all variables are nonnegative.

Optimal solution:

$z = 93.78$	$e_2^+ = 0$	$C_1 = 1287.55$
$e_1^+ = 0$	$e_2^- = 256.97$	$R_{11} = .8$
$e_1^- = 256.97$	$R_1 = .96$	$C_{11} = 643.78$

(M12) min $z = -4894.8e_1^+ + 4894.8e_1^- + e_2^+$

such that

$$
\begin{aligned}
R_2 - e_1^+ + e_1^- &= .98407 \\
C_2 - e_2^+ + e_2^- &= 455.48 \\
R_2 &= R_{21}R_{22} \\
R_2 &\geq .95 \\
R_2 &\leq 1 \\
R_{21} &\geq .9 \\
R_{21} &\leq .98
\end{aligned}
$$

$$R_{22} \geq .99$$

$$R_{22} \geq 1$$

$$e_i^+ e_i^- = 0 \quad i = 1,2$$

as well as equations (12), (13), and (16) and all variables are nonnegative.

Optimal solution:

$z = 32.04$	$e_2^+ = 0$	$R_{22} = .99747$
$e_1^+ = 0$	$e_2^- = 0$	$C_2 = 455.48$
$e_1^- = .00655$	$R_2 = .97752$	$C_{21} = 60$
	$R_{21} = .98$	$C_{22} = 395.48$

APPENDIX B

(M21) min $z = -5e_1^+ + 5e_1^- + e_2^+ - e_2^-$

such that

$$
\begin{aligned}
H \quad - \quad e_1^+ \; + \; e_1^- \quad &= \quad .98407 \\
C_1 \; - \; e_2^+ \; + \; e_2^- \quad &= \quad 538.59 \\
C_1 \quad &= \quad C_{11} + C_{12} \\
C_1 \quad &\geq \quad 525 \\
F_1 \quad &\geq \quad 2 \\
F_2 \quad &\geq \quad 2 \\
e_i^+ e_i^- \quad &= \quad 0 \quad i = 1,2
\end{aligned}
$$

as well as equations (21)-(26) and all variables are nonnegative.

Optimal solution:

$z = 73.25$	$C_1 = 525$	$F = 4.71$
$e_1^+ = 0$	$C_{11} = 199.54$	$F_1 = 2$
$e_1^- = 17.37$	$C_{12} = 325.46$	$F_2 = 2.71$
$e_2^+ = 0$	$A = 130$	$H = 165.97$
$e_2^- = 13.59$	$B = 200$	

APPENDIX C

(M22) min $z = .0926e_1^- - e_2^+ + e_2^-$

such that

$$S - e_1^+ + e_1^- = 6480$$

$$C_2 - e_2^+ + e_2^- = 375$$

$$C_2 = 29.4V + 18W$$

$$W \geq 8$$

$$e_i^+ e_i^- = 0 \quad i = 1,2$$

as well as equation (27) and all variables are nonnegative.

Optimal solution:

$z = 12.69$ $\quad C_2 = 375$

$e_1^+ = 0$ $\quad S = 6437.17$

$e_1^- = 103.05$ $\quad W = 8$

$e_2^+ = 0$ $\quad V = 7.857$

$e_2^- = 3.15$

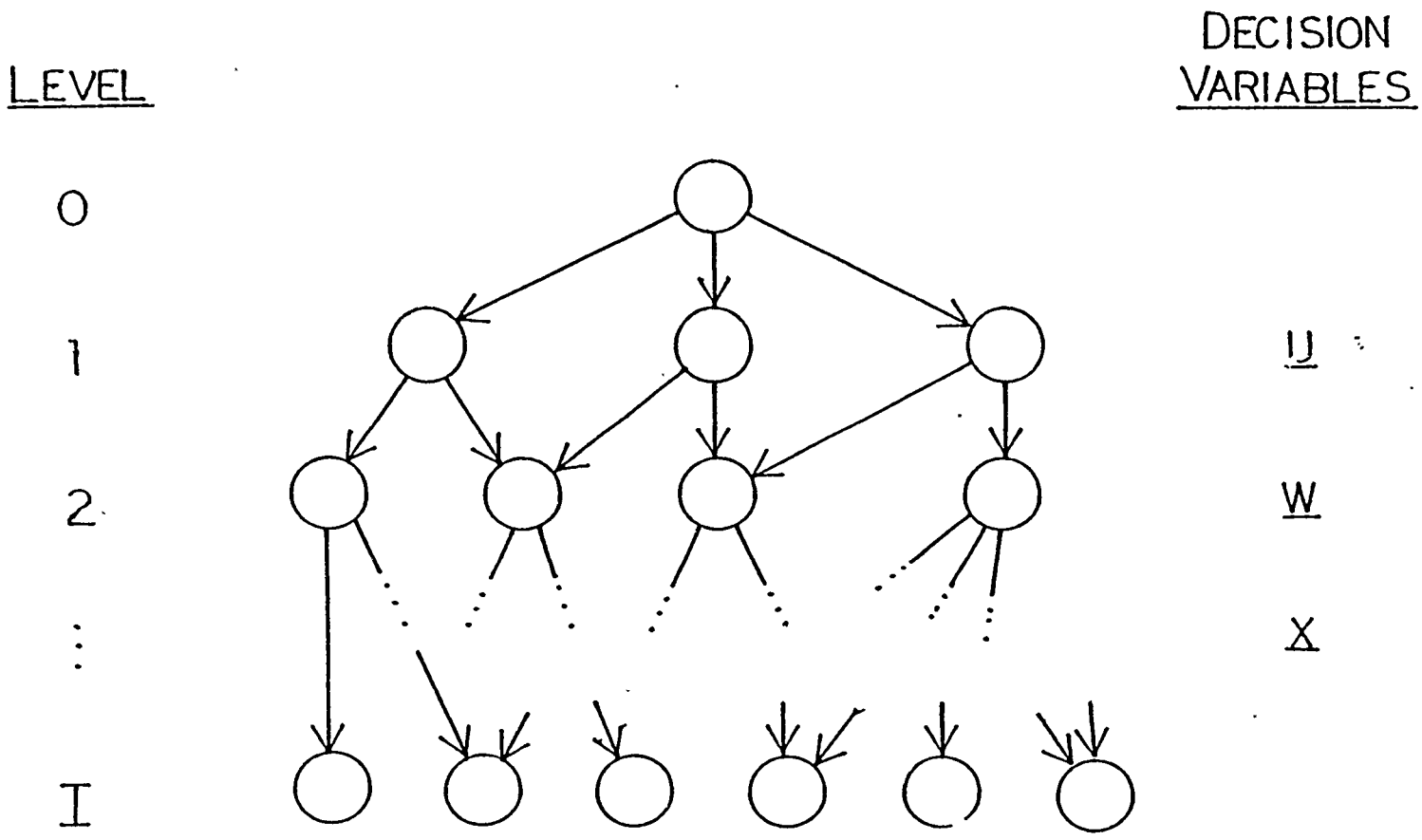


Figure 1--Decomposition Network

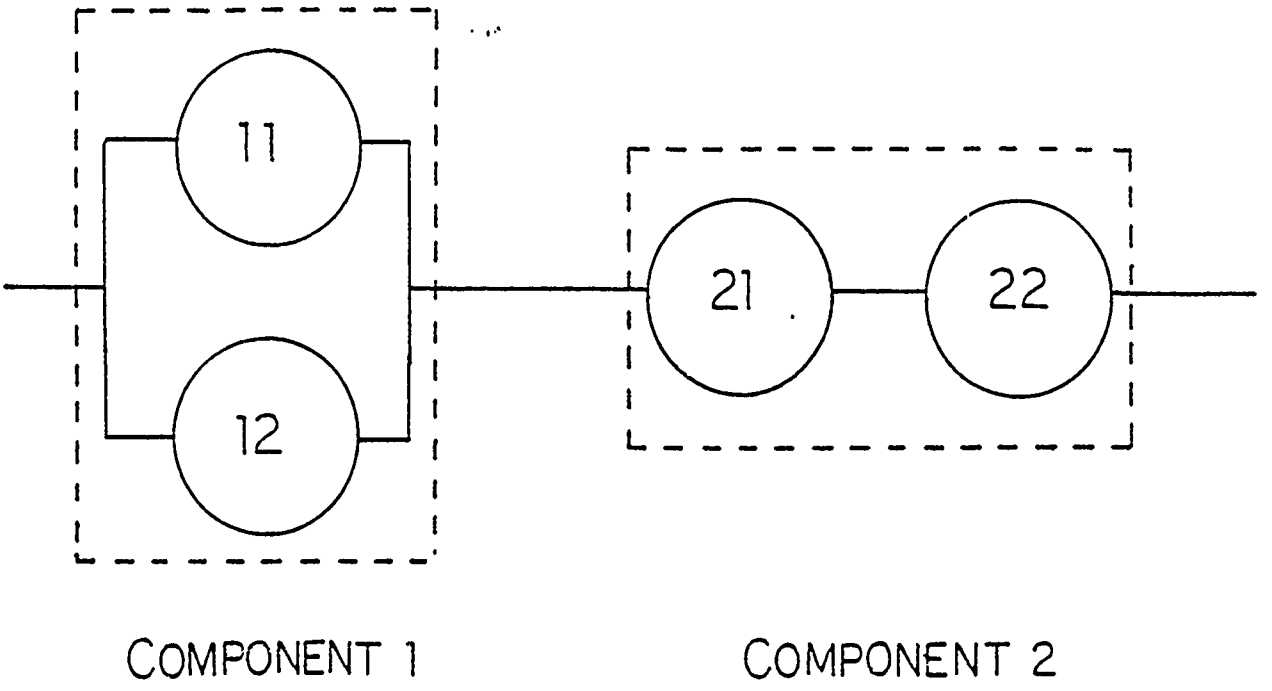


Figure 2a—A Conceptual System

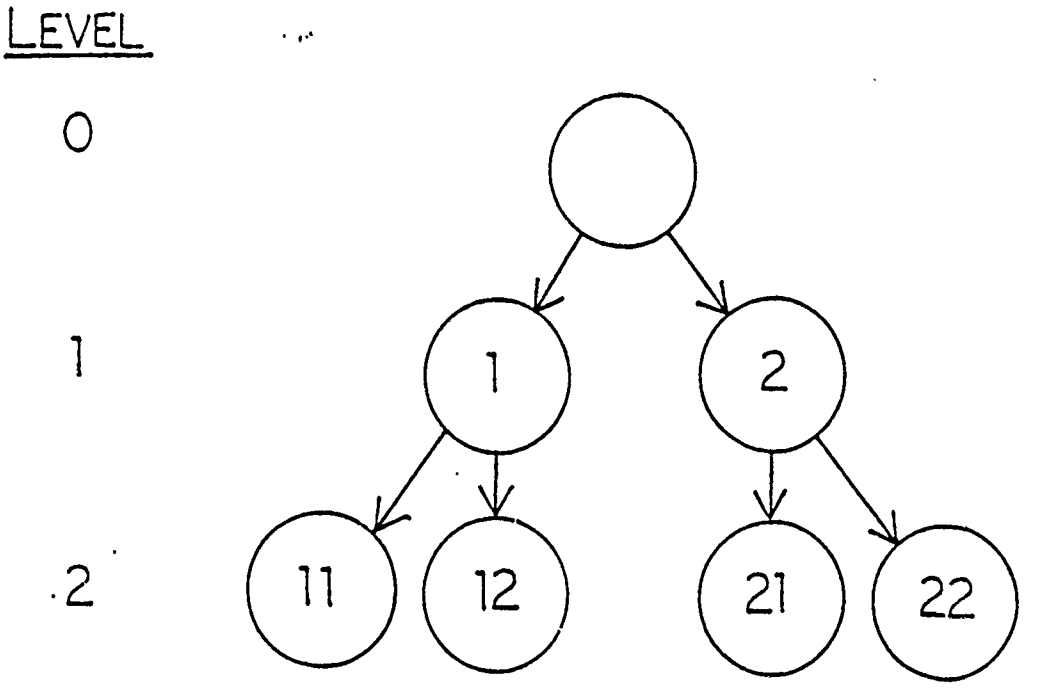


Figure 2b--Network Decomposition
of the Conceptual System

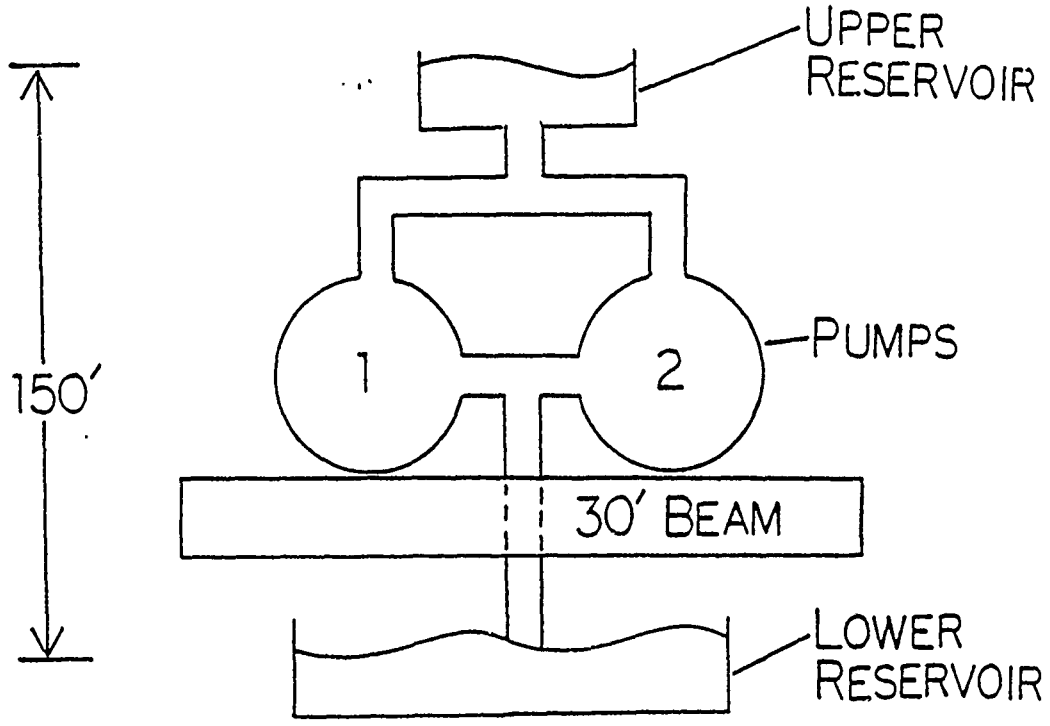


Figure 3a--Water Pumping System

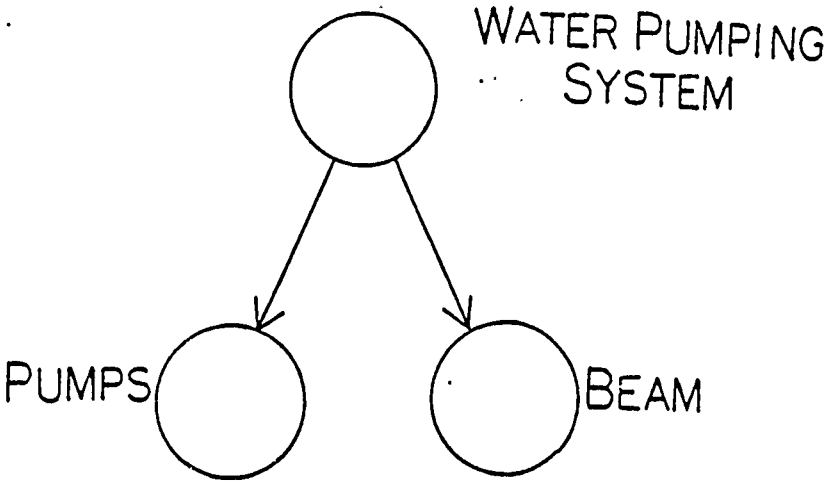


Figure 3b--Network Decomposition of the Water Pumping System

FINAL REPORT NUMBER 121
REPORT NOT AVAILABLE AT THIS TIME
Dr. Robert K. Young
210-9MG-099

1988 USAF-UES FACULTY RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Universal Energy Systems, Inc.

FINAL REPORT

SOLVENT EXTRACTION OF BORON FROM INDUSTRIAL WASTEWATERS

Prepared by:

Steven C. Chiesa, Ph.D., P.E.
Assistant Professor
Department of Civil Engineering
Santa Clara University
Santa Clara, CA 95053

Contract No: 5210-9MG-102

March 30, 1990

SOLVENT EXTRACTION OF BORON FROM INDUSTRAL WASTEWATERS

Steven C. Chiesa
Department of Civil Engineering
Santa Clara University

ABSTRACT

As the demand for irrigation water throughout the western portion of the United States grows, the use of reclaimed wastewater for agricultural use is being contemplated with increasingly greater frequency. The suitability of reclaimed wastewater for irrigation is strongly dependent on the post-treatment concentration of certain effluent constituents. Boron is one of these critical effluent constituents. When present in sufficiently high concentrations, boron is toxic to many forms of plant life, including many agriculturally valuable crops. Boron behaves as a conservative pollutant in conventional secondary wastewater treatment systems with little net removal normally expected or observed. Pretreatment of industrial wastewaters to remove boron may, accordingly, be necessary where high effluent boron concentrations limit the beneficial uses of reclaimed wastewaters and/or surface water receiving conventionally treated effluents.

This research project evaluated liquid/liquid extraction as a means of removing boric acid from industrial wastewaters. An Air Force photography lab wastewater was used as a test case. Two organic solutions were examined for their relative ability to serve as effective extractants for boron. A solution containing 4-t-butylcatechol as the principal reactive agent was found to be the best extractant for treating low ionic strength aqueous solutions. The extractant could be stripped of boron during subsequent contact with a dilute hydrocholoric acid solution.

INTRODUCTION

Many communities in the western United States are contemplating the use of reclaimed wastewater for irrigation. The suitability of reclaimed wastewater for this type of disposal/resource recovery option is strongly influenced by the concentration of residual constituents in the treated effluent. Boron is one particular element which, when present in sufficiently high concentrations, can potentially render a reclaimed wastewater unfit for irrigation purposes. Boron phytotoxicity has been well documented in both agricultural and water quality oriented literature sources [1-3]. Concentrations in irrigation waters as low as 0.5 mg/L have been recommended to protect boron sensitve crops while other crops have been shown to tolerate irrigation waters containing up to 4.0 mg/L of boron [2,3]. Anticipated effluent boron concentrations should be compared against these guidelines to determine if pretreatment and/or alternative disposal options should be explored. Even where direct irrigation with reclaimed wastewater is not contemplated, receiving stream water quality standards may dictate that boron source control be considered.

The need to investigate pollution control technologies designed and operated specifically for boron removal is related to the conservative nature of this element in conventional secondary treatment facilities. Waggott [3], Dagon [4] and Lapp and Cooper [5] have examined the fate of environmentally relevant forms of boron in a variety of secondary and tertiary treatment operations and found little net removal from aqueous carrier streams. Conventional biological treatment processes, in addition, tolerate boron concentrations well in excess of those identified as being inhibitory to even the more resistant types of plant life [6,7]. Given the aforementioned constraints, mandated pretreatment of industrial wastewaters for boron

removal becomes one of the few viable options available to sewerage system managers when effluent boron levels must be controlled. Ion exchange, using boron selective resins [5,8-10], and liquid/liquid extraction operations [11-16] both have been employed to remove boron from aqueous carrier streams. While ion exchange systems have proven to be capable of maintaining very high boron removal efficiencies, the economics of this option are strongly influenced by preliminary treatment requirements and resin regeneration chemical costs [5,9,10]. Use of liquid/liquid extraction for boron removal has been limited to concentrated brines where boric acid recovery was used to offset costs associated with phase separation operations. Few full-scale boron removal facilities have been constructed based on these technologies and fewer, if any, have been designed to operate with low ionic strength industrial wastewaters as the aqueous boron carrier phase.

This investigation was conducted specifically to determine the ability of liquid/liquid extraction technology to effect the removal of boron from low ionic strength aqueous carrier solutions. Boric acid was selcted for evaluation as it is the most frequently encountered form of boron in both domestic and industrial wastewaters. Recent innovations in liquid/liquid extraction technology have also made a wider variety of waste streams potentially amenable to this method of treatment. The development and production of high quality hydrophobic and hydrophilic porous membrane systems, in particular, have extended the potential concentration range of liquid/liquid extractions to encompass relatively dilute influent carrier solutions [17-21]. Successful application of this technology for boron removal is dependent on identifying a minimally water soluble organic solvent which is capable of efficient extraction of boron from an aqueous carrier and chemically compatible with the membrane material.

MATERIALS AND METHODS

Test Extractants. The two organic extractant solutions evaluated as part of this investigation are indicated in Table 1. Both extractant systems have been reported in the literature as being capable of effectively removing boron (as boric acid) from brines and highly acidic solutions.

2-ethylhexanol (Alrich Chemical Company) represented a simple, single component organic extraction system.

4-t-butylcatechol (TBC) (Aldrich Chemical Company) was used in conjunction with tri-(n-octyl)amine (Sigma Chemical Company) to reduce the requirements for co-extraction of aqueous phase cations [11]. Both components were dissolved in Shell 42 (Shell Oil company), a high-boiling point alkylated aromatic hydrocarbon solvent. Preparation of the mixture followed the procedure described by Grinstead [11]. Extractions with TBC-containing solutions were conducted only after the aqueous phase pH had been adjusted to either 4.5 or 5.0 to prevent oxidation of the catechol [11]. Preliminary extractions carried out with aqueous phase pH values lower than 4.0 resulted in significant reductions in extraction efficiency (.ie. lower mass distribution coefficients). Subsequent experimentation with the TBC-based extractant focussed on conditions yielding optimum partitioning.

Test Aqueous Solutions. Most extraction runs were conducted with a synthetic wastewater solution containing 10 to 50 mg l^{-1} boron (as boric acid) and 500 mg/L calcium chloride in distilled water. pH was adjusted with 1.0 N hydrochloric acid as required.

Boron extraction was also evaluated using effluent from an Air Force base photo processing facility. Boron-containing compounds are used in photo processing operations as part of photo developing and fixing solutions. Here again, boron existed in solution as boric acid. Previous sampling of this

Table 1. Boron Extractants Evaluated

Extractant	Reported Value(s) of Boron Mass Distribution Coefficient	Stripping Solution	Reference
2-Ethylhexanol	15.5 - 16.8	caustic	[13]
0.2 M TBC + 0.2 M TOA in Shell 42	61	acid	[11]

TBC = 4-t-butylcatechol

TOA = tri-(n-octyl)amine

waste stream indicated that boron concentrations generally ranged between 3.0 and 50.0 mg/L. The 15 liter sample of wastewater collected for these experiments contained 4.4 mg/L boron. A concentrated boric acid stock solution was used to increase the initial boron concentration for many of the batch extractions. Aluminum sulfate (as alum, $Al_2(SO_4)_3*18H_2O$) was added to the photo processing wastewater to reduce remove color related compounds which apparently contributed to the formation of a potentially problematic film observed in a number of preliminary batch extractions. Extractions were conducted on samples generated after settling and filtering the alum-treated effluent. Boron concentrations were determined and final pH adjustments were made only after the alum pretreatment step.

Characterization of Extractants. The potential efficiency of an organic solvent in a liquid/liquid boron extraction operation is related to the partitioning of boron between the immiscible solvent and aqueous phases. A system distribution coefficient was used in these experiments to quantify partitioning between the two phases. The coefficient was numerically equal to the ratio of target solute equilibrium concentrations in the organic solvent and aqueous phases [22].

Batch extractions were conducted by contacting/mixing equal volumes of wastewater and organic solvent (phase ratio = 1:1). The majority of extractions were conducted using 20 ml (total volume) Mixor liquid/liquid extraction systems. Approximately 200 piston strokes were used as part of the contact phase to insure that equilibrium conditions were achieved. The Mixor systems provided suffient sample volumes to analyze the post-extraction aqueous phase for residual boron and the solvent phase for boron re-extraction potential. Re-extraction of boron from solvent phases was conducted in 10 ml (total volume) Mixor systems. Boron distribution

coefficients between organic solvent phases and stripping solutions were again determined using post-contact equlibrium boron concentrations.

A limited number of extractions were conducted using 150 ml of aqueous and solvent components. These larger volume extractions were used detect emulsion formation or other potentially problematic side reactions. Separation of the two phases after mixing was completed using a separatory funnel. All extractions (and re-extractions) were conducted at room temperature ($22\pm2^{o}C$).

Continuous Flow Experiments. Continuous flow liquid/liquid extraction experiments were conducted using bench-scale contained liquid membrane extraction modules. The prefabricated systems (Hoechst Celanese model number E-400/0.7) consist of two intertwined bundles of porous, hydrophobic hollow fibers contained within a polypropylene chamber. The modules allow countercurrent flow of aqueous waste and stripping streams within the two fiber bundles. The solvent phase, contained by the polypropylene chamber, surrounds the fiber bundles and acts as a carrier medium between the two aqueous phases. Any ionic or molecular form soluble in the organic extractant will diffuse across the waste-solvent interface, diffuse through the stationary solvent phase and then diffuse into the stripping solution until equilibrium conditions are achieved. By maintaining aqueous phase operating pressures slightly greater than the organic phase operating pressure, solvent flow through the fiber pores is prevented and aqueous phase-solvent phase interfaces are continuously maintained at the fiber surfaces. Such membrane systems provide a mechanistically simple means of non-dispersive extraction of a target solute from an initial carrier phase to a second, immisible, solvent phase. Use of the continuous flow stripping solution provides in-situ regeneration of the organic extractant [23].

Extractions were conducted by recycling waste and stripping solutions between completely mixed reservoirs and the extraction module. Flow rates for waste and stripping streams were equal and constant. The waste volume treated in each run was 2.0 liters. The stripping solution volume was also 2.0 liters. pH control in the wastewater reservoir was accomplished, where specified, by adding concentrated solutions of sodium hydroxide or hydrochloric acid to maintain pH within 0.1 pH units of the target value.

Analytical Procedures. Boron concentrations were determined by the carmine method [24]. pH was determined with an Orion Model SA 230 pH meter.

RESULTS

Batch Extractions Using the Synthetic Wastewater. Results of the individual boron extraction runs using 2-ethylhexanol as the principal extractant for an aqueous phase pH of 4.5 are presented in Figure 1. A boron mass distribution coefficient of 0.29 (r^2 = 0.994) was determined for all data generated with aqueous phase pH values between 4.0 and 7.0. Results for the 2-ethylhexanol-based extractions verified previous observations related to the effect of pH on boron partitioning. This value for the mass distribution coefficient was significantly lower than the range of values reported by Kristanova *et al.* [13] using the same extractant. The lower distribution coefficient observed in this work is probably the result of the relatively low hydrogen ion concentrations and ionic strengths provided in these experiments compared with those employed by others [12-16].

Results of boron extraction runs using TBC as the principal extracting agent in contact with a synthetic waste solution at pH 4.5 are presented in Figure 2. A calculated boron mass distribution coefficient of 30.4 (r^2 = 0.986) represents a significant improvement over the value determined for 2-

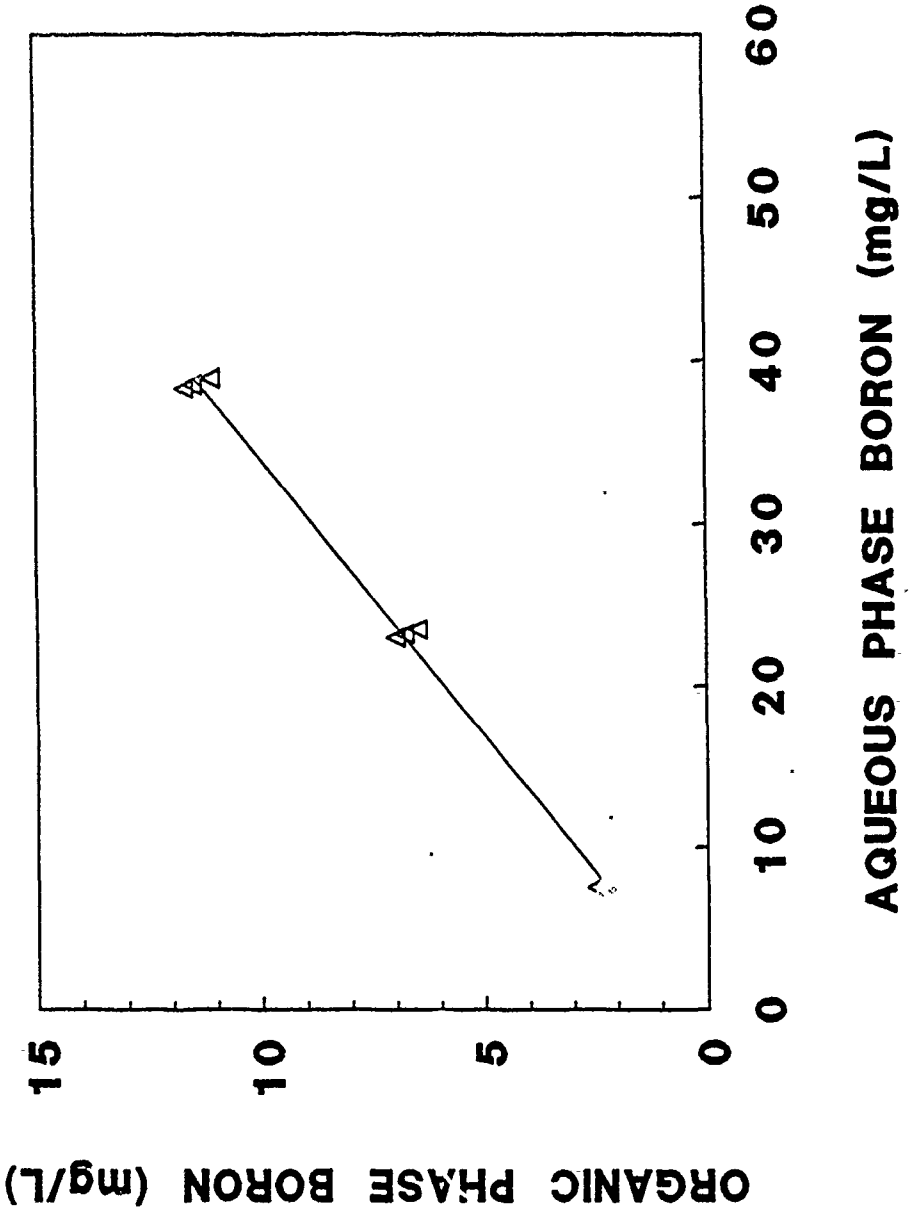


Figure 1. Equilibrium Data for Batch Extraction of Boron with 2-Ethylhexanol. Aqueous phase pH = 4.5 , phase ratio = 1:1.

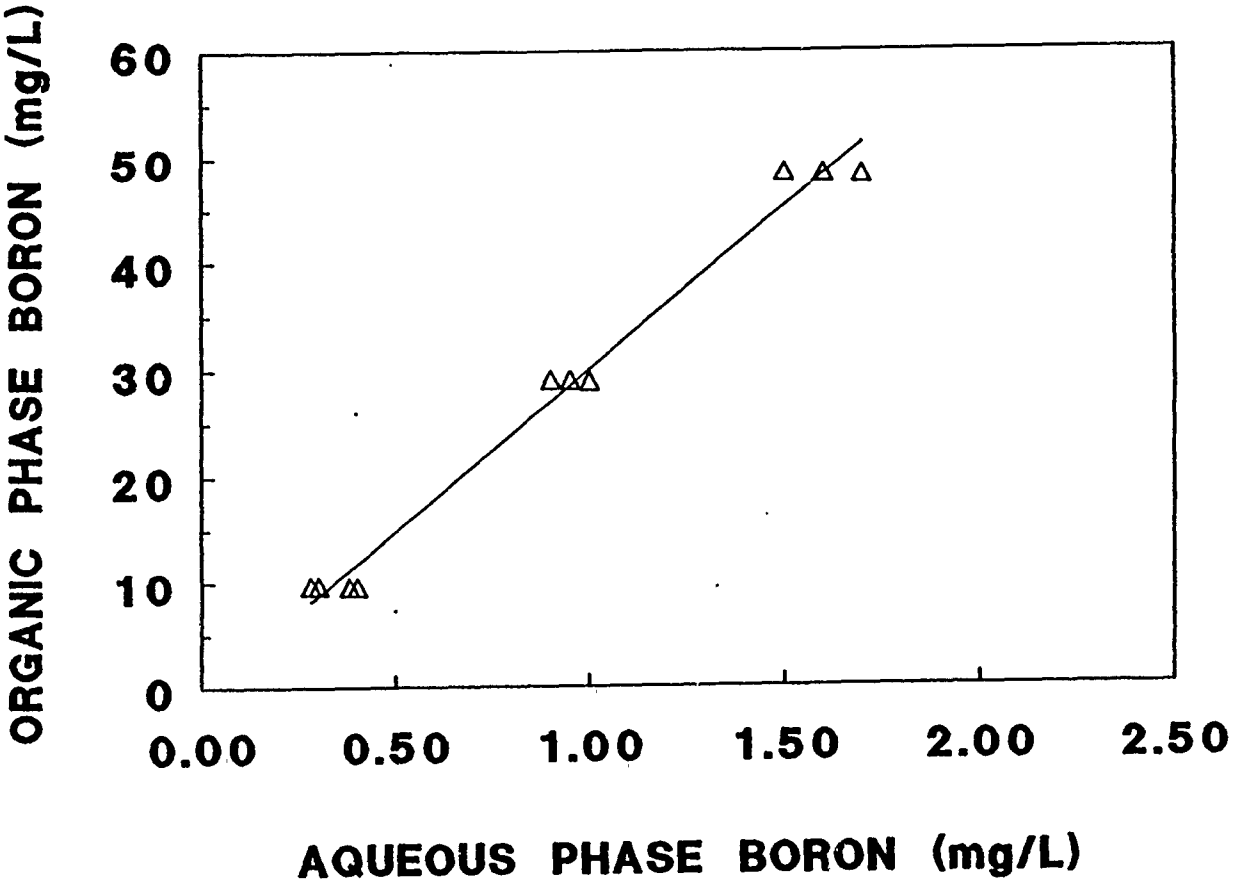


Figure 2. Equilibrium Data for Batch Extraction of Boron with 0.2 M TBC / 0.2 M TOA in Shell 42. Aqueous phase pH = 4.5 , phase ratio = 1:1.

ethlyhexanol. The value also compares reasonably well with values reported in the literature for boron extraction from calcium and magnesium chloride brines [11]. TBC-based boric acid extractions were apparently much less sensitive to variations in ionic strength than are 2-ethylhexanol-based extractions.

Batch Extractions Using Photo Processing Wastewater. The photo processing wastewater evaluated had a slight yellowish color and a pH of 6.5. Solution pH was reduced to 5.5 and 4.5 with the addition of 80 and 500 mg/L aluminum sulfate, respectively. With this range in alum dosage, precipitation reactions were observed at pH values of 5.5 and greater. Where precipitation did occur, there was essentially complete removal of color from waste samples. No color removal was recorded for conditions not yielding precipitates. Precipitation and removal of precipitated solids increased the speed of subsequent aqueous-organic solvent phase separations.

Boron mass distribution coefficients for solvent-photo processing wastewater batch extractions were essentially the same as the values determined for the synthetic wastewater. Other ionic components in the photo processing wastewater did not apparently contribute to any "salting out" effect. Likewise, the alum dosages utilized did not influence equilibrium boron concentrations.

Solvent Re-Extraction Experiments. Concentrating boron in a stripping solution is an essential element of any proposed liquid/liquid extraction treatment system. If boron partitioning between the solvent phase and aqueous stripping solution is not significantly biased towards the aqueous phase, large volumes of the stripping solution must be used to sequester the boron removed from wastewater streams in the intial extraction step. Dilute stripping solutions will result in higher diposal/processing costs.

Results of solvent re-extraction experiments are presented in Table 2. For 2-ethylhexanol, the mass distribution coefficient for the solvent-caustic stripping solution system was identical to that determined for the test wastewaters. This implies a volume of stripping solution well in excess of the volume of waste treated would be necessary to signifcantly reduce effluent boron concentrations.

The TBC-based extraction solution was effectively stripped of boron by a 2.0 N hydrochloric acid solution. When taken in conjuction with the waste-solvent mass distribution coefficient, boron concentrations in the stripping solution could easily be magnified by several orders of magnitude compared to effluent boron levels. Small volumes of a concentrated boric acid solution could then be generated for possible boron recovery.

Extractions Using Contained Liquid Membrane Systems. Operation of the contained liquid membrane systems was designed to qualitatively demonstrate the ability of such systems to remove target waste impurities. Results for test runs using each organic extractant are presented in Figures 3 and 4. The run conducted with 2-ethlyhexanol, as shown in Figure 3, proved that the contained liquid membrane system was capable of tranfering boron from the wastewater sample to a "remote" stripping solution using the immisible organic solvent as a carrier/shuttle. The equilibrium boron concentrations were consistent with the volumes of the two aqueous solutions and values of the mass distribution coefficients determined during earlier batch testing. Equilibrium condistions were effectively reached after six hours or nine turnovers of the wastewater reservoir volume.

Contained liquid membrane system operation with the TBC-based organic extractant, as presented in Figure 4, is typical of two runs made with the prefabricated extraction modules. In both cases, no reduction in wastewater

Table 2. Stripping Potential of Solvent-Extracted Boron

Solvent	Stripping Solution	Initial Solvent Phase Boron Conc. (mg/L)	Post-Extraction Boron Conc. (mg/L) Solvent	Strip	Boron Mass Distribution Coefficient
2-Ethylhexanol	5.0 N NaOH	11.5	2.7	8.7	0.31
		6.8	1.7	5.1	0.33
		2.3	0.5	1.8	0.28
0.2 M TBC + 0.2 M TOA in Shell 42	2.0 N HCl	48.5	0.5	48.0	0.010
		28.9	0.4	28.5	0.014

TBC = 4-t-butylcatechol

TOA = tri-(n-octyl)amine

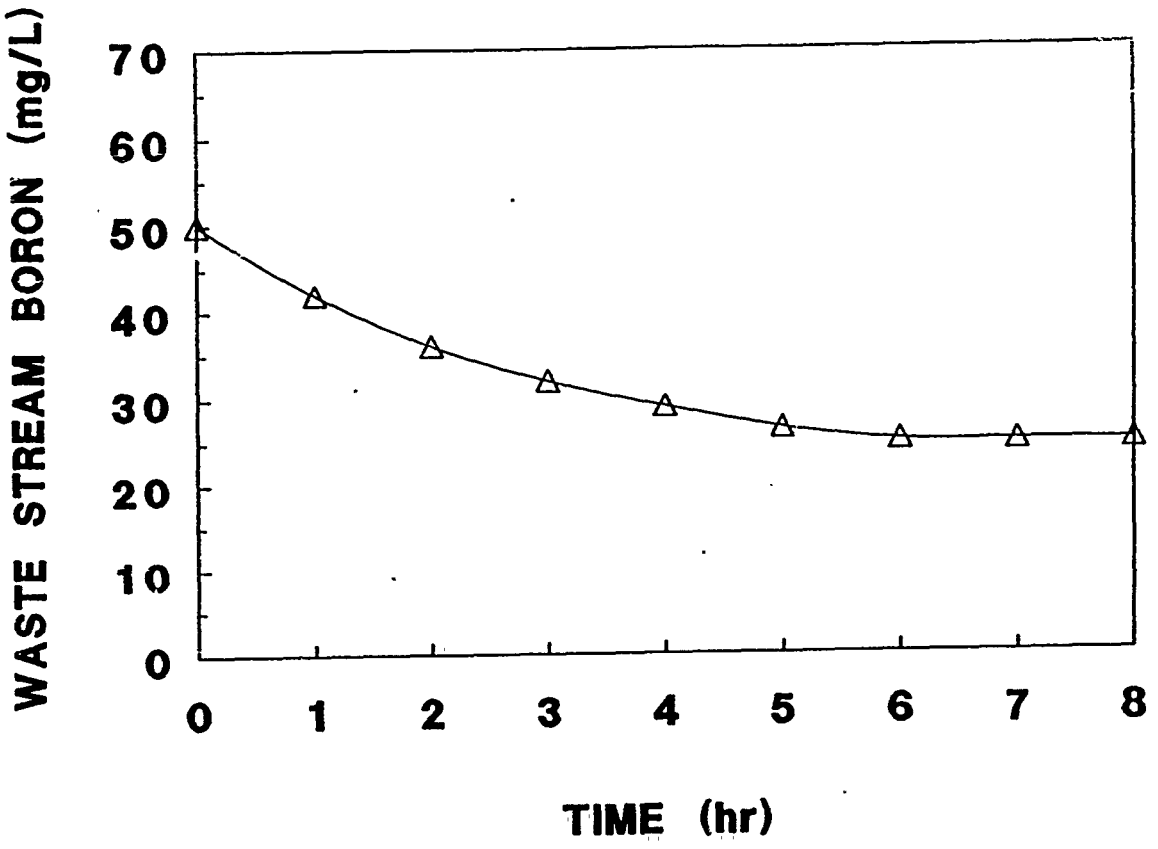


Figure 3. Boron Extraction from Synthetic Wastewater Using A Contained Liquid Membrane Extraction Module. Extractant = 2-ethylhexanol, wastewater pH = 5.0, waste volume treated = 2.0 L, stripping solution = 5.0 N NaOH, stripping solution volume = 2.0 L.

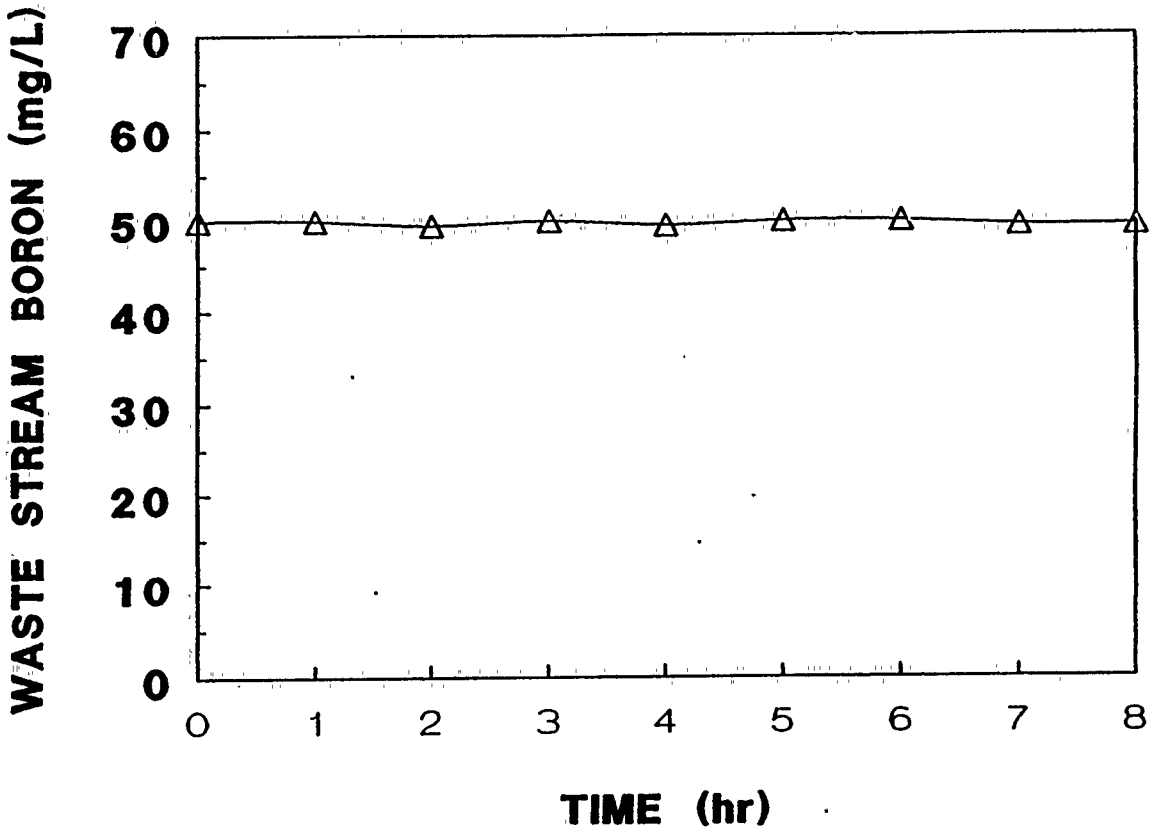


Figure 4. Boron Extraction from Synthetic Wastewater Using A Contained Liquid Membrane Extraction Module. Extractant = 0.2 M TBC / 0.2 M TOA in Shell 42, wastewater pH = 5.0, waste volume treated = 2.0 L, stripping solution = 2.0 N HCl, stripping solution volume = 2.0 L.

boron concentration was observed. This was particularly surprising seeing that equilibrium of the synthetic wastewater with the solvent alone (no stripping) would have resulted in 35 percent removal of the waste boron. The first extraction run provided control of the wastewater reservoir pH at a value of 5.0. In examining the system after the first run, it was found that maintaining pH in the waste solution reservoir at a target value did little to control the pH in the extraction module. Due to high fiber surface area and the rapid production (and/or transfer) and low buffering capacity, the pH of the aqueous waste stream was rapidly reduced to approximately 2.0 and the initial extraction step was apparently inhibited. The second extraction run incorporated a well-buffered wastewater sclution to minimize pH reductions in the extraction module. The absence of lack of boron removal realized in this run, however, was again inconsistent with batch extraction results.

The unexpectedly poor performance of the contained liquid membrane extraction system with the TBC-based extractant may be related to two possible phenomenon. In one cases, a stagnant aqueous film with a consisitently low pH may have formed at the wastewater-solvent interface, preventing effective partitioning between these two phases. Alternatively, the aromatic hydrocarbon solvent (Shell 42) used in the extractant mixture may have adversely affected the fiber pore structure thereby preventing mass transfer of the boron from one fiber bundle to the other.

Equilibrium boron concentrations determined for aqueous solutions in waste and stripping solution reservoirs were used in mass balance determinations for total system boron. Final mass balances closed to within three percent of the initial mass of boron in the system. The solvent phase volume was less than 40 ml and contained only a small fraction of the boron at equilibrium.

DISCUSSION

Removal of boron from industrial wastewater streams using liquid/liquid extraction appears to be technologically viable. While a number of extractants have been shown to be efffective partioning agents for boron contained in brines and highly acidic solutions, many of these extractants will not be as effective in removing boron from dilute waste streams. Low mass distribution coefficients associated with boron extraction of dilute wastewaters may eliminate many extractants from consideration unless exceptionally effective stripping solutions can be identified. Use of liquid/liquid extraction to remove boron from dilute industrial wastewaters, however, should focus on those extractants exhibiting high boron mass distribtion coeffients under field conditions together with an ability to be effectively stripped of boron by a second aqueous stream. Such a combination of characteristics will maximize boron concentrations in stripping streams, minimize disposal costs and increase the potential for boric acid recovery. The TBC-based extractant system evaluated as part of this investigation appears to meet these general criteria.

Contained liquid membrane systems appear to be an attractive alternative to conventional liquid/liquid extraction technology and could compete with ion exchange or adsorption for the removal of a variety of pollutants. No special facilities are required for phase separation and relatively small volumes of solvent are required. Mathematical models are currently being refined in an efffort to help engineers design and/or predict the performance of full-scale systems [25].

Any assessment of contained liquid membrane technology for the removal of boron or any other target waste constituent, however, must consider a number of related factors which may influence system technical feasibility

and cost. Organic extractants must be chemically compatible with the fiber material. Preliminary treatment requirements must be identified particularly as they relate to potential clogging problems within the hollow fibers. Larger diameter fibers can be used to minimize clogging by particulate material but only at the expense of reduced fiber volume-specific surface areas and lower mass transfer rates. Reactions in thin, stagnant difussion layers at phase boundaries must also be identified to insure that partitioning will not be impeded of inhibited. Chemical costs associated with the use of a particular extractant (.ie. for pH control, etc.) must be quantified. The impact of small quantities of extractant in the treated effluent stream must be evaluated in order to prevent creating greater environmental problems than the one solved by treatment. Lastly, unintentional removal of waste constituents other than target pollutants should also be investigated for potential impacts on stripping solution disposal/resource recovery options. In summary, while contained liquid membrane systems offer a number of desirable characteristics, extensive pilot-scale testing should be conducted on any waste stream before full-scale application of the technology is implemented.

V. CONCLUSIONS:

The following conclusions have been reached based on the experimental results detailed in previous paragraphs.

* Boron, when present in a dilute aqueous solution as boric acid, can be extracted using a suitable organic solvent. Alkyl catechols, such as 4-t-butylcatechol, appear to be capable of removing boron from low ionic strength aqueous streams in liquid/liquid extraction processes.

* A porous hollow fine fiber membrane system, based on contained liquid membrane extraction technology, can effect the removal of boron from dilute, low ionic strength wastewaters. The organic solvent solution making up the contained liquid membrane, must be compatible with the hollow fiber material and sufficent pretreatment must be provided to prevent clogging of individual fibers. Reactions potentailly impacting the composition of stagnant phase boundary layers must also be evaluated to insure extraction.

* Alum pretreatment of a test photo processing wastewater appeared to eliminate any potential problems with interfacial film formation. Formation of stable emulsions was never observed. Extraction of boron from the photo processing wastewater was identical to that observed for a low ionic strength synthetic wastewater.

REFERENCES:

1. Wilcox, L.V., "Water Quality from the Standpoint of Irrigation," J. Am. Water Works Assoc., 1958, Vol. 50, pp. 650-654.

2. McKee, J.M. and R.W. Wolf, "California Water Quality Criteria," Publication 3A, Agricultural Water Supply (State of California), 1963, pp. 106-114.

3. Waggot, A., "An Investigation of Potential Problem of Increasing Boron Concentrations in Rivers and Water Courses," Water Res., 1969, Vol. 3, pp. 749-765.

4. Dagon, T.J., "Biological Treatment of Photographic Processing Effluents," Eastman Kodak Company Technical Report No. J-46, 1967.

5. Lapp, T.W. and G.R. Cooper, "Chemical Technology and Economics in Environmental Perspective - Task II - Removal of Boron from Wastewater," USEPA Report No. 560/1-76-007, 1976.

6. Banerji, S.K. , et al., "Effect of Boron on Aerobic Biological Waste Treatment," Proceed. 3rd Purdue Ind. Waste Conf., 1968, pp. 956-965.

7. Webber, W.G., et al., "Study on the Effect of Boron Toxicity on an Activated Sludge System," Proceed. 31st Purdue Ind. Waste Conf., 1976, pp. 743-752.

8. Roberts, R.M., "Deboronation Process by Ion Exchange," Ind. Engng. Chem. Prod. Res. Dev., 1971, Vol. 10, pp. 356-357.

9. Wong. J.M., "Boron Control in Power Plant Reclaimed Water for Potable Use," Environ. Prog., 1984, Vol. 3, pp. 5-11.

10. Okay, O., et al., "Boron Pollution in the Simav River, Turkey and Various Methods of Boron Removal," Water Res., 1985, Vol. 19, pp. 857-862.

11. Grimstead, R.R., "Removal of Boron and Calcium from Magnesium Chloride Brines by Solvent Extraction," Ind. Engng. Chem. Prod. Res. Dev., 1972, Vol. 11, pp. 454-460.

12. Grizo, A.N. and N.T. Petrovska, "Extraction of Boric Acid from Dilute Sulfuric Acid Solutions with Organic Extracting Agents. I. Extraction with 2-Ethylhexanol," Glas. Hem. Drus. Beograd, 1983, Vol. 48, pp. 513-518; Chem. Abs., 1984, Vol. 100: 92168.

13. Kristanova, G.N., et al., "Extraction of Boron from Bulgarian Natural Brines with 2-Ethylhexanol," Izv. Khim., 1984, Vol. 17, pp. 211-217; Chem. Abs., 1985, Vol. 102: 81057.

14. Vinogradov, E.E., et al., "Extraction of Boric Acid form Sulfuric Acid Solutions," Zh. Neorg. Khim., 1986, Vol. 31, pp. 1004-1009; Chem. Abs., 1986, Vol. 104: 194187.

15. Shamiryan, P.S., et al., "The Salting-Out Effects of Cations on Extraction of Boric Acid with 2-Ethylhexanol from Aqueous Sulfate Solutions," Arm. Khim. Zh., 1986, Vol. 39, pp. 570-574; Chem. Abs., 1987, Vol. 106: 144888.

16. Eto, Y., et al., "Treatment of Water Containing Boron," Japan Patent 62,121,689; Chem. Abs., 1987, Vol. 107: 182850.

17. Cooney, D.O. and C. Jin, "Solvent Extraction of Phenol from Aqueous Solution in a Hollow Fiber Device," Chem. Eng. Commum., 1985, Vol. 37,

pp.173-191.

18. D'Elia, N.A., et al., "Liquid-Liquid Extractions with Microporous Hollow Fibers," J. Membr. Sci., 1986, Vol. 29, pp. 309-319.

19. Frank, G.T. and K.K. Sirkar, "An Integrated Bioreactor-Separator: In-Situ Recovery of Fermentation Products by a Novel Membrane-Based Dispersion-Free Solvent Extraction Technique," Biotech. Bioeng. Symp. No. 17, John Wiley & Sons, New York, NY, 1987, pp. 303-316.

20. Prasad, R. and K.K. Sirkar, "Microporous Membrane Solvent Extraction," Sep. Sci. Technol., 1987, Vol. 22, pp. 619-640.

21. Muscatello, A.C., et al., "Supported Extractant Membranes for Americium and Plutonium Recovery," Sep. Sci. Technol., 1987, Vol. 22, pp. 843-853.

22. Dilts, R.V., Analytical Chemistry, Van Nostrand Co., New York, NY, 1974.

23. Sengupta, A., et al., "Separation of Solutes from Aqueous Solutions by Contained Liquid Membranes," Amer. Inst. Chem. Eng. J., Vol. 34, 1988, pp. 1698-1708.

24. American Public Health Association, et al., Standard Methods for the Examination of Water and Wastewater, 16th Ed., Washington, D.C., 1985.

25. Personal Communication, Brad Reed, Senior Applications Engineer, Hoechst Celanese, 1990.

FINAL REPORT

1988-89 Mini Grant Program

Comparison of Asbestos Analysis by SEM-EDXA and TEM-SAED

Contract # F49620-88-C-0053/SB5881-0378

Submitted to

Universal Energy Systems, Inc.
4401 Dayton Xenia Road
Dayton, Ohio 45432-1894

Submitted by

Larry R. Sherman
Department of Chemistry
University of Scranton
Scranton, Pennsylvania 18510-2192

22-FEB-1990

ABSTRACT

Although some changes were required in the original protocol in order to complete the research work in this project, the objectives as outlined in the proposal were completed. Unfortunately, the study did not reveal a distinct advantage of SEM analyses over TEM analyses for asbestos as anticipated at the time the proposal was written. Several distinct characteristics were evident from the work: (1) TEM analysis shows a bias toward analyzing chrysotile fibers; (2) SEM shows a bias toward analyzing amosite fibers (amphiboles); (3) both methods have the same relative reliability in determining asbestos in a sample whose fiber count is greater than the action level (30 fibers/100 fields); (4) there is a small cost advantage in SEM over TEM instrumentation but considerably less than anticipated; (5) SEM may be an excellent method for providing information to AF safety engineers about an asbestos problem.

I. INTRODUCTION

The basic format laid out in the original proposal (dated October 4, 1988) was used in this work except for the following: forty five air samples instead of 20-30 were collected by OEHL/SA, Brooks Air Force Base, TX 78235 for TEM and SEM asbestos identification comparison. All samples had been previously analyzed by PLM and had an action level greater than 30 fibers/ 100 fields (PLM field).

In the original proposal, the TEM analyses was supposed to be performed by the P.I. and his graduate assistant; because of difficulties at the University of Scranton, the work was not performed as described. Furthermore, early in the study, the P.I. realized that he had insufficient training in TEM analyses to obtain "good" results. As such, the actual analyses were performed by a skilled TEM technician under the P.I.'s supervision using NIOSH procedure 7402.

Similar difficulties at OEHL/SA required a change in their protocol. The samples were not analyzed on the Brooks SEM prior to transmitting the samples to the P.I. Because of the lack of prior analyses, the P.I. traveled to Brooks AFB to personally analyze each SEM samples on the Amray 1820 SEM during January, 1990. Because of the change in format, i.e. the inability to perform two independent studies, the data was not transmitted to Mr. T. Thomas as blind results. All datum was in the hands of the P.I. throughout the work and some bias may have been introduced into the work (see section III-D)

II. PROCEDURES

All analyses were performed according to standard methods (1,2,3) which will not be repeated here.

Sample 116 was damaged in transit to Brooks and no SEM data was obtained for this sample. Ten SEM samples (#107,108,120,123,131,133,136,138,139 and 141) were rerun when initial evaluation indicated inconsistencies (see Table 1).

Because of cost restraints, only 40 of the 45 samples were subjected to TEM analysis. The five samples not run on the TEM were selected because preliminary PLM data indicated that they would show little or no asbestos. SEM analysis of one of these unsubmitted samples indicated amosite and chrysotile (sample #138) and chrysotile and anthopholite in another sample (#141). Three samples that were not run on the TEM (#129,132 and 135) indicated no asbestos. Two samples were not determined on the TEM (#120 and 134); sample 120 had too much interference from calcium silicate to make a valid identification of the fibers; the fibers in sample 134 were "packed" on the grid and could not be resolved.

III. RESULTS

A. Samples containing asbestos.

TEM analysis of 32/40 of the samples (80%) indicated either chrysotile or amosite fibers. Usually both fiber types were present. Normally all fibers on ten grid openings were evaluated for asbestos, this corresponds to a PLM field. Only 29/44 samples indicated asbestos by SEM; however, the number does not give a full story which will be critiqued in section III-C. Since

SEM does not concentrate the fibers in a sample and does not have a PLM equivalence, a greater scanning time was required for real samples than with a TEM.

Furthermore, because of the time restraint placed upon the P.I. in performing the SEM analysis, only four fibers were originally selected from each sample for qualitative analysis. More fibers were analyzed to correct for anomalies or when inconsistencies were present, e.g. sample 108 (see section V). The number of fibers given in Table 1 is the number for which the X-ray analysis was performed by TEM or SEM and not the total number investigated (the number is given in one column for each method). Many fibers, which would be counted by PLM evaluation, were rejected because it was evident to the P.I. that these fibers were not asbestos. Furthermore, when several fibers indicated the presence of amosite in the sample, many fibers which were probably amosite, were not analyzed to facilitate completing the work in the available time. If every fiber had been counted and tabulated, better quantitative comparison between TEM and SEM numbers might be possible. Unfortunately, this factor was not discerned until the data was being evaluated.

B. Number of asbestos fibers in the samples.

In the TEM analysis 2739 fibers were investigated and in the SEM work 222 fibers were studied. The difference in the number is more due to the time available for the work than any other factor. Almost three months were available for analyzing the TEM samples, but only 18 days were available for the SEM work. Some of the time available for the SEM study was required to reprogram the Amray and LAS systems. If the same amount of time were

available for both studies, a more comparable number of fibers would be in both data banks. Furthermore, a larger number of fibers were "passed over" in the SEM work than in the TEM study.

In the TEM study, 29.3% of the fibers were identified as chyrostile, 24.5% were amosite and 0% were other asbestos fibers. A total of 53.8% of the analyzed fibers were asbestos. In the SEM work, only 3.1% were identified as chrysotile, 56.7% as amosite, 2.25% as other asbestos fibers for a total of 62.0%.

In the TEM work, a large number of small fibers were read. Their dimensions were below 0.5 um in length and 0.2 um in diameter, which places them below the NIOSH limit for determination. The small fibers increased the total number of fibers read on the TEM but contributed little to sample understanding, i.e. which and how much of each asbestos type was present in the samples. Furthermore, as mentioned earlier, many SEM fibers were not given an elemental analysis because the P.I. knew they could not be asbestos and would not warrant the instrumental time. In both investigations, undue bias was given to the morphology of the fibers visible in the electron microsopy field and tended to yield a distorted result as discussed in section III-D.

C. Samples Showing No Asbestos.

In the TEM analyses, 8/40 samples yielded zero asbestos fibers. In the SEM analyses, 14/44 indicated no asbestos. However, three of the SEM samples which indicated zero were not analyzed on the TEM and direct comparison cannot be made, thus 12/41 of the duplicated analyzed samples tested zero and can be

used for evaluation.

SEM analysis of sample #120 indicated both amosite and chrysotile, whereas the TEM did not indicate any asbestos. Eight fibers were read on the SEM; one was chrysotile and another two fibers were identified as amosite. Twenty two fibers were read on the TEM, none indicated asbestos.

No asbestos fibers were found in Samples #123,127,130 and 139 by SEM, yet the TEM gave positive results. The PLM counts were 49.5, 82.5, 96 and 32.5 fibers respectively for these samples. TEM analysis of samples #127 and 130 indicated two and one chrysotile fibers respectively and no amosite in either sample. It is dubious that this low count has relevance and the results would depend upon where the samples were cut from the filter. Because of the question in comparing the relevance of the TEM and SEM for these low levels trial results, the samples were deleted from the final evaluation. The TEM analysis of sample #139 indicated nine chrysotile fibers and one amosite. As indicated in other sections of this report, TEM favors determination of chrysotile whereas SEM favors determination of amosite; thus the discrepancy between the systems are within the limitations of the methods for the samples. SEM analysis of sample 123 indicated no asbestos, although nine of the fibers with asbestos morphology were analyzed, none fit into the element ratio or cation/anion ratios demanded for the seven common types of asbestos. The TEM analysis indicated thirty chrysotile and one amosite fiber. This can only be deemed an analytical error.

If the low fiber count samples are eliminated from the evaluation, the study indicates that the TEM work failed to

identify asbestos in one sample (#12) and the SEM work had the same difficulty for a different sample (#123). For the TEM work 8/40 or 20% of the samples showed no asbestos; whereas after corrections 8/41 or 19.5% showed no asbestos in the SEM work. For all practical purposes these are the same numbers. Seven samples, which contained zero asbestos in both studies, were identical. If both sample #120 and 123 were deemed incorrect, it would appear as if the cross reference reliability for negative results is about 90% for both methods.

D. Comparison of Amphibole Analyses.

Because chrysotile contains a tubular morphology (see McCrone's Particle Atlas for examples), it is extremely easy to identify chrysotile fibers on a TEM grid. Because of the tubular characteristic, the counting and identification of chrysotile in the TEM study appears to be biased. Furthermore, chrysotile fibers are very small (there are always more fibers which were less than 0.5 um in length than fibers greater than 0.5 um). After the presence of chrysotile was assured in the sample, many of the fibers were deemed chrysotile because of this property without determining the crystal structures. 29.3% of the identified fibers were chrysotile and only 24.5% were amosite. No other asbestos fibers were identified.

On the other hand, the straight fiber property of amphiboles makes these fibers easy targets for identification in SEM work, 85% of the positively identified fibers were amosite, only 10% were chrysotile and 5% were other amphiboles. The latter is significant since it shows the only real variance in the two methods, three fibers were identified by SEM as crocidolite, one

as tremolite and one as anthopholite. It is possible that these were missed in the TEM analyses because of the sample selection; however, considering the bias which appears to exist in the two methods, it is highly likely that they were misidentified as amosite or totally ignored in the TEM assay process.

IV. STANDARDS

All analytical instruments must be standardized on a routine basis, the Amray 1820 and Tracor-Northern X-ray Analyzer are no exception. These instruments had not been standardized for asbestos analyses in the sixteen months since the P.I. had used them for the initial analyses. Although a complete restandardization was not made, a half day was devoted to reprograming the Amray SEM-EDXA system and another half day spent in correcting the computer library on the LAS system. Difficulty was also encountered with sample #108 and illustrated the type of problems that arise when a system is not maintained on a "ready" basis. None of the problems were formable, but they required about two and a half days of the P.I.'s time which could have been devoted to assay work.

Furthermore, if the method is used on a routine basis, it must be remembered that drift in the SEM, overlap of bands due to matrix effects (e.g. different types of cellulose acetate filters) and changes in manufacturer's software will probably require a monthly update of the computer programs. This is absolutely necessary in order to produce the confidence or assurance of the results required by OEHL.

V. LIBRARY

The Tracor-Northern library is capable of analyzing all elements using either K, L or M X-ray lines; however, it only produces elemental assays. The LAS computer has been programmed to accept TN data and determine the element/silicon and cation/anion ratio to help the analyst make a decision whether the analyzed fibers are asbestos or non-asbestos material. The LAS system contains a library for analyzing only the seven common asbestos fibers(3).

Since papers are beginning to appear which indicate that most mineral fibers which have the ERM criteria and, not only asbestos fibers, causes lung diseases, I feel it would be wise to expand the LAS library to include other common mineral fibers like fiberglass (silicon dioxide materials) and ceramic fibers (aluminum silicate materials). The latter materials were commonly found in the samples used in this study. They could not be identified by the TEM method used in this study but showed up in the SEM work. They were not observed in the original procedure set-up in 1988 because it was assumed that the aluminum was background from the SEM studs. This latter interference can now be avoided.

VI. PROBLEMS

A. Sample 108

TEM analysis of sample #108 indicated no asbestos fibers. Initial SEM analysis of sample #108 indicated the presence of ferroactinolite (FeAC), an asbestos fiber not normally found in air samples. The lack of TEM data to support the SEM data is not unusual considering the problems discussed in section III-D.

Sample #108 was rerun on the SEM when the discrepancy was realized. The rerun indicated no asbestos; however, there was doubt in the mind of the P.I. A standard FeAC sample was prepared and reentered into the LAS library. The new standard and suspected fiber were fingerprinted. When the two samples were compared side by side, it was evident that the suspect fiber was not FeAC. Although this is a standard analytical method for many analytical techniques, like IR, it is not as common in electron microscopy. The need for this type of verification of "real" samples increased the time required to perform a specimen analyses by about two hours.

From the current studies, I would guess that approximate 10-15% of an analyst's time will be needed for updating and restandardizing the equipment and computer program if a "ready" system is to be maintained.

B. Number of Fibers

If the method is used for routine analysis, the amount of time involved in preparing a sample and scanning the fields may not justify looking at all samples with only 30 fibers/PLM field. A count of 50/PLM field is probably a better figure. Even if every sample with only 30 fibers is investigated, a sliding scale of fiber analyses should be used to assure good results and reduce excessive analyst's time. The following scale is suggested: at thirty fibers/PLM field only four fibers need to be analyzed to determine whether the sample contains asbestos; at 100 fibers/PLM field, ten to 15 fibers need to be analyzed; over 200/PLM, about 20 fibers need to be analyzed. If the rejection

rate based on operator expertise is greater than 50% before analyzing four fibers, the number of asbestos fibers are probably too small to warrant further consideration. Negative samples should be reported but require no action.

VII. SUGGESTIONS FOR FUTURE USE OF THIS WORK.

A. Standard SEM Method

The wide discrepancy in the TEM versus SEM numbers is partly due to the lack of a standard method for searching, counting and analyzing fibers by SEM. Because this was a qualitative study, the SEM study was performed on random fibers, whereas the TEM was run on a fixed number of fibers on equivalent PLM fields. The SEM fields need to be standardized and marked and only those fibers in the field should be analyzed; every fiber in the marked field needs to be counted and determined if not by X-ray, at least by morphology. A more systematic method for reading fibers by SEM will produce better comparison of TEM/SEM quantitative numbers.

B. Use of the Method

Although the SEM method used here is not approved by EPA for asbestos analysis, the P.I. feels that it is a good method for routine analysis of air samples, not only at OEHL/SA but also in other laboratories. It is a viable method for monitoring an asbestos environment. Therefore I suggest that the following be considered: when a sample exceeds the action limit, a SEM analysis be performed on the sample to determine the approximate % of asbestos on the filter. This SEM results should be transmitted to the environmental engineer submitting the samples in order to assist him in making a decision on how to correct a

possible health hazard. The following is an example of a potential problem:

PLM analysis of sample #134 indicated 274 fibers/PLM field. TEM analysis indicated 37% of the fibers are either chrysotile or amosite and SEM analysis indicates that as many as 50% of the fibers could be asbestos. The P.I. has no knowledge about where or how the sample was collected, but the electron microscopy indicates that a potential health hazard exists. The base engineer should be made aware of this information. OEHL/SA is capable of performing this service and should do so.

VIII. PERSONNEL

A. SEM Technician.

From the experiences learned in this work, it is evident that an experienced SEM technician is required to perform the work requested in the previous sections. The technician needs expertise in determining which fibers should be analyzed or not analyzed and how to make a decision whether a sample should be rerun. Fingerprinting technique requires SEM-EDXA expertise and the procedure will often be needed for the exact asbestos fiber identification. Daily use of the instrument is imperative for personnel to remain familiar with the technique, correct for drift and library maintenance.

B. Time.

As detailed in an earlier report (6), rapid analysis is possible, but this study indicates it can lead to uncertainty in the results. A fiber can be analyzed in about 10 minutes under ideal conditions but more times is required to assure that the

answer has the degree of confidence expected of a laboratory like OEHL/SA. Eight to ten samples/day is the maximium that can be run; only half this number can be run when standardization or an update is necessary. Bugs still need to worked out of the system. They do not appear to be a serious drawback to using the method but will inhibit productivity during initial routine application.

IX. COSTS

The funds allotted for this project were spent according to the budget revision submitted on October 3, 1989. A detailed expenditure is given in the appendix.

Using a skilled TEM technician, the costs of the TEM analyses were reduced from approximately $500/sample to $200/sample. The SEM analyses cost $58/sample but this figure does not reflect overhead, equipment replacement, laboratory space, etc. When the latter costs are included in the SEM figure, the cost/sample will double.

Based on the original work (6) the SEM method should reduce the cost of asbestos monitoring by as much as 90%. However, when real samples were determined, as compared with the ideal samples used during the summer fellowship, the cost of performing an analysis was significantly higher. It is now estimated that when the same number of fibers are determined by the SEM as by a TEM, the cost savings is only about 25%.

X. CONCLUSIONS

No distinct advantage or disadvantage can be shown for either TEM or SEM from this study. When an experienced operator performs the asbestos analyses, the same degree of results can be

produced at a 95% confidence level. The margin of error is approximately the same as expected in most analytical work. Several differences exist for each method, TEM appears to possess a bias for reporting chrysotile whereas SEM appears to possess a bias for reporting amphiboles. The SEM methods may have a greater discriminating ability in identifying individual amphiboles, e.g. tremolite and crocidolite, but this may exist only because of the method used for selecting fibers in analysis in this study. Since a standard counting procedure does not exist for SEM, it is possible that the discrepancy in chrysotile versus amphibole results may disappear when a standard procedure is developed.

Originally, it was anticipated that the work would lead to a peer reviewed publication regardless of the results; however, two recent publications (4,5) have appeared on the comparison of TEM and SEM analyses. The recent papers and the work on this project seriously overlap each other and further evaluation of the data is required before a decision can be made as to whether this work has produced sufficient new information to warrant a publication.

XI. ACKNOWLEDGEMENTS

The P.I. wishes to acknowledge assistance by Ms. Jane Hubbard during the TEM work and by personnel at Brooks AFB during January, 1990 during the SEM work, expecially T. Thomas and K. Roberson. Without the technical and moral assistance of these people, the project could not have been completed.

XII. REFERENCES

1. NIOSH Method 7400 (PLM)

2. NIOSH Method 7402 (TEM)

3. L.R. Sherman, K.T. Roberson and T. Thomas, "Qualitative analysis of asbestos fibers in air, water and bulk samples using SEM-EDXA." J. Penna. Acad. Sci. (1989) 63: 28-33.

4. J.A. Helsen, P. Van De Velde, A. Kuczumow and A. Deruyttere, "Surface characteristics of asbestos fibers released from asbestos cement products." Am. Ind. Hyg. Assoc. (1989) 50:655-663.

5. P.N. Breyssee, J.W. Cherrie, J. Addison and J. Dodgson, "Evaluation of airborne asbestos concentration using TEM/SEM during residential water tank removal." Ann. Occup. Hyg. (1989) 33:243-256 (1989).

6. L.R. Sherman, "Determination of asbestos fibers in environmental samples using SEM-EDXA." 1988 USAF-UES Summer Faculty Research Program.

TABLE 1
Analyses of Samples

Project Sample Number	SC Number	PLM Fiber Counts	TEM Fiber Counts		SEM Fiber Counts		Other Asbestos
			Chrysotile	Amosite	Chrysotile	Amosite	
101	63915	237	29/125	58	0	3/7	
102	63916	30.5	0/10	0	0	0/4	
103	64174	123.5	0/10	0	0	0/4	
104	61175	71.5	0/10	0	0	0/4	
105	64362	385	71/107	10	0	3/4	
106	65716	213	71/157	21	0	2/5	
107	69081	366	0/14	0	0	0/4	
108	69082	387	0/10	0	0	0/8	
109	69810	207	25/120	16	0	1/4	
110	69811	121	8/198	11	0	4/4	
111	70462	294	29/99	18	0	3/4	
112	70463	58.5	0/35	19	0	2/4	
113	70464	622.5	0/115	81	0	3/4	
114	70465	552.5	0/123	72	0	3/5	
115	70951	46	36/135	19	0	1/5	
116	70952	65	27/113	33	lost sample **		
117	00076	51	1/42	3	0	2/4	1/4
118	00077	77	6/62	9	0	1/4	
119	00309	47	17/135	4	0	1/4	2/4
120	02884	302	0/?	0	1/8	2/8	
121	02888	131	79/112	8	0	2/4	
122	16280	206	5/105	57	0	1/5	
123	03168	49.5	15/61	2	0	0/9	
124	03169	43	15/89	9	1/4	2/4	
125	03170	85	0/11	0	0	0/4	
126	51664	400	12/106	53	0	0/4	1/4
127	52187	82.5	2/10	0	0	0/4	
128	52202	39.5	1/10	1	0	1/4	
129	52987	39	NIL	NIL	0	0/4	
130	57080	96	1/10	0	0	0/5	
131	54192	236	55/103	23	0	2/8	
132	55032	32	NIL	NIL	0	0/6	
133	56238	119	65/75	6	0	3/4	
134	57248	274	HIGH*	*	1/5	2/5	
135	57428	31	NIL	NIL	0	0/4	
136	57431	280	0/10	0	0	0/8	
137	57583	505	32/83	21	0	1/4	
138	58472	31	NIL	NIL	1/8	1/8	
139	58575	32.5	9/19	1	0	0/8	
140	58578	89	50/58	0	0	3/4	
141	59220	35.5	NIL	NIL	2/9	0/9	1/9
142	58672	113.5	7/61	22	0	3/4	
143	59232	56.5	80/102	4	0	2/4	
144	59543	35.5	32/59	10	0	2/4	
145	59545	39.5	25/35	1	1/5	0/5	

* Too many fibers on grid could not analyze.
** Sample was damaged in transit could not perform SEM analysis.
Nil indicates the sample was not subjected to TEM analysis.

TABLE II.

Summary of Asbestos Data

	TEM	SEM
Number of samples evaluated	38	44
Number which contained asbestos	32	30
Corrections for low values and duplicate samples	30	29
Zero samples	8	14
Correction for low values	10	11
Difference in sample after correction	1	1
% Chrysotile fibers	29.5%	3.1%
% Amosite fibers	24.5%	56.7%
% Other asbestos fibers	0%	2.2%
TOTAL fibers identified as asbestos	53.8%	62.0%

APPENDIX A

Expenditure of Funds

Salaries - P.I./Staff/Students	$2,891.59
Fringe benefits - F.I.C.A.	257.00
Professional Services - TEM analysis	8,000.00
Travel SA - SEM analysis	1,526.00
Laboratory supplies	88.56
Equipment - Anatech sample preparer	4,600.00
Consumable supplies - telephone, xerox, etc.	215.85
Overhead	2,421.00
TOTAL	$20,000.00

AN EXAMINATION OF KRIGING TECHNIQUES FOR GROUND WATER MONITORING

By
Gary Stevens
Department of Statistics
Oklahoma State University
Stillwater, OK 74048

AN EXAMINATION OF KRIGING TECHNIQUES FOR GROUND WATER MONITORING

By

Gary Stevens
Department of Statistics
Oklahoma State University
Stillwater, OK 74078
(405) 744-5684

I. INTRODUCTION

In many applications of Kriging in hydrologeological settings, the number of observations in the realization is small (Stevens(1988)). Therefore, the purpose of this paper is to study the small sample properties of Kriging. In particular, the small sample properties of the cross validation method for estimating the zone of influence parameter of the correlation parameter. These properties are examined by means of a controlled simulation study. The results of this simulation study are presented in section V. In section II, the basic ideas of Kriging are presented and in section III, these are modified for use in analyzing small samples. In section IV, the turning bands method for simulation of spatial processes is presented. This is the method used in this paper to simulate the spatial processes.

AN INTRODUCTION TO KRIGING

Spatial estimation has its roots deep in ore estimation problems. D.G. Krige, a South African mining engineer, in the early 1950's applied moving average techniques to the problem of ore estimation (David(1977), Journel(1978)). In the 1960's, G. Matheron, a French engineer, formalized the approach. Matheron (1971) introduced the concept of regionalized variables to represent functions whose values do not follow a deterministic function over some range of values. He developed a theory for spatially distributed random variables satisfying some basic assumptions. Matheron referred to his spatial estimation techniques as "Kriging", in honor of Krige. Kriging today has many applications in mining, meteorology, forestry, mapping of ground water potentials, air pollution concentrations, geologic surfaces, and oil field delineations.

Kriging is essentially a stochastic method of interpolating sparse data that is spatially correlated. The interpolated or estimated values are then used to estimate the contours of the spatial process. Kriging, as will be seen later, consists of basicly two areas. The first one is to estimate a measure of correlation between the spatial observations. This measure is generally a function of the distance between observations. The second area is to estimate the Kriging weights that are used in the prediction. These weights are a measure of the relative influence of each observation on the interpolations. These weights are dependent on the measure of correlation.

Generally, in statistics, it is assumed that there are n independent and identically distributed (iid) random variables each with density $f_X(x)$. Also assume that

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \qquad (1)$$

that is, the random variables are normally distributed. The theory of estimation and hypothesis testing has been extensively developed for this situation. Now suppose there is just have one observation X_1 from a normal population. What can be done ? The mean, μ, can be estimated, but the variance, σ^2 cannot. One cannot test hypotheses either. This is the situation that arrises in spatial estimation. That is, only one realization of a spatial process is observed from a collection of all possible realizations of the process. This collection of all possible realizations is called an ensemble.

Assume a realization $Z(\underline{s}_1), \ldots, Z(\underline{s}_n)$ of a spatial process at locations $\underline{s}_1, \ldots, \underline{s}_n, \underline{s} \in \mathbb{R}^m$ has been observed. Generally m=2 or 3. In order to make valid statistical inferences from a single realization of the process, several assumptions need to be made. In Kriging, these assumptions are

$$E(Z(\underline{s})) = \mu \qquad \forall \underline{s} \in \Omega \subset \mathbb{R}^m \qquad (2)$$

$$\mathrm{Var}(Z(\underline{s})-Z(\underline{s}+\underline{h})) = 2\gamma(\underline{h}). \qquad (3)$$

The quantity $2\gamma(\underline{h})$ is called the variogram ($\gamma(\underline{h})$ is called the semi-variogram). Note that $2\gamma(*)$ depends only the distance $\underline{h}$ between $Z(\underline{s})$ and $Z(\underline{s}+\underline{h})$ and not on the location $\underline{s}$. These two assumptions lead to a class of intrinsic stationary processes. This class of processes is broader than the class of second order stationary processes and strictly contains the class of second order stationary processes. A process is said to be second order

stationary if (2) holds, the covariance function exists and

$$Cov(Z(\underline{s}),Z(\underline{s}+\underline{h})) = C(\underline{h}). \qquad (4)$$

If C(*) exists then $\gamma(\underline{h}) = C(\underline{0})-C(\underline{h})$. This relationship can best be illustrated by the following figure.

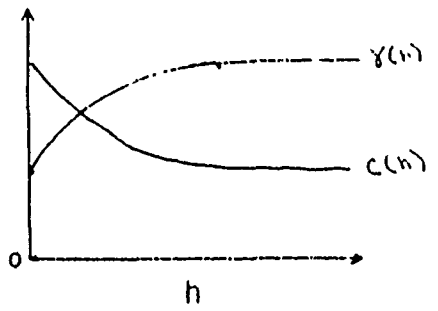


Figure 1. Relationship between covariance and variogram

If the variogram depends only on $|\underline{h}|$, that is, the norm of the distance, i.e. $\gamma(\underline{h}) = \gamma(|\underline{h}|)$, the variogram is said to be isotropic. If the variogram depends on both $|\underline{h}|$ and θ, the direction, the variogram is said to be anistropic. The variogram has the following properties:

P1: The variogram starts at 0, i.e. $\gamma(\underline{h}) = 0$

P2: The variogram is an even function, i.e. $\gamma(\underline{h}) = \gamma(-\underline{h}) > 0$

P3: The variogram must satisfy the conditional negative semi-definiteness condition, i.e.

$\sum_{i=1}^{k} \sum_{j=1}^{k} a_i a_j \gamma(\underline{s}_i-\underline{s}_j) \leq 0$ for any finite number of spatial locations $\{\underline{s}_i: i=1,\ldots,k\}$ and any real numbers $\{a_i: i=1,\ldots,k\}$ satisfying $\Sigma a_i=0$.

P4: When the variance exists, the variogram has a limit

i.e. $\lim_{\underline{h}\to\infty} \gamma(\underline{h}) = \mathrm{Var}(Z(\underline{s}))$

P5: The variogram increases faster than a parabola, i.e.

i.e. $\lim_{|\underline{h}|\to\infty} \frac{\gamma(\underline{h})}{|\underline{h}|^2} = 0.$

In practice. the variogram of the process is not known and therefore, it must be estimated. A nonparametric unbiased estimator of the variogram is

$$2\hat{\gamma}(\underline{h}) = \frac{1}{|N(\underline{h})|}\sum_{N(\underline{h})} (Z(\underline{s}_i)-Z(\underline{s}_j))^2 \qquad (5)$$

where the average is taken over $N(\underline{h}) = \{(\underline{s}_i,\underline{s}_j):\underline{s}_i-\underline{s}_j \le \pm\underline{h}\}$ and $|N(\underline{h})|$ is the number of distinct elements in $N(\underline{h})$. This estimator is influenced by extreme values of Z; therefore, Cressie(1985) has proposed the following robust estimator of the variogram

$$2\tilde{\gamma}(\underline{h})=\left(\frac{1}{|N(\underline{h})|}\sum_{N(\underline{h})} |Z(\underline{s}_i)-Z(\underline{s}_j)|^{1/2}\right)^4 *(.457+.494/|N(\underline{h})|)^{-1} \qquad (6)$$

Once the point estimate of the variogram has been calculated, a parametric model is fit to the estimated variogram. Some practical rules for calculating the variogram are :

R1: Calculate the variogram for distances $\underline{h}$ such that $|N(\underline{h})| > 30\text{-}50$.

R2: Calculate the variogram for $\underline{h}$ such that $|\underline{h}| < L/2$, where L is the largest distance.

R1 will help control the variation of he estimated variogram. R2 is practical since we are only interested in the correlation of observations that are close to each other.

Parametric estimation of the variogram is important since this estimate is used in obtaining the weights of the observations that will be used in Kriging. There are several models that are commonly used in spatial statistics. They are

M1: Linear model

$$\gamma(\underline{h})=c_o+b|\underline{h}|$$

M2: Exponential model

$$\gamma(\underline{h})=c_o+c_*[1-\exp(-|\underline{h}|/a)]$$

M3: Spherical model

$$\gamma(\underline{h})=c_o+c_*[(3/2)(|\underline{h}|/2)-(1/2)(|h|/a)^3] \qquad 0<|\underline{h}|\leq a$$

$$=c_o+c_* \qquad |\underline{h}|>a$$

M4: Power model

$$\gamma(\underline{h})=c_o+a|\underline{h}|^b \qquad 0\leq b<2$$

M5: Gaussian model

$$\gamma(\underline{h})=c_o+c_g(1-\exp(-|\underline{h}|^2/a^2))$$

M6: Logarithmic model

$$\gamma(\underline{h})=c_o+c_l\ln(|\underline{h}|).$$

In all the above models c_o is the nugget effect, i.e.

$$c_o = \lim_{|\underline{h}|\to 0} \gamma(\underline{h}).$$

c_o is the value of the variogram as the distance goes to zero. Note this is a contradiction to P1. This is brought about by "nuggets" or micro scale variation, which causes a discontinuity at the origin. It is easier thought of as a measurement error, that is, sampling the same point several times and not getting the same answers.

The value of a in the above models is called the range. It defines the lag at which $Z(\underline{s})$ and $Z(\underline{s}+\underline{a})$ are no longer correlated.

It is often thought of the parameter that allows one to determine which data should be included in the Kriging predictor, i.e. only data in the range of the predicted value need to be retained. As Cressie (1987) points out, this should be just a guideline not a hard and fast rule.

The nugget effect, range and another parameter called the sill are illustrated in Figure 2.

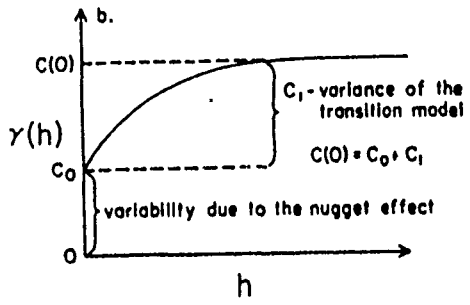


Figure 2. Variogram for spherical model

Most of the time the parameters are estimated by trial and error, cross validation, or eyeball methods. This is fine if one has good eyes and lots of practice. However, the statistical properties of these estimators are unknown and cannot be derived. Cressie (1985) derives generalized least squares and weighted least squares estimates of the parameters of the model. These types of estimators are consistent and asymptotically normal.

The problem in spatial statistics can be described as follows, given a set of spatial observations $Z(\underline{s}_1), \ldots, Z(\underline{s}_n)$ it is desired to predict the value of the process at some unknown

location $\underline{s}_0$, that is, predict $Z(\underline{s}_0)$. If it is assumed that the observations are normally distributed then the best predictor is $E(Z(\underline{s}) | Z(\underline{s}_1),\ldots,Z(\underline{s}_n))$. Since only a second order moment structure has been assumed, i.e, the mean and variance, then the predictors are restricted to linear predictors of the form

$$\hat{Z}(\underline{s}_0) = \sum_{i=1}^{n} \lambda_i Z(\underline{s}_i), \tag{7}$$

where the λ_i's are the Kriging weights assigned to the $Z(\underline{s}_i)$'s. It is desired to have $\hat{Z}(\underline{s}_0)$ be unbiased, that is, $E(\hat{Z}(\underline{s}_0) = \mu$. To insure that this is true, the following restrictions need to be put on the λ_i's

$$\sum_{i=1}^{n} \lambda_i = 1. \tag{8}$$

Also the predictor should be minimum variance. So the problem of Kriging can be stated as follows. Find

$$\hat{Z}(\underline{s}_0) = \sum_{i=1}^{n} \lambda_i Z(\underline{s}_i), \tag{9}$$

such that

$$\min_{\lambda} \; E(\hat{Z}(\underline{s}_0)-Z(\underline{s}_0))^2 \tag{10}$$

subject to

$$\sum_{i=1}^{n} \lambda_i = 1. \tag{11}$$

The λ's can be found by solving the following system of equations

$$\sum_{j=1}^{n} \lambda_j \gamma(\underline{s}_i-\underline{s}_j) + \beta = \gamma(\underline{s}_0-\underline{s}_i) \qquad i=1,\ldots,n \tag{12}$$

$$\sum_{i=1}^{n} \lambda_i = 1. \tag{13}$$

where β is the Lagrange multiplier. These can be written in matrix form as

$$\Gamma\underline{\lambda}=\underline{\gamma} \tag{14}$$

where $\underline{\lambda} = (\lambda_1,\lambda_2,\ldots,\lambda_n,\beta)'$, $\underline{\gamma} = (\gamma(\underline{s}_1-\underline{s}_0),\ldots,\gamma(\underline{s}_n-\underline{s}_0),1)'$ and

$$\Gamma = \begin{cases} \gamma(\underline{s}_i-\underline{s}_j) & i=1,\ldots,n \quad j=1,\ldots,n \\ 1 & i=n+1 \quad j=1,\ldots,n \\ 0 & i=n+1 \quad j=n+1 \end{cases} \qquad (15)$$

and Γ is a symmetric $(n+1)\times(n+1)$ matrix. The minimized mean square error is given by (Kriging variance)

$$\sigma_k^2(\underline{s}_0) = \sum_{i=1}^{n} \lambda_i \gamma(\underline{s}_i-\underline{s}_0) + \beta. \qquad (16)$$

This is Kriging for the simple ideal conditions and constant mean and isotropic variogram.

Often in statistics the ideal assumptions are not always satisfied. In spatial statistics the assumption of stationary in the mean $E(Z(\underline{s})) = \mu$ is not always satisfied. Often in fact,

$$E(Z(\underline{s})) = d(\underline{s}), \qquad (17)$$

where d(*) is called the drift function. Non stationarity in the means creates major problems in estimation of the variogram and therefore the Kriging predictors. One problem is that the estimates of the variogram $\gamma(\underline{h})$ given by 5) and 6) are no longer unbiased. Several ways of Kriging in the presence of drift have been proposed.

The first method is called universal Kriging (Matheron(1969)) and is an extension of ordinary Kriging. Here it is assumed that d(*) can be represented by a polynomial of finite order. This leads to modeling $Z(\underline{s})$ as

$$Z(\underline{s}) = \sum_{m=1}^{n} a_m f_m(\underline{s}) + \varepsilon(\underline{s}), \qquad (18)$$

where $\{f_m(\underline{s}); m=0,\ldots,k\}$ are known deterministic functions whose linear combinations account for the spatial drift and the $\{\varepsilon(\underline{s}), \underline{s}\in\mathbb{R}^n\}$ is a stochastic "error" process which satisfies equations 1) and 2) with $\mu=0$. Universal Kriging predicts $Z(\underline{s}_0)$ by

a linear combination of $Z(\underline{s}_1), \ldots, Z(\underline{s}_n)$, i.e

$$\hat{Z}(\underline{s}_o) = \sum_{i=1}^{n} \lambda_i Z(\underline{s}_i), \tag{19}$$

Suppose $f(\underline{s})$ is a polynomial of order l $(l=0,\ldots,k)$. Then the unbiasedness condition becomes

$$E\left[\sum_{i=1}^{n} \lambda_i Z(\underline{s}_i) \right] = \sum_{l=0}^{k} a_l f_l(\underline{s}_o). \tag{20}$$

A sufficient condition for this is

$$\sum_{i=1}^{n} \lambda_i f_l(\underline{s}_i) = f_l(\underline{s}_o) \qquad l=0,1,\ldots,k. \tag{21}$$

There are now (k+1) Lagrange multipliers $\beta_o,\ldots,\beta_k$. This yields the following set of equations

$$\Gamma_u \underset{\sim u}{\lambda} = \underset{\sim u}{\gamma}. \tag{22}$$

These look just like the ordinary Kriging equations where now

$$\underset{\sim u}{\lambda} = (\lambda_1,\ldots,\lambda_n,\beta_o,\ldots,\beta_k)' \tag{23}$$

$$\underset{\sim u}{\gamma} = (\gamma(\underline{s}_1-\underline{s}_o),\ldots,\gamma(\underline{s}_n-\underline{s}_o),f_o(\underline{s}_o),\ldots,f_k(\underline{s}_o))' \tag{24}$$

and

$$\Gamma_u = \begin{cases} \gamma(\underline{s}_i-\underline{s}) & i=1,\ldots,n \quad j=1,\ldots,n \\ f_{i-n-1}(\underline{s}_j) & i=n+1,\ldots,n+k+1 \quad j=1,\ldots,n \\ 0 & i=n+1,\ldots,n+k+1 \quad j=n+1,\ldots,n+k+1 \end{cases} \tag{25}$$

where Γ_u is a symmetric (n+k+1 × n+k+1) matrix.

There are two basic problems associated with universal Kriging. The first is k, the order of the polynomial is never known and has to be guessed. The second is that $\gamma(*)$ is never known and must be estimated from the residuals. The residuals are the data - guessed drift. Matheron(1971) pointed out that these estimates are biased and unless the variogram of the errors is known or calculated in a direction where there is no drift, Universal Kriging is hard to apply.

The second method involves the use of intrinsic random

functions of order k (Matheron (1973)). The basic idea behind this approach is that low-order generalized increments of the variables filters out the polynomials in the expectation. (This is just the n-dimensional analog of the Box-Jenkins approach to time series analysis.) This also allows $\{\varepsilon(\underline{s});\underline{s}\in\mathbb{R}^n\}$ to follow a more general class of stochastic processes than the intrinsic stationary class. If a generalized increment of order k is used on the data, Kriging is still possible when the stochastic process model is a intrinsic random function of order k. An intrinsic random function is defined below. For clarity and ease of notation we will restrict ourselves to $\mathbb{R}^2$. Let $\underline{s}=(x,y)$ and $\underline{s}_i=(x_i,y_i)$; i=1,...,n. Let X be the n × (p+1) regressor matrix corresponding to a polynomial of order k. For example

$$k=0 \quad X=\{1\} = \begin{bmatrix} 1 \\ 1 \\ . \\ . \\ . \\ 1 \end{bmatrix} \qquad k=1 \quad X=\{1,x_i,y_i\} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ . & . & . \\ . & . & . \\ . & . & . \\ 1 & x_n & y_n \end{bmatrix}$$

$$k=2 \quad X=\{1,x_i,y_i,x_i^2,x_iy_i,y_i^2\} = \begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & x_1y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2y_2 & y_2^2 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & x_n & y_n & x_n^2 & x_ny_n & y_n^2 \end{bmatrix}$$

Assume that X is of full column rank. Matheron (1973) defines the random variable $\underset{\sim}{\nu}'Z$ to be a generalized increment of order k if the coefficients $\underset{\sim}{\nu}$ satisfy $X'\underset{\sim}{\nu}=0$.

An intrinsic random function of order k (IRF-k) is defined to be any process $\{Z(\underline{s});\underline{s}\in\mathbb{R}^2\}$ for which

$$H^{\nu}(\underline{u}) \equiv \sum_{i=1}^{n} \lambda_i Z(\underline{s}_i+\underline{u}) \qquad \underline{u}\in\mathbb{R}^2 \tag{27}$$

is such that both $E(H^{\nu}(\underline{u}))$ and $COV(H^{\nu}(\underline{u}),H^{\nu}(\underline{u}+\underline{v}))$ do not depend on $\underline{u}$ for any $\underline{s}_i$ and any generalized increment vector $\underset{\sim}{\nu}$ of order k. Matheron (1973) proves that an IRF-k possesses a generalized covariance function $\{K(\underline{h});\underline{h} \in \mathbb{R}^2\}$ such that for data $\underline{Z}$ and generalized increment vectors $\underset{\sim}{\nu}$ and $\underset{\sim}{\nu}^*$ of order k

$$COV(\underset{\sim}{\nu}'\underline{Z},\underset{\sim}{\nu}^{*\prime}\underline{Z}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \nu_i \nu_j^* K(\underline{s}_i-\underline{s}_j). \tag{28}$$

The generalized covariance is unique up to a polynomial of order k and satisfies $K(-\underline{h})=K(\underline{h})$.

Kriging can be done using $K(\underline{h})$ instead of $\gamma(\underline{h})$. Note for an IRF-0 $K(\underline{h})=\gamma(\underline{h})$. There are various of classes of functions that satisfy the conditions of a generalized covariance, but one class with nice properties for identification and estimation is the class of polynomial generalized covariances. There are other classes of covariances which are also useful (Cressie (1987)). Although the mathematics for this type of process is well developed, practically the approach described above has several problems. The first is that, in general, it is not possible to estimate the generalized covariance nonparametrically and there is no easy to interpret the parameters of the generalized covariance. Secondly, the order of the IRF-k is not known and must be guessed. Cressie(1987) has proposed a nonparametric graphical method for helping to determine the order. The third problem is that the edge effects can lead to a big reduction in the number of points that can be Kriged in this way.

A third way of handling the drift in the means proposed by Cressie (1986), is called median polish. Median polish eliminates the drift and ameliorates the bias problem. To do median polish one needs data that is collected on a regular or near regular grid. One successively subtracts the median from each row and column until the median for each row and column is zero. Then one can perform Kriging on the residuals of median polish.

The final method of correcting for drift is to estimate the drift and use a drift corrected variogram. One can estimate the drift by

$$\hat{d}(\underline{h}) = \frac{1}{|N(\underline{h})|} \sum_{N(\underline{h})} (Z(\underline{s}_i+\underline{h})-Z(\underline{s}_i)). \qquad (29)$$

and the drift corrected variogram by

$$\hat{\gamma}^*(\underline{h}) = \frac{1}{|N(\underline{h})|} \sum_{N(\underline{h})} (Z(\underline{s}_i+\underline{h})-Z(\underline{s}_i)-\hat{d}(\underline{h}))^2. \qquad (30)$$

This approach is simpler and easy to use but still suffers from the bias is the estimator.

PRACTICE OF KRIGING WITH A SMALL NUMBER OF OBSERVATIONS

In the practice of Kriging it is often the situation that only a small realization of the process is observed, that is n<50. In these situations the experimental point variogram often exhibits erratic behavior and is hard to estimate parametrically. Therefore, an alternative approach must be taken. In order to facilitate this, it is assumed that the mean of the process is a constant and that the covariance function, C(h), exists, i.e.

$$C(h)=E[(Z(\underline{s})-\mu)(Z(\underline{s}+\underline{h})-\mu)]. \qquad (31)$$

With these assumptions, the Kriging problem can now be formulated in terms of linear models.

Let $\underline{Z}$ be a vector of n random variables with mean $\underline{\mu}$ and variance - covariance matrix V. That is,

$$\underline{Z} \sim (\underline{\mu}, V) \qquad (32)$$

$\underline{Z}$ can now be partitioned into two parts, $\underline{Z}_1$ of dimension n-1 and $\underline{Z}_2$ of dimension 1. Now.

$$\underline{Z} = \begin{bmatrix} \underline{Z}_1 \\ Z_2 \end{bmatrix} \sim \left(\begin{bmatrix} \underline{\mu}_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{bmatrix} \right) \qquad (33)$$

where V_{11}, V_{12}, and V_{22} are the appropriate partitions of V. The problem of Kriging is now to predict Z_2 given a set of observations $\underline{Z}_1=\underline{z}_1$. That is, need to estimate $E(Z_2|\underline{Z}_1=\underline{z}_1)$. It is known that

$$E(Z_2|\underline{Z}_1=\underline{z}_1) = \mu + V_{12}' V_{11}^{-1}(\underline{z}_1-\underline{\mu}_1). \qquad (34)$$

Since the estimator of $E(Z_2|\underline{Z}_1=\underline{z}_1)$ is restricted to a linear, then it is known that the best linear unbiased estimator (BLUE) is the least squares estimator given by

$$\hat{Z}_2 = \hat{\mu} + v'_{12}+V_{11}^{-1}(\underline{z}_1-\hat{\mu}) \qquad (35)$$

where

$$\hat{\mu} = (\underline{J}' V_{11}^{-1}\underline{J})^{-1}\underline{J}' V_{11}^{-1}\underline{z}_1. \qquad (36)$$

and $\underline{J}$ is a n×1 vector of 1's. Here also it is assumed that V is known.

This looks than the Kriging equations from before. However, it can be shown that when the covariance exists then the Kriging equations reduce to the above.

Estimation of the Covariance Function

The inherent assumption in this problem is that the covariance between $Z(\underline{s}+\underline{h})$ and $Z(\underline{s})$ depends only on $h=|\underline{h}|$, the distance between the two points. Since there are a small number of points and the sample point covariance estimate has poor structure, a simple correlation function will be fit to the data. Some common correlation functions are given in Table 1. The parameter b^{-1} is called the range, correlation zone, or zone of influence.

There are several methods to estimate b. The method of cross validation has a lot of appeal in this situation, because of the limited amount of data. Cross validation is a sequential procedure whereby each datum in turn is removed from the set of observations and it's value is then predicted by the model based

on the remaining observations. Then we choose the model that yields the best fit by some "goodness of fit" criteria. Again there are many such criterion, the one used in this study is the minimum residual variance criterion. The residuals are defined as $\hat{\varepsilon}(\underline{s}) = (\hat{Z}(\underline{s})-Z(\underline{s}))$. The model chosen is the one that minimizes

$$Q = (1/n) \sum_{\underline{s}} (\hat{\varepsilon}(\underline{s})-\bar{\varepsilon})^2 \qquad (37)$$

where $\bar{\varepsilon}$ is the mean of the residuals.

TURNING BANDS METHOD FOR SIMULATION OF A SPATIAL PROCESS

The turning bands method of multidimensional simulation is a method for simulation of spatial processes first introduced by Matheron (1973). The method of based on the theory of random functions. The basic concept is to transform the multidimensional simulation into a sum of a series of unidimensional simulations, while preserving the true characteristics of the field.

The idea of a random field is a generalization of the idea of a random variable. Let $\underline{x} = (x_1, x_2, \ldots, x_n)$ be a point in $\mathbb{R}^n$ and $Z(\underline{x})$ be a random variable corresponding to the point $\underline{x}$. Then a random field in $\mathbb{R}^n$ is defined to be the set $\{(\underline{x}, Z(\underline{x})); \underline{x} \in \mathbb{R}^n\}$.

Define the covariance function $C(\underline{x})$ to be

$$C(\underline{x}_1, \underline{x}_2) = E[(Z(\underline{x}_1) - E(Z(\underline{x}_1))(Z(\underline{x}_2) - E(Z(\underline{x}_2))] \quad (38)$$

A random field is said to be second order stationary if

$$E(Z(\underline{x}) = \mu \qquad \forall \underline{x} \in \mathbb{R}^n \quad (39)$$

and

$$C(\underline{x}_1, \underline{x}_2) = C(\underline{x}_1 - \underline{x}_2) = C(\underline{h}), \quad (40)$$

that is, the covariance of any two points depends only on the distance between the points. A second order stationary process is said to be isotropic if

$$C(\underline{h}) = C(r) \quad (41)$$

where $r = |\underline{h}|$. That is, the covariance function does not depend on the direction $\underline{h} = \underline{x}_1 - \underline{x}_2$ of the distance vector, but only the vector length $|\underline{h}|$. Some common covariance functions are presented in Table 1. Since it has been noted by Manteglou and Wilson(1982) and Krajewski and Duffy(1988) that for small sample sizes it is

almost impossible to distinguish between the exponential and the Bessel covariance functions, the focus of this study will be on the exponential and Gaussian covariance functions. The radial spectral density will be defined later. The parameter b^{-1} is called the correlation length. Note this is just the zone of influence parameter from the previous section.

Covariance Model	Covariance Function	Radial Spectral Density
Exponential	$\sigma^2 \exp(-br)$	$\dfrac{\omega/b}{b[1+(\omega/b)^2]^{3/2}}$
Gaussian	$\sigma^2 \exp(-b^2 r^2)$	$(1/2b)(w/b)\exp[-(\omega/2b)^2]$
Bessel	$\sigma^2 br K_1(br)$	$\dfrac{2(\omega/b)}{b[1+(\omega/b)^2]^2}$

Table 1. Two dimensional covariance functions and radial spectral densities.
(note: K_1 is a modified Bessel function of the first kind of order zero.)

The Turning Bands Method

It will be assumed that the spatial process to be simulated is second order stationary and isotropic. It is also assumed the process is a zero mean Gaussian process. The turning bands method does not directly simulate the two or three dimensional fields directly but instead performs simulations along several lines using a unidimensional covariance function corresponding to the given two or three dimensional covariance function. The value of the simulated field at each point is just a weighted sum of

corresponding values of the line process.

Let $\mathbb{F}$ represent the two or three dimensional spatial process to be simulated. Choose an arbitrary origin in $\mathbb{R}^n$ (n=2 or 3) and generate lines so that the direction vectors $\underline{u}$ are uniformly distributed on the unit circle (two dimensional) or the unit sphere (three dimensional). Along each line i, generate a second order stationary unidimensional process having zero mean and covariance function $C_i(\xi)$, where ξ is the coordinate on a line i (the relationship between the unidimensional covariance function C_i and the covariance function of the field C will be derived later). gonally project onto line i, those points from $\mathbb{F}$, where values are to be simulated and assign to them the corresponding values from the one dimensional process. Let P be the point in $\mathbb{F}$ having location vector $\underline{x}_p$, then the assigned value from line i will be $z_i(\xi_{pi})$, where $\xi_{pi} = \underline{x}_p \cdot \underline{u}_i$ is the projection of $\underline{x}_p$ onto line i. Take L such lines as i, then for every point P in $\mathbb{F}$ there are L values of $z_i(\xi_{pi})$, for i=1,...,L. Then the simulated value of the process assigned to point P is

$$z_s(\underline{x}_p) = \frac{1}{\sqrt{L}} \sum_{i=1}^{L} z_i(\underline{x}_p \cdot \underline{u}_i) \qquad (42)$$

where the subscript s means simulated.

The field given by (42) has zero mean. However, it is desired to the form of the unidimensional covariance function $C_i(\xi)$ that yields the desired covariance function $C(\xi)$ for the field. Let $Z(\underline{x}_1), Z(\underline{x}_2)$ be two simulated values from $\mathbb{F}$, then the covariance function $C_s(\underline{x}_1, \underline{x}_2)$ is given by

$$C_s(\underline{x}_1,\underline{x}_2) = E[Z(\underline{x}_1)Z(\underline{x}_2)]$$

$$=(1/L)\sum_{i=1}^{L}\sum_{j=1}^{L} E[Z_i(\underline{x}_1\cdot\underline{u}_i)Z_j(\underline{x}_2\cdot\underline{u}_j)]. \qquad (43)$$

Since the realizations along two different lines are independent the expected value is zero unless i=j. Then (43) reduces to

$$C_s(\underline{x}_1,\underline{x}_2) = (1/L)\sum_{i=1}^{L} E[Z_i(\underline{x}_1\cdot\underline{u}_i)Z_i(\underline{x}_2\cdot\underline{u}_i)]$$

$$= (1/L)\sum_{i=1}^{L} C_i(\underline{h}\cdot\underline{u}_i) \qquad (44)$$

where $\underline{h} = \underline{x}_1-\underline{x}_2$.

Letting $L \longrightarrow \infty$ and using the law of large numbers

$$C_s(\underline{h}) = \lim_{L\to\infty}\left\{\frac{1}{L}\sum_{i=1}^{L} C_i(\underline{h}\cdot\underline{u}_i)\right\} = E(C_i(\underline{h}\cdot\underline{u}_i)$$

$$= \int_C C_i(\underline{h}\cdot\underline{u})\ f(\underline{u})\ d(\underline{u}). \qquad (45)$$

where C is the unit circle or sphere, $f(\underline{u})$ is the density function of $\underline{u}$ which is $1/2\pi$ (two dimension) or $1/4\pi$ (three dimensions). Since the mathematics for the two and three dimensional cases are different, they will be studied separately.

Three Dimensional Fields

The simulated covariance function $C_s(r)$ given by equation (45) can be written as

$$C_s(r) = (1/4\pi)\ \int_{\text{unit sphere}} C_i(\underline{h}\cdot\underline{u})d\underline{u}. \qquad (46)$$

Transforming to polar coordinates, (46) can be written as

$$C_s(r) = (1/4\pi)\int_0^{2\pi}\int_0^{\pi} C_i(r\cdot\cos\phi)\sin\phi\ d\phi\ d\Theta. \qquad (47)$$

Letting $\xi = r\cos\phi$ (47) can be written as

$$C_s(\xi) = 1/r\int_0^{r} C_i(\xi)d\xi \qquad (48)$$

From (48) we get the relationship between the one and two dimensional covariances as

$$C_1(\xi) = -\frac{d}{d\xi}\left[\xi C(\xi)\right] \tag{49}$$

Thus the univariate process can easily be simulated using standard statistical techniques. This nice relationship, however, does not hold true for two dimensional processes.

Two Dimensional Processes

For two dimensional fields, the simulated covariance function $C_s(r)$ given by (45) can be written as

$$C_s(r) = (1/2\pi) \int_{\text{unit circle}} C_1(\underline{h}\cdot\underline{u})d\underline{u} \tag{50}$$

Transforming to polar coordinates (50) becomes

$$C_s(r) = (1/2\pi) \int_0^{2\pi} C_1(r \sin\theta)d\theta \tag{51}$$

where $\underline{h}\cdot\underline{u} = r \sin\theta$ and $d\underline{u} = d\theta$. Letting $\xi = r \sin\theta$, (51) becomes

$$C(r) = (2/\pi) \int_0^r \frac{C_1(\xi)}{(r^2-\xi^2)^{1/2}}\, d\xi. \tag{52}$$

This leads to the relationship between one and two dimensional covariances as

$$\int_0^r \frac{C_1(\xi)}{(r^2-\xi^2)^{1/2}}\, d\xi = (\pi/2)C(r). \tag{53}$$

This equation is not easily solved, and although certain solutions have been found, they do not have useful meanings for hydrogeological applications.

For this reason, a more preferred method is to use a spectral method to generate the line process and still preserve the two dimensional covariance structure.

Spectral Method for Generating Two Dimensional Fields

Let the covariance function ,$C(\underline{h})$, of the two dimensional process be continuous and tend to zero fast enough as $|\underline{h}| \longrightarrow \infty$. Then $C(\underline{h})$ can be expressed as a Fourier integral given by

$$C(\underline{h}) = \int_{\mathbb{R}^2} e^{i\underline{h}\cdot\underline{\omega}} S(\underline{\omega}) d\underline{\omega} \qquad (54)$$

where $i = \sqrt{-1}$, $\underline{\omega} = [\omega_1, \omega_2]$ a vector of frequencies and $\underline{h}\cdot\underline{\omega}$ is the inner product of the vectors $\underline{h}$ and $\underline{\omega}$. The function $S(\underline{\omega})$ is the spectral density function of the two dimensional process and is given by the Fourier transform of $C(\underline{h})$;

$$S(\underline{\omega}) = 1/(4\pi) \int_{\mathbb{R}^2} e^{-i\underline{h}\cdot\underline{\omega}} C(\underline{h}) d\underline{h}. \qquad (55)$$

If the field is isotropic then $S(\underline{\omega})=S(\omega)$ where $\omega=|\underline{\omega}|$. Then (54) and (55) become (Shoenberg(1938))

$$C(r) = \sigma^2 \int_0^\infty f(\omega) J_0(\omega r) d\omega \qquad (56)$$

$$f(\omega) = (\omega/\sigma^2) \int_0^\infty C(r) J_0(\omega r) dr \qquad (57)$$

where $r=|\underline{h}|$ and $J_0(*)$ is a Bessel function of the first kind of order 0 and $f(\omega)$ is the radial spectral density defined by

$$f(\omega) = \frac{2\pi\omega S(\omega)}{\sigma^2}. \qquad (58)$$

If the covariance $C(\underline{h})$ is known, the (57) can be used to calculate the corresponding radial spectral density (common ones are given Table 1.).

The next problem is to determine the univariate spectral density $S_1(\omega)$ that yields the two dimensional spectral density $S(\omega)$. From Monteglou and Wilson (1982) we get

$$S_1(\omega) = (\sigma^2/2)\, f(\omega). \qquad (59)$$

Given the spectral density of the unidimensional process, the process can easily be simulated along the turning bands lines using the spectral method of Shinozuka and Jan (1972). If the unidimensional covariance function is $C_i(\xi)$ and the corresponding spectral density function is $S_i(\omega)$ then the unidimensional process on line i can be generated by

$$z_i(\xi) = 2 \sum_{k=1}^{M} [S_i(\omega_k)\Delta\omega]^{1/2} \cos(\omega'_k \xi + \phi_k) \qquad (60)$$

where the ϕ_k are independent random angles uniformly distributed between 0 and 2π, $\omega_k = (k-1/2)\Delta\omega$ and $\omega'_k = \omega_k + \delta\omega$ for $k=1,\ldots,M$. It is assumes that $S_i(\omega)$ is insignificant outside $(-\Omega,\Omega)$. The discretization frequency $\Delta\omega$ is defined as $\Delta\omega = \Omega/M$, where M is the number of harmonics used. The frequency $\delta\omega$ is a small random frequency added to avoid periodicities. Shinozuka and Jan (1972) prove that the process given by (50) has zero mean and covariance function $C_i(\xi)$, as $M \longrightarrow \infty$, $\Omega \longrightarrow \infty$ and $\Delta\omega \longrightarrow 0$. The process is also ergodic and Gaussian.

This method of generating spatial process has been studied by Monteglou and Wilson (1981).

V. SIMULATION STUDY AND RESULTS

The purpose of this simulation study is to examine the small sample properties of Kriging. Since Kriging depends directly on the correlation function, the properties of Kriging depend directly on the properties of the estimation of the correlation function. To that end, this simulation study investigates the small sample properties of the cross validation estimation method. The cross validation method is used to estimate the parameter of the correlation function. This method is studied in relationship to the value of the parameter, sample size, and error variance.

The spatial processes were generated using the turning bands method given in section IV, with the exponential correlation function. The coordinates were generated randomly on a 10x10 grid. Although the restriction on the grid size is not necessary for the turning bands method, it gives a frame of reference for making inference (Masry(1971)).

The values of the correlation parameter were 1,5 and 9, the sample sizes were 10,15, and 20, and the values of the error variance were 1,5 and 9. All possible combinations of the parameter values, sample size, and error variance were simulated 1000 times. The simulations were run as a batch job on a VAX 1170 at Oklahoma State University and took 60 days to run. The mean and variance of the 1000 estimates of the correlation parameter are given table 3.

B

σ^2	1	5	9
1	2.78 10.35	5.60 16.84	6.42 16.23
5	2.98 11.62	5.80 16.58	6.37 16.21
9	2.96 11.71	5.68 16.42	6.35 16.59

N=10

B

σ^2	1	5	9
1	2.47 10.01	5.29 16.55	6.60 16.08
5	2.52 11.10	5.59 16.50	6.49 16.12
9	2.65 11.55	5.62 16.46	6.41 16.49

N=15

B

σ^2	1	5	9
1	2.03 9.25	5.15 15.97	7.25 15.35
5	2.20 10.50	5.34 16.24	7.01 15.71
9	2.35 10.95	5.45 16.35	6.87 15.84

N=20

Table 3. Means and Variances for the cross validation estimation of the range parameter

From the tables, it can been seen that when the true value of B is small relative to the grid size (B=1) the cross validation estimation method tended to over estimate it. When the true value of B was equal to five, the estimated values of B were close to the true value, and when the true value of B was large relative to the grid size (B=9) the true value was under estimated. Also as the variance of the process increased the estimated values of B deviated more from the true values and the variance of the estimates also increased. As expected as the sample size increased the estimated values of B approached the true value of B and the variance of the estimates decreased.

It appears from the simulation study that when the true value of the parameter is less than half the grid size that the Kriging procedure is robust with respect to estimation of the correlation parameter, and thus robust with respect to prediction of unsampled locations.

More work needs to be done with respect to other correlation functions and a wider range of parameter values. Due to the length of time for the simulations and the short length of this grant, these aspects were not able to be investigated.

Finally, it is the opinion of this author, based on the results of this simulation study, that Kriging is a robust procedure and performs extremely well when the realization size is small.

REFERENCES

Clark, G.P. and Dane, J.H. (1988) "Experiences with Kriging in Soil Sciences",Auburn University Department of Agronomy, Technical Report.

Cooper, R.M. and Istok, J.D. (1988) "Geostatistics Applied to Groundwater Contamination, I: Methodology", Journal of Enviromental Engineering,114,2,270-284.

Cooper, R.M. and Istok, J.D. (1988) "Geostatistics Applied to Groundwater Contamination, II: Application", Journal of Enviromental Engineering,114,2,287-299.

Cressie,N.A.C. (1984) "Fitting Variograms by Weighted Least Squares", Mathematical Geology,17,5,563-586.

Cressie,N.A.C. (1986) "Kriging Nonstationary Data",Journal of the American Statistical Society,81,625-634.

Cressie,N.A.C. (1987) "Spatial Statistics and Ordinary Kriging", Tec. Rep.#87-22,Statistical Laboratory,Iowa State University.

David,M, (1977) Geostatistical Ore Reserve Estimation, Elsevier, Amsterdam.

Journal,A. and Huijbregts,C. (1978) Mining Geostatistics: London Academic Press.

Krajewski,W.F. and Duffy,C.J. (1988)"Estimation of Correlation Structures for a Homogenous Isotropic Random Field: A Simulation Study",Computers and Geosciences,14,113-122.

Manteglou,A. and Wilson,S.L (1982)"Turning Bands Method for Simulation of Random Fields, Using Line Generation by Spectral Methods",Water Resources Research,18,1379-1394.

Masry,E. (1971)"Random Sampling and Reconstruction of Spectra", Information and Control, 19,275-288.

Matheron,G. (1973)"Intrinsic Random Functions and Their Applications", Advances in Applied Probability,5,439-468.

Myers,D.E. et al (1982)"Variogram Models for Regional Ground Water Geochemical Data",Mathematical Geology,14,629-644.

Stevens,G. (1988)"Analysis of Contaminated Ground Water Using Kriging Methods", AFOSR/UES Final Report.

Shinozuka,M. and Jan,C.M. (1972)"Digital Simulation of Random Processes and its Application",Journal of Sound Vibration, 25,113-122.

Shoenberg,I.J. (1938)"Metric Spaces and Completely Monotone Functions",Annals of Mathematics,39,811-840.

Zirschly,J.H. and Harris,D.J. (1986)"Geostatistical Analysis of Hazardous Waste Data",Journal of Enviromental Engineering, 112,770-785.

CORTISOL PREVENTION OF CHRONIC BERYLLIUM DISEASE IN POSTPARTUM RATS: A PILOT STUDY

Submitted by

Shirley A. Williams-Scott, Ph.D.

Stillman College, Tuscaloosa, Alabama 35403

Research sponsored under Grant #760-6MG-078 by the Universal Energy Systems, Inc., and the Air Force Office of Scientific Researh

CORTISOL PREVENTION OF CHRONIC BERYLLIUM DISEASE IN POSTPARTUM RATS: A PILOT STUDY

ABSTRACT

This study was designed to investigate the possible effect of cortisol on the prevention of the onset of chronic beryllium disease following parturition and the effect of beryllium exposure (300ug/kg) on reproductive parameters of the female rat, e.g, litter number, litter weight, and general appearance of the litter. The female rats were grouped according to the treatment plan: Group 1--normal impregnated rats (N), Group 2--beryllium exposed impregnated rats (B-1), and Group 3--beryllium exposed impregnated rats treated postpartum with 25mg/kg of hydrocortisone (B-2), at 3 day intervals for 28 days. There was no significant difference, ($P<0.05$), between the weight gain and overall health status of the female rats in each treatment group. In addition, the litter size, weight and general health status were the same for all groups of rats. Although there were some individuals who failed to deliver pups after cervical plugs were noted, and cyanosis was present in some pups of the beryllium exposed groups of animals, no conclusive evidence could be noted to associate beryllium as a causative agent. Gross examination, of the lung tissue of the beryllium treated rats, noted white nodular areas dispersed on the surface of the lung tissue, when compared to the normal group. Because of technical difficulties in the storage of the lung tissue, histological analysis, if at all possible, has not been completed.

CORTISOL PREVENTION OF CHRONIC BERYLLIUM DISEASE IN POSTPARTUM RATS: A PILOT STUDY

Introduction:

Beryllium is important in industry, especially the aerospace industry, because of its physical properties : low density; remarkable tensile strength, resistance to fatigue, wear and corrosion, as well as good electrical and thermal conductivity--all properties which it confers on its alloys.

During the mining, manufacturing and processing of beryllium, the substance becomes airborne as a dust, fume or mist (1,2). The main routes of intake for men as well as animals are inhalation and ingestion (3,4). While the absorption of ingested beryllium is probably quite insignificant, the chemical properties of the substance allow it to undergo transformation from soluble to insoluble forms in the lung--facilitating greater retention time. Exceptions to this are the sulfate and flouride compounds of beryllium. An appreciable amount is lost from the lung tissue with time following exposure (5).

Beryllium has been cited as a causative agent in producing systemic disorders that may involve almost any organ or tissue (2,6). However, it has an especially marked predilection for the lungs--producing fibrogenic characteristics, chronic granulomatous changes and marked interstitial infiltration of lung tissues.

The pathogenesis of this disease is poorly understood. Suggestions are that immunologic mechanisms, e.g., cellular immune reactivity with blast transformations of lymphocytes (5,7); and certain forms of stress, including injuries, surgery, infection, and

physiological trauma or stress, such as pregnancy (8,9), cause beryllium translocation, lysosomal instability and consequent cell death resulting in the onset of the chronic disease.

Other suggestions include the premise that decreased adrenal hormone levels destabilize cell membranes leading to lysosomal rupture. Treatment with steroids prevents further progress of degenerative processes associated with the chronic illness (7).

Since it is known that there is significant changes in hormonal levels during pregnancy, including cortisol levels, to further elucidate the mechanism(s) of beryllium action in the onset of the chronic disease, this study was designed to investigate the possible effect of cortisol on the prevention of the onset of chronic berylliosis following parturition and the effect of beryllium exposure on reproductive parameters of the female rat, e.g., litter number, size of pups of litter, and general appearance of pups born to litter.

Materials and Methods:

Normal, mature white female rats , weighing 275-300 gms were used in the experiments. The animals were supplied by the University of Alabama at Tuscaloosa. The animals were maintained on a 12 hour light cycle and on a standard rat diet, provided by the University's Animal Facility. The rats were grouped based on the following treatment plan: Group 1--normal impregnated rats (N), Group 2--beryllium exposed impregnated rats (B-1), and Group 3--beryllium exposed impregnated rats treated postpartum with 25mg/kg of hydrocortisone (B-2), (supplied by Abbot Pharmaceuticals, Inc., Chicago, Illinois), at 3 day intervals for 28 days.

All groups of rats were bred with proven male rats, supplied by the University's Animal Facility. Two females were placed in a cage with a male. Pregnancy was determined by the presence of a cervical plug. After the appearance of the cervical plug, each female rat was placed in a cage and checked periodically for general health status. Groups B-1 and B-2 were bred 4 weeks post-beryllium exposure. Group N was bred within the same time span. The animals of each group were weighed periodically; after mating--considered the initial period, day 10-11 of gestation--mid-pregnancy, and at the time of weaning. All groups were examined for reproductive parameters: number of pups per litter, size and weight of the pups in the litter, and the general appearance and health status of the litters.

Rats in Groups B-1 and B-2 were exposed to 300 ug/kg of beryllium sulfate (supplied by Aldridge Chemical Co., Madison, Wisconsin), utilizing a modified intratracheal instillation method of W.G. Morrow (10), Jaffe and Free (11) and Nicholson and Kinkead (12). Rats were anesthetized with BIO-TAL, (Thiamylal Sodium--0.44 ml of a 4% solution per kg of body weight), to prevent excessive depression of respiration. Each rat was suspended vertically on an animal support stand with gauze attached to the incisor teeth, holding the mouth open. A spatula held to a commerical illuminator with nylon clamps was inserted into the rat's mouth, keeping the tongue flat and allowing for direct observation of the pharynx. With the device held in one hand, the other hand administered the beryllium sulfate solution into the trachea, utilizing a 18 x 3" gauge-curved ball-tipped needle. Following instillation of the solution, the animals were laid on a heating mat until arousal.

The weight of the female rats and pups were measured with a caged-pan balance. The length of the newborn pups were ascertained by measuring the animals from crown to rump in centimeters.

After the completion of the weaning period, the female rats of each group were sacrificed by CO~2 inhalation. The lungs were immediately removed, inflated with and stored in 10% buffered formalin. The specimen were kept under refrigeration at 5°C for histological examination.

All numerical data were analyzed utilizing the SIGSTAT standard t-test [P<0.05]. For unequal samples the F' (folded) statistic for equality of variances was used.

Results:

BIO-TAL produced an anesthetic state of short duration and minimal excitement in the B-1 and B-2 experimental groups, facilitating the intratracheal injection of the beryllium solution. Previous trials with phenobarbital and ketamine produced excessive depression of respiration and resulted in the death of the experimental rats. The animals were aroused within 30-45 minutes post-injection of BIO-TAL.

One female of Group B-1, and 2 females of Group B-2, died day 1 following the intratracheal exposure to beryllium. When compared to Group N, the animals of Groups B-1 and B-2 exhibited symptoms of irritable and jumpy behavior, post-injection, for approximately 5 to 10 days. In addition, some of the animals of Groups B-1 and B-2 experienced sneezing and coughing, which lasted for the first 2-3 days post-exposure to beryllium.

All members of each group of female rats, possessed a cervical plug after placement in the cage with proven male rats. However, there was no significant difference, [P<.05], in the reproductive performance of each group of rats. Although cervical plugs were noted, three (38%) of the females of Group N, did not deliver pups at the end of the expected gestational period. Two females of Group B-1 (33%), possessing cervical plugs, did not deliver any pups. Fifty-percent of the animals in Group B-2 produced pups at the end of the gestational period. Two (33%) of the animals of this group possessed cervical plugs, however, pups were not produced at the end of gestation. One member, 17%, of Group B-2, displayed symptoms of pregnancy, but died after the mid-point of the gestational period.

There was no significant difference between the weights and weight gain of individuals in each group of experimental rats (Table 1). The initial weight of animals in all groups, post-mating, ranged from 308 gms to 327 gms. The weight of the animals ranged from 337-340 gms during the mid-pregnancy stage and 381-409 gms at the time of weaning.

The overall health status of the female rats during and after the gestational period was good--including healthy coat of hair, normal pigmentation of the eyes and extremities and behavior.

There was no significant difference between the litter size of the normal and treatment group of animals (Table 2). The mean number of pups born to each group of animals, N, B-1 and B-2, were 7, 8 and 9 pups, respectively. The averaged length of the pups at day 4-5 ranged from 4.5 to 5.1 centimeters, when measured from crown to rump. The average weight of all pups, 4-5 days old, ranged from 8.3 to 10.3 gms.

The overall appearance of the pups for each group was good--without malformations or difficulty in surviving. However, one pup in Group B-1 and two pups in Group B-2 were delivered with blue heads. Within a period of 6-10 days, the pups recovered from the cyanotic state. At the time of weaning, all pups displayed overall good health, normal extremities and color and behavior. The averaged weight for pups, in all groups at the time of weaning, ranged 100.8 gms to 137.8 gms (Table 3).

Gross examination of the lung tissues removed from members of Groups B-1 and B-2, when compared to Group N, noted some white nodular areas dispersed on the surface of the lung tissue. The color of the lungs of all groups were the same. The lung tissue of all groups were stored in 10% buffered formalin and placed in a refrigerator unit, which malfunctioned and the tissues were inadvertently frozen.

Discussion:

The four basic responses of the respiratory tract to the deposition of inhaled particles are: immunologically induced airway constriction; pharmacologically induced airway constriction--release of serotonin and histamine; acute irritation and reflex broncho-constriction; and non-specific responses. Specifically, exposure to beryllium produces non-specific acute and subacute bronchitis and pneumonitis with irritated respiratory mucosa, coughing and moderate breathlessness within 72 hours of heavy exposure (6,13), as noted in some of the female rats treated with beryllium in Groups B-1 and B-2.

In addition, it has been suggested that individuals possibly have

marked predisposition or susceptibility and sensitivity for contracting the disease in response to beryllium exposure, in that 1-4% of exposed individuals actually acquire the disease (14). While the exact level of beryllium which constitutes toxic or lethal doses are not yet clear, individual susceptibility and sensitivity to beryllium exposure which causes the development of toxic reactions and even death may account for the death of the female rats in Groups B-1 and B-2.

The appearance of cervical plugs, which are indicative of pregnancy, and the failure to produce pups at the end of the expected gestational period may be due to the development of a pseudo-pregnant state in the rats, which in many instances have a duration of 12-16 days (15). On the other hand, the females might have been actually pregnant and resorption of the fetuses occurred. Since the same condition was noted in the normal group of animals, it is difficult to ascertain that beryllium exposure precipitated the condition in the treatment groups. Additional study would have to be undertaken, including a much larger number of--at least fifty animals per group--individuals to begin to correlate beryllium exposure to a decrease or failure in reproductive performance.

Cyanosis and clubbing of the fingers have been seen in about 30-40% of individuals (6.2% of all cases) who developed chronic beryllium disease (3,6). The association of beryllium exposure to the presence of cyanosis in the pups of the female rats in Groups B-1 and B-2, again would be difficult to make. However, since it is known that in some instances, beryllium, in particular beryllium sulfate, does not remain in all instances in the lungs and that the substance

undergoes translocation to produce systemic disease, the question which may be raised is --does beryllium cross the placental barrier and affect the lungs of the unborn fetus? On the other hand, cyanosis occuring in the pups, may be due to the degeneration of the female rats' lung tissue and decreased oxygen delivery to the pups.

Cortisol levels are known to increase and to play a role in different forms of stress, whether it is physical or neurogenic. In addition, it has been shown in several species of animals that the hormone increases during pregnancy and experiences a sharp decline as the result of the onset of labor. The increased levels of cortisol have also been implicated as a factor in causing the onset of labor (16). Increased or high levels of cortisol have been suspected in playing a role in anti-inflammatory effects (17) and causing stabilization of lysosomal memoranes. While a decrease in the hormone has been suggested as a trigger to initiating the onset of the chronic beryllium disease, treatment with cortisol, as mentioned earlier, has been shown to arrest the effects of beryllium on lung tissue. Unfortunately, still very little is known about the mechanism of cortisol's anti-inflammatory effect.

Freezing of tissues often result in artifacts which make histological interpretation difficult. Histological analysis of the preserved lung tissue removed from the rats in the experimental groups is not presented. However, the tissues have been forwarded to Major Harold Davis, Chief of Comparative Pathology, Brooks AFB, TX, for preparation and examination. If a reasonable assessment of the tissue can be made, the findings will be forwarded as an addendum to this report.

Conclusions:

Although some of the resulting characteristics attributed to beryllium exposure were noted in this study, e.g., cyanosis in the pups of female treated with beryllium, failure of the females to produce pups after possible impregnation, and death in some of the female rats treated with beryllium, no conclusive evidence can be noted to associate beryllium as a causative agent in these occurrences when compared to the normal group of rats.

Recommendations:

The study of the mechanism of action of beryllium's effects on the lung tissue and the prevention of the onset of the chronic disease by cortisol should be undertaken as possibly a two-fold effort: 1.) designing a similar study, utilizing at least 50 animals per experimental group inorder to provide a stronger and broader base from which to make correlations; 2.) developing invitro studies to ascertain the effect(s) of beryllium on membrane formation in the lysosome and other tissue fractions, and to determine which constituent of the membrane is most affected by the presence of beryllium; 3.) developing in vitro studies to determine the effects of varying levels of cortisol on lysosomal membranes and membrane constituents, in the absence of and presence of beryllium.

Acknowledgments:

Special thanks to Mr. Rodney C. Darrah, Universal Energy Systems, Inc., and the Air Force Office of Scientific Research for sponsorship of my research grant #760-6MG-078. My deepest appreciation is given to Major Harold Davis, Dr. Bruce Poitrast and Dr. Morris Simon for their technical assistance and to Dr. James Neville, Mr. Godwin Inejeto and Mr. Gerald Moseley for their assistance and support in this project.

Table 1. Averaged weights (gms) of female rats during treatment period with beryllium sulfate (300ug/kg): N represents the normal group; B-1 represents the group of rats exposed intratracheally to beryllium sulfate; B-2 represents the group exposed intratracheally to beryllium sulfate and treated postpartum with hydrocortisone (25mg/kg) for 28 days.

Treatment Period	Group	#Animals	Mean Wt.	STD DEV	STD ERR	T	DF	P
Initial	N	8	317.5	32.0	11.3	.665	7.3	.527
	B-1	8	308.6	19.8	7.0			
	N	8	317.5	32.0	11.3	-.556	7.0	.596
	B-2	8	326.6	32.5	11.5			
Mid-Pregnancy	N	7	340.8	39.6	14.9	-.127	11.4	.903
	B-1	7	343.3	31.5	11.8			
	N	7	340.8	39.6	14.9	.176	11.0	.877
	B-2	6	337.2	34.9	14.2			
Weaning	N	8	387.5	58.5	20.7	-.816	12.0	.450
	B-1	6	409.1	40.5	16.5			
	N	8	387.5	58.5	20.7	.241	11.0	.833
	B-2	5	381.2	35.8	16.0			

Table 2. Mean number of pups born to the litters of each treatment group. N represents the normal group of rats, B-1 represents animals treated with beryllium and B-2 represents animals treated with beryllium and administered hydrocortisone (25mg/kg) postpartum.

Group	#Animals	Mean # pups	STD DEV	STD ERR	VARIANCES	T	DF	PROB
N	8	6.4	6.2	2.2	unequal	-.337	10.98	.767
B-1	6	7.5	6.2	2.5	equal			
N	8	6.4	6.2	2.2	unequal	-.593	6.63	.629
B-2	5	9.0	8.6	3.8	equal			
B-1	6	7.5	6.2	2.5	unequal	-.326	7.11	.755
B-2	5	9.0	8.6	3.8	equal			

Table 3. Averaged total weight (gms) of litters born to the normal group of female rats (N), beryllium exposed rats (B-1), and beryllium exposed rats treated with 25mg/kg of hydrocortisone postpartum.

Group	Mean * # pups	Mean Wt/litter	STD DEV	STD ERR	VARIANCES	T	DF	PROB
N	7	704.7	714.3	252.5	unequal	-.925	9.7	.396
B-1	8	1103.5	856.6	349.7	equal			
N	7	740.7	714.3	252.5	unequal	-.545	7.1	.606
B-2	9	963.8	899.6	402.3	equal			
B-1	8	1103.5	856.6	349.7	unequal	.262	8.5	81
B-2	9	963.8	899.6	402.3	equal			

* Mean # of pups rounded to whole number.

REFERENCES

1. Mayers, M.R., Occupational Health: Hazards of the Work Environment, The Williams and Wilkins Company, Baltimore, Maryland, p. 223, 1969.

2. Dickerson, O.B., "Metal and Metalloids: Antimony and its Compounds" in Occupational Medicine: Principles and Practical Applications, Carl Zenz (ed.). Year Book Medical Publishers, Inc., Chicago, Illinois, pp. 613 and 626, 1982.

3. Preuss, O.P., "Beryllium and Its Compounds" in Occupational Medicine: Principles and Practical Applications, Carl Zenz (ed.), Year Book Medical Publishers, Inc., Chicago, Illinois, p. 619, 1982.

4. Health Assessment Document for Beryllium: Review Draft, Springfield, Virginia (U.S.) Environmental Protection Agency, Research Triangle Park, N.C., U.S. Department of Commerce National Technical Information Service, p. iv., 1984.

5. Sterner, J.H. and M. Eisenbud, Epidemiology of Beryllium Intoxication, Ind. Hyg. and Occup. Med. 4: 123-131, 1951.

6. Freeman, D.G. and H.L. Hardy, Beryllium Disease: The Relation of Pulmonary Pathology to Clinical Course and Prognosis Based on A Study of 130 Cases from the U.S. Beryllium Case Registry, Human Path., 1: 25-44, 1970.

7. Deodhar, S.D., B. Barna and H.S. Van Ordstrand, A Study of the Immunologic Aspects of Chronic Berylliosis, Chest 63: 309-313, 1973.

8. Kronenberg, M.H., Working Conditions for Female Employees, J. Amer. Med. Assoc., 124: 677-683, 1944.

9. Hardy, H.L., Beryllium Poisoning-Lessons in Control of Man-Made Diseases, New Eng. J. Med., 273: 1188-1199, 1965.

10. Morrow, W.G., A Method for Intratracheal Instillation in the Rat, Lab. Anim. Sci., 25: 337-340, 1975.

11. Jaffe, R.A. and M.J. Free, A Simple Endotracheal Intubation Technic for Inhalation Anesthesia of the Rat, Lab. Anim. Sci., 23: 266-269, 1973.

12. Nicholson, J.W. and E.R. Kinkead, A Simple Device for Intratracheal Injections in Rats, Lab. Anim. Sci., 32: 509-510, 1982.

13. Levy, Staurt A., "Occupational Pulmonary Diseases" in Occupational Medicine: Principles and Practical Applications, Carl Zenz (ed.), Year Book Medical Publishers, Inc., Chicago, Illinois, p. 139, 1982.

14. De Nardi, J.M., Long-Term Experience with Beryllium Disease, A.M.A. Arch. Ind. Health, 19: 110-116, 1959.

15. Baker, H.J., J.R. Lindsay and S.H. Weisbroth (eds.), The Laboratory Rat: Research Applications, Vol. 2, American College of Laboratory Animal Medicine Series, Academic Press, New York, p. 85, 1980.

16. Yoshida, T., H. Suzuki, Y. Hattori and K. Noda, Hormonal Changes around the Parturition in Rats, Tohoku J. Exp. Med., 135: 87-91, 1981.

17. Guyton, Arthur C., Textbook of Medical Physiology, 4th ed., W.B. Saunders Company, Philadelphia, pp. 893-895, 1971.

RESEARCH INITIATION GRANT REPORT

UNITED STATES AIRFORCE SCHOOL OF AEROSPACE MEDICINE

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Prepared by:	Ronald Bulbulian, Ph.D.
Academic Rank:	Associate Professor
Department and	Exercise Physiology Laboratory
University:	University of Kentucky
Research Location:	University of Kentucky 201 Seaton Building Exercise Physiology Lab Lexington, KY 40506-0219
USAF Researcher:	Loren Myhre
Date:	February 1990
Contract No:	F49620-88-C-0053/SB5881-0378
Subcontract No.:	S-210-9MG-057, Amendment No. 0001

Blood Flow Distribution In The Non-Working Forearm During Exercise

by

Ronald Bulbulian

ABSTRACT

A non-invasive method for determining forearm muscle blood flow (MBF) was investigated by combining Laser doppler velocimetry for skin blood flow (SBF) and strain gauge plethysmography for forearm blood flow (FBF).

The results from data gathered on 5 male and 5 female moderate to well conditioned adults indicate that an exercise intensity graded response for SBF is clearly evident at 30, 50, and 70% VO_2 max. FBF measures are highly variable with several subjects showing significant increases in FBF not associated with SBF elevations. The derived variable, MBF, shows an identifiable trend toward muscle vasodialation with increasing duration of exercise to steady state core temperature (Tes) and with increasing intensity of exercise. We conclude from the data that MBF responses to exercise are variable and not predictable among subjects and that FBF is not a good indicator of SBF during

exercise in normal exercise conditions which are not artificially controlled.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Universal Energy Systems must be mentioned for their concern and help to me in all administrative and directional aspects of this program.

The help of Don Tucker in overcomming many technical roadblocks, and Dr. Loren Myhre in providing encouragement and timely help was greatly appreciated.

Professional interaction with Susan Bomalaski, Dr. Sally Nunnley, Dr. S. Constable, and Mr. John Garza were very helpful and provided added enlightenment. Appreciation is also expressed to Mary Amos, Drew Pringle, and Joe Essex for their assistance in data gathering and analysis. Finally our thanks to Mary Clark for her work in preparation of this manuscript.

INTRODUCTION

Exercise induced redistribution of visceral blood flow to supply the needs of working muscles is well established. However, the redistribution of cardiac output (Q) from non-working muscle to the metabolically active tissues is less established. Much of the literature reporting on this important area of research is anchored by studies which attempt to fraction the skin and muscle blood flow (SBF, MBF) in the non-working forearm. One group has measured the absence of SBF through epinephrine iontophoresis which is less than totally effective (25) and, the second report has predicted SBF from oxygen saturation of deep and superficial veins (2). Furthermore, the reports and much of the current work employs exercise parameters which are ill-defined and inadequate. Nevertheless, the current research continues to buttress its case on these reports and the belief that non-working MBF is unchanged or decreased during lower body leg exercise and that forearm blood flow (FBF) evaluation can be interpreted as elevation of SBF (8,18).

Greenleaf et al. (6) have recently shown an increased FBF in the resting muscle during supine and sitting ergometer exercise. This report has been interpreted by

Lamb (14) to provide evidence for increase in MBF to non-exercising muscle during exercise. This interpretation may be partly due to reports of skin vasoconstriction (i.e., lower finger blood flow) in the hand during bicycle work (4) or similar vasoconstriction during rest to exercise transition at the onset of work (1,2), but it is certainly misinterpretation of Greenleaf's data which was taken 1 minute post maximal exercise and most likely represents a hyperemic response following work cessation (1,2).

Given present concerns with cardiovascular drift (CVD) during prolonged exercise and the need to better understand circulatory adjustments to accommodate metabolic and thermoregulatory needs (18), the methodological shortcomings (i.e., exercise quantitation, Q measurement, assumptions) in the landmark exercise-SBF studies of Zelis et al. (25) and Bevegard and Shepherd (2) make a reexamination of FBF distribution necessary.

METHODS

Subjects

10 moderate to well conditioned subjects were chosen for the study (5 male and 5 female). Each subject was briefed regarding the procedures of the study and provided informed consent.

Research Procedure

The study employed a counterbalanced repeated measures research design with each of ten subjects acting as their own control. Following extensive screening which included cholesterol, ECG, and health screening to confirm suitability for participation in the study, all subjects performed a maximal exercise test on a mechanically braked cycle ergometer (Quinton model 844) to determine maximal oxygen consumption (VO_2 max). Subjects were clothed in shorts and running shoes (females wore cut short tank top shirts) in order to maximize convective and evaporative cooling. Cooling was assisted by exposing each subject to rapid air flow by means of a household fan. The results of the maximal cycle ergometer test were used to calculate relative exercise workloads for testing subjects at 30%, 50%, and 70% of VO_2 max. All tests were conducted under similar environmental conditions in a temperature regulated environment (23 + 1 and 30-35% relative humidity). The

submaximal test conditions were completed at least 3 days apart (usually 7 or more days) to minimize fatigue and possible confounding effects of training. Nude body weights were measured before and after each submaximal trial and no fluids were provided during the course of an exercise trial.

Data Collection Protocol

For each of the 30%, 50%, and 70% VO_2 max exercise protocols subjects were instrumented for the various measurement devices and rested for 10 minutes before the start of exercise. Cardiac output (Q), blood pressure (BP), esophageal temperature (Tes), skin temperature (Tsk), and forearm blood flow (FBF) were measured every 5 minutes throughout the test. VO_2 and skin blood flow (SBF) were measured continuously. Additional measures were taken throughout each trial as needed to confirm unusual findings or inconclusive recordings due to movement artifact.

The Exercise Apparatus

The cycle ergometer with toe clips was housed within a body stabilization apparatus constructed to keep the torso as near to motionless as possible. This was achieved by securing the torso of the subjects to a padded backrest by means of a 15 cm wide velcro belt fastened over one shoulder and under the opposite arm. The belt was sufficiently secure to provide stabilization without hindering

respiration. To further assist in stabilization, the hips were strapped securely to the bicycle seat which allowed a counterforce to the leg extension during the heavier workloads (i.e. 70%) which the body weight alone could not support. These measures allowed for the optimal conditions needed to measure SBF and FBF in the relaxed (non-exercising) forearm during leg exercise on a bicycle. To further reduce movement artifact the arms were rested 90 abducted and the elbows 90 flexed on an arm rest at the side of the subject at heart level to negate hydrostatic forces. Finally, the right hand of each subject was gently molded into an appropriate amount of potters clay to eliminate finger movements which caused artifact recordings in the plethysmographic FBF record.

Oxygen Consumption (VO_2)

VO_2 (ml/kg/min) was measured continually during baseline metabolism determination and during exercise on a breath-by-breath basis using a Sensormedics Horizon 4400 metabolic cart. The screen display was set to output integrated minute values and allowed fine tuning of the metabolic cost of exercise to more precisely match the pre-determined VO_2 values at 30%, 50%, and 70% VO_2 intensity for the relative submaximal exercise loads.

Analog input for heart rate (HR) was also recorded in digital format to back up to printed output for future reference.

During all tests, following 10 minutes of baseline VO_2 determination, the subjects were asked to pedal at 60 rpm and the pre-determined workload was applied. The mechanical workloads were adjusted when necessary to maintain the required metabolic work level during Tes monitoring in order to achieve a near thermic (Tes) steady state. Bicycle ergomter work loads were adjustable to the nearest 100 kg-m/min (16 watts).

Temperature Measurement

To facilitate rapid achievement of steady state core temperature.(Tes) disposable YSI 491 A esophageal probes were placed 38-40 cm deep nasally and adjusted slightly to obtain an optimal temperature recording (highest Tes near right arm) . The probe was taped in place for the remainder of the exercise period. Temperatures were recorded from a YSI 12 channel Tele-Thermometer (Model 44TA) every 5 minutes and care was taken to not allow a subject's swallowing to alter correct temperature recording. Simulatneous Tsk measurements were made from the same YSI Tele-Thermometer attached to a 709B skin temperature probe secured to the dorsal surface of the forearm next to the Laser Doppler skin

blood flow probe. The attachment site was approximately over the belly of the brachioradialis muscle. The probe was attached to the forearm with a concentric fastener which allows for free ventilation about the skin surfaces surrounding the thermocouple.

Cardiac Output

Q was measured from aortic flow velocity and acceleration determined by continuous wave Doppler echocardiography (Quinton-ExerDop) at ultrasound transmission frequency of 3.0 MH_z. The Doppler shifted blood velocity was measured from the suprasternal notch region with the angulation adjusted to about 0-20 degrees to align with the long axis of the ascending aorta. The signal was monitored by audio output from the Doppler system and digital output transmitted to an oscilloscope. The highest velocity and lowest spectral dispersion were used as criteria for signal acceptance. The beat-by-beat data output for each 30 second sampling period was printed and the average of the 10 highest unrejected, stroke distances recorded for that period were used to calculate Q.

Aortic root diameters were measured with an Irex III echocardiographic system and recorded on magnetic tape. The diameters were determined with the Doppler probe held in the parasternal short axis position. The cross section was

carefully examined to ensure accurate measurement of the aortic root by visual reference of the valve cusps.

Blood Pressure (BP)

Systolic and diastolic BP measurement was made at each stage of exercise and rest at 5 minute intervals by brachial artery ausculation with a Colin STBP 680. Automated BP monitor measurements were made at the 4th minute of each stage in order to avoid interference with the FBF measurements which were initiated during the last minute of each 5 minute period. All BP measures were taken on the left arm of the subjects.

Heart Rate (HR)

HR was measured using a cardiostress ECG cart with standard frontal plane leads (disposable electrodes) placed bilaterally at the shoulders and hips. The output was displayed on a Quinton ECG monitor and channelled into a Grass 4 channel polygraph for hardcopy ECG recording and subsequent linking to the Horizon metabolic cart where a digital HR record was printed on a minute by minute basis.

Skin Blood Flow (SBF)

SBF was measured on the belly of the brachioradialis muscle of the right forearm in the previously described fully relaxed position. The flat 15 mm probe of an LD5000

Med Pacific Skin Doppler device was secured with double-sided adhesive tabs to the shaved forearm after being lightly preped with electrode gel. The instrument was calibrated prior to every experimental trial and the SBF data were derived from a minimum 2 minute record (free of FBF recording artifact) at the end of each 5 minute recording period. SBF was measured as the mean amplitude determined from the strip chart record and was a highly reliable measure of SBF expressed in arbitrary units.

Forearm Blood Flow (FBF)

FBF was measured by venous occlusion adjacent to the site of the SBF measure with a double-stranded mercury strain gauge plethysmograph (Hokanson EC-4). The sensitive device was calibrated with a 1% marker prior to each recording on a Grass polygraph with two records being made during the 5th minute of every 5 minute recording period. Venous occlusion was achieved by a carefully calibrated air cylinder system attached to two solenoid switches which when activated/deactivated would relay 50 mm Hg air pressure to and away from an inflation bladder on the subject's right arm. A 30 mm Hg occlusion pressure was quickly achieved (less than 2 seconds) with the steady state 50 mm Hg pressure being reached by 4-5 seconds. The measurement of FBF was made by drawing a tangent line from the initiation

of venous occlusion to the initial arterial inflow curve, disregarding inflation cuff / pressure artifact when present, and relating this slope to the calibration standard preceeding each record. The data are expressed in ml/100 ml tissue and the average of two acceptable records for each time period were entered into analysis.

Statistical Analysis

The data were coded and entered into an XEDIT file management system on the University of Kentucky IBM mainframe and analyzed by Statistical Analysis System (SAS) repeated measures ANOVA to evaluate differences among the means for SBF and FBF across exercise intensity. Descriptive data were also generated on all other factors to help explain the findings of the blood flow responses. The $p < .05$ level of significance was deemed acceptable for interpretation.

RESULTS

The descriptive characteristics of the 5 male and 5 female subjects used in the study are combined in Table 1. Metabolic responses to graded bicycle exercise show a predictable linear response with workloads evaluated in linear fashion from rest to 30% VO_2 max and subsequently, to 50 and 70% VO_2 max. Because of close monitoring, the minute by minute VO_2 the steady state VO_2 measures very closely approximate the target VO_2 for each relative workload (Table 2). As would be predicted, the Tes increased also with energy output from rest through various levels of exercise (Figure). The work levels of 50% and 70% VO_2 max did not result in significantly different Tes. Tsk did not show significant variation between 50% and 70% levels of exercise as they reached similar decreasing values (Figure 2). In contrast, the 30% exercise Tsk values show higher levels throughout the exercise period.

FBF when expressed in ml/100cc shows either depressed or constant initial FBF in transition from rest to all three exercise levels (Figure 3). The pattern continues through 10 minutes of exercise for all three exercise bouts but a shar significant increase in FBF occurs at 15 minutes for the high intensity exercise group and is sustained thereafter ($p < .05$). The 30 and 50% VO_2 max intensities of exercise produced no subsequent changes in FBF over the

30 minutes of exercise which few subjects had to complete before reaching a steady Tes. When FBF was expressed as a % change from resting baseline, (Figure 4) the values show great variation in the means and standard errors (+ SE). However, the FBF change plots tend to support the gradation in FBF responses to grading intensity of exercise to a greater extent.

In contrast to FBF, SBF responses show a clear and significant separation of the three intensities of exercise over time with graded increases corresponding to the energy cost of exercise (Figure 5). The data expressed as % change in SBF are essentially identical but show substantial individual variation. At 30% VO_2 max the changes range from -10% to + 830% and values inbetween. But at higher (70%) intensity, after a transcient cutaneous vasoconstriction the SBF changes all reflect on elevation in all nine subjects (Figures 6 and 7).

The purpose of this study was to assess MBF during various intensities of exercise during conditions simulating normal thermoregulatory control (i.e. ambient skin environment and SBF). Figures 8, 9, and 10 show the derived MFB % change (FBF% minus SBF% change) over time to steady state Tes at 30%, 50%, and 70% VO_2 max. The data suggest that MBF% change shows slight depression at low intensity exercise with some subjects showing sustained elevations

At 50% VO_2 max. The response is a largely one of muscular vessel constriction and depressed MBF. At the highest level of exercise the MBF response is by in large unchanged (same as 50%) but two subjects show consistant elevation (100% to 200%).

DISCUSSION

The purpose of this experiment was to study the effect of thermoregulatory adjustments in skin and muscle blood flow to varying levels of leg exercise.

Numerous experiments have examined the effects of absolute levels of exercise exposure on several cardiovascular parameters and have observed great variability in the data. The data have generally favored the use of relative loading experiments for exercise research (6,13,18).

The reasoning is centered around the concept that physiological responses to exercise are controlled by an intensity loading that is related to the individuals fitness level. Thus, two individuals working at a 150 watt power output may display different heart rates with the less fit individual showing the greater HR response. Conversely, two individuals working at the same exercise HR will not be expected to demonstrate the same thermoregulatory responses to exercise (1,6,21). Not withstanding these observations, a substantial amount of work has recently been published which has not taken these factors into account (8-12,20,21). Also, in an effort to illusidate mechanisms of thermoregulation and sweating, experiment unusual procedures have been employed, which have incorrectly described the

normal circulatory response during exercise (8-12,14,20,21).

The protocol chosen for this study uses the sitting exercise position which most closely approximates the normal orthostatic challenge accompanying exercise. This position has been reported to attenuate skin blood flow responses during exercise and more closely illustrates the common exercise response (12). The subjects in the experiment were also exposed to a moderate flow of air to enhance evaporative cooling during exercise (18). This had the effect of establishing lower skin temperatures (Tsk) during the study which in turn should have reduced skin blood flow (SBF). However, although the normal eveluations in SBF may have been attenuated, the data nevertheless, show a graded response to exercise with the highest SBF levels being achieved in the 70% VO_2 max condition (Fig. 5).

The data regarding forearm blood flow (FBF) are less consistent with good reason. Venous occlusion plethysmography measures blood volume accumulation in the occluded limb. However, this accumulation represents blood entering the limb (arterial inflow) segment from vascular muscle and cutaneous sources. The differential response of the skin and inactive muscle to circulatory demands during exercise confound this measure and can result in tremendous variability. This variability, however, is not the result

of MBF effects on the SBF (17). The results in Figure 3 document this variability and provide initial evidence for a need to reexamine our understanding of SBF and MBF during exercise in different exercising populations and varying exercise conditions (6,14,20). The data seem to show that FBF has a tendency to vary according to the intensity of exercise being undertaken by the subjects. The lowest (30%) intensity level of exercise shows little or no change in FBF and possibly a tendency toward vasoconstriction. The 50% VO_2 max of exercise does not show a reduction in FBF but the flow values hover around the baseline level. The highest intensity exercise showed an initial slight reduction (p = NS) but the flow increased to a near two fold increase at steady state. The reduction in the FBF values at minutes 25 to 30 represent statistical artifact to the lower number of subjects represented in the latter time periods (subjects reached steady state Tes sooner). To clarify and better understand the nature of this response, the data were converted to a relative FBF response with each subject's FBF reported a a percent change in his/her baseline resting FBF (Fig. 4). Here the data show a definite drop in the FBF during initial satges of exercise at light intensity but the data are still confounded by one subject who showed an inordinate response in FBF during this trial. The individual data best characterize

the response to this exercise challenge in the ten subjects used in the study (Fig. 11). Seven of the ten subjects in the study showed diminished or unchaged responses in FBF with one subject completing the trial with a 70% reduction at the end of 15 minutes. The results for the 50% intensity trial are a little more difficult to interpret since six of ten subjects chose to complete the trial consecutively after having obtained a steady state response to the 30% exercise workload. Although the kinetics of the response are unclear, the steady state responses to a plateau Tes and the corresponding FBF and SBF values a.e reliable. The data in this case show a slight elevation with exercise which is less variable than the same subject's response to the lighter workload. Expression of the response as a relative (%) change is of no consequence (Figures 3 and 4). Once again, however, the individual plots by subject are more revealing. Three of the four subjects initiating their work at this workload (50%) showed a transient depression in FBF. By 20 minutes there were still 3 subjects showing a depressed FBF with only one subject showing a negative (-41%) final value (Fig. 12).

The results of the high intensity work bout differed significantly from the moderate and low intensity bouts but once again the variability was much greater notwithstanding the use of relative exercise loading to suppress the

variability (1,5). While some of the subjects still showed a vasoconstriction response during the initial stages, the FBF increased thereafter and reached values above 300% in three of the subjects. Another four subjects showed elevated values, and two again terminated with depressed negative values. These data underscore the individual variabilities displayed by subjects when exercising under more normal (less controlled/artificial) exercise conditions and partially support Wenger et. al. who have found no evidence for a negatively linear (vasoconstrictor) effect of increasing exercise intensity on FBF (24).

These findings bring into question the commonly accepted practice of assuming changes in FBF to be a result of SBF changes during leg exercise (8-12,20,23). The result of correlational studies where laser doppler SBF measures have shown high correlation with simultaneous plethysmographic FBF measures has strengthened the practice even though the data have only been fully verified in the none exercise state (11). Research has demonstrated a reduction in the MBF of inactive muscle during leg exercise. This has been documented in man and primate (7,10,25). Given the great variations in FBF in the presence of consistent increases in SBF graded with exercise, one must reconsider the use of venous occlusion plethysmography in the assessment of SBF. The present results suggest a very

strong muscle constriction when no changes in SBF are accompanied by reductions in FBF or the possibility of an elevated MBF when the elevated SBF is accompanied by an elevated FBF. In the present study only one subject showed a depressed % SBF which was sustained throughout the 30% intensity trial (-5.8%). Five subjects showed elevated SBF accompanied with depressed FBF at the lowest intensity of exercise. Zelis et. al. have reported that about 50% of the increase in FBF during exercise can be attributed to increases in the SBF (25). These findings are not consistently supported by the present data. Assuming that the FBF and SBF data are directly comparable, some reports have set the resting flows for both tissues at about 4-5 ml/100cc/min (18,22,23), when expressed as relative (%) changes, the derived differences between the two measures should represent a derived MBF. Clearly in half the cases this manipulation does not show a consistent reduction in MBF as a response to exercise (Figures 14-15). It can be argued that the SBF contribution to the volume changes seen in the plethysmographic record are minimal. A crude calculation of the cross sectional area represented by the mercury strain gauge in most of the subject's forearms was about 49 square cm, allowing for skin, fat, and bone construction to the area would reduce the muscle area to about 33 square cm. A generous assumption for skin

thickness of 3 mm. would put the skin's cross section for the same diameter segment at about 7.5 square cm. If these calculations are reasonable, it would be hard to suppose that FBF could adequately represent SBF assessment even if MBF was predictable and consistent. To further illustrate the problem, if the cross sectional variation was taken into account by weighing the SBF contribution to FBF as a ratio of the cross sectional areas, the FBF-SBF plots would be represented by the Figures 14-16. These data, if reliable, clearly show an increasing MBF in some subjects and a trend toward dimished vasoconstrictor tone in the muscle with increasings levels of work. The present findings are inconsistent with the findings of Zelis et. al. and Bevegard et. al. and do not support the current practice of indirect measurement for SBF by plethysmographic FBF measurement.

In conclusion, the present findings suggest that present methods using laser doppler techniques for SBF measurement seem to better indicators of SBF during exercise and that the use of plethysmography in measuring SBF during exercise needs to be reevaluated carefully under tightly controlled conditions. Studies which have used plethysmography or have not taken into account the fitness or conditioning level of the subjects used could be in error due to fitness induced changes such as sweat rate alterations (3), environmental and/or cardiovascular

interactions (13-15,16,18), effects of obesity on individual responses (23), and changes in muscular conductance which are certain to effect the thermoregulatory (cardiovascular) responses to an exercise challenge (19).

REFERENCES

(1) Astrand, P. O., and K. Rodahl. Textbook of Work Physiology. McGraw-Hill, New York, NY. 1977.

(2) Bevegard, B. S., and J. T. Shepard. Reaction in man of resistance and capacity vessels in forearm and hand to leg exercise. J. Appl. Physiol. 21: 123-132, 1966.

(3) Buono, M. J., and N. T. Sjoholm. Effect of physical training on peripheral sweat production. J. Appl. Physiol. 65(2): 811-814, 1988.

(4) Christensen, E. H., and M. Nielsen. Investigation of the circulation in the skin at beginning of muscular work. Acta. Physiol. Scand. 4: 162, 1942.

(5) Greenleaf, J. E., C. J. Greenleaf, D. H. Card, and B. Saltin. Exercise-temperature regulation in man during acute exposure to simulated altitude. J. Appl. Physiol. 26(3): 290-296, 1969.

(6) Greenleaf, J. E., L. D. Montgomery, P. J. Brock, and W. Van Beaumont. Limb blood flow: Rest and heavy exercise in sitting and supine positions in man. Aviat. Space Environ. Med. 50: 701-717, 1979.

(7) Hohimer, A. R., J. R. Hales, L. B. Rowell, and O. A. Smith. Regional distribution of blood flow during mild dynamic leg exercise in the baboon. J. Appl. Physiol.: Respirat. Environ. Exercise Physiol. 55(4): 1173-1177, 1983.

(8) Johnson, J. M. Responses of forearm blood flow to graded leg exercise in man. J. Appl. Physiol.: Respirat. Environ. Exercise Physiol. 46(3): 457-462, 1979.

(9) Johnson, J. M., D. S. O'Leary, W. F. Taylor, and M. K. Park. Reflex regulation of sweat rate by skin temperature in exercising humans. J. Appl. Physiol.: Respirat. Environ. Exercise Physiol. 56(5): 1283-1288, 1984.

(10) Johnson, J. M., and L. B. Rowell. Forearm skin and muscle vascular responses to prolonged leg exercise in man. J. Appl. Physiol. 39:920-924, 1975.

(11) Johnson, J. M., W. F. Taylor, A. P. Shepard, and M. K. Park. Laser-Doppler measurement of skin blood flow: comparison with plethysmography. J. Appl. Physiol.: Respirat. Environ. Exercise Physiol. 56(3): 798-803, 1984.

(12) Johnson, J. M. and M. K. Park. Effect of upright exercise on threshold for cutaneous vasodilation and sweating. J. Appl. Physiol. 50: 814-818, 1981.

(13) Kenney, W. L., E. Kamon, and E. R. Buskirk. Effect of mild essential hypertension on control of forearm blood flow during exercise in the heat. J. Appl. Physiol.: Respirat. Environ. Exercise Physiol. 56(4): 930-935, 1984.

(14) Lamb, D. R. Physiology of Exercise. MacMillan Publishing Co., New York, NY, 1984. p. 150.

(15) Nadel, E. R. Body fluid and electrolyte balance during exercise: competing demands with temperature regulation. Thermal Physiology, 365-376, 1984.

(16) Nadel, E. R. Temperature regulation during exercise. In Houdus, Y. and Gulen, J. D. (eds.): New Trends in Thermal Physiology. Paris, Masson, 1978, pp. 143-153.

(17) Saumet, J. L., D. L. Kellogg, Jr., W. F. Taylor, and J. M. Johnson. Cutaneous laser-doppler flowmetry: influence of underlying muscle blood flow. J. Appl. Physiol. 65(1): 478-481, 1988.

(18) Shaffrath, J. D., and W. C. Adams. Effects of airflow and work load on cardiovascular drift and skin blood flow. J. Appl. Physiol. 56:1411-1417, 1984.

(19) Snell, P. G., W. H. Martin, J. C. Buckey, and C. G. Blomqvist. Maximal vascular leg conductance in trained and untrained men. J. Appl. Physiol. 62(2): 606-610, 1987.

(20) Taylor, W. F., J. M. Johnson, W. A. Kosiba, and C. M. Kwam. Graded cutaneous vascular responses to dynamic leg exercise. J. Appl. Physiol. 64(5): 1803-1809, 1988.

(21) Taylor, W. F., J. M. Johnson, D. S. O'Leary, and M. K. Park. Modification of the cutaneous vascular response to exercise by local skin temperature. J. Appl. Physiol.: Resprat. Environ. Exercise Physiol. 57(6): 1878-1884, 1984.

(22) Tur, E., M. Tur, H. I. Maiback, and R. H. Guy. Basal perfusion of the cutaneous microcirculation: measurements as a function of anatomic position. The J. Investigative Dermatology. 81: 442-446, 1983.

(23) Vroman, N. B., E. R. Buskirk, and J. L. Hodgson. Cardiac output and skin blood flow in lean and obese individuals during exercise in the heat. J. Appl. Physiol. 55(1): 69-74, 1983.

(24) Wenger, C. B., M. F. Roberts, J. A. J. Stolwijk, and E. R. Nadel. Forearm blood flow during body tempertature transients produced by leg exercise. J. Appl. Physiol. 38(1): 58-63, 1975.

(25) Zelis, R., D. T. Mason, and E. Braunwald. Partition of blood flow to the cutaneous and muscular beds of the forearm at rest and during leg exercise in normal subjects and in patients with heart failure. Circulation Research 24:799-806, 1969.

Table 1. Descriptive Characteristics of 5 Male and 5 Female Subjects Combined.

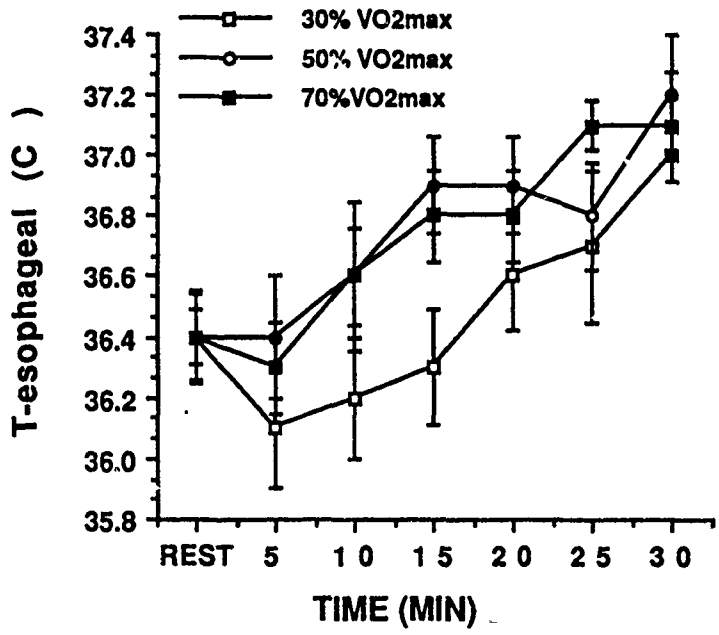
VARIABLE	MEAN +/- SD	RANGE
Age (yrs)	29.4 +/- 7.6	18 - 40
Height (cm)	175.2 +/- 6.8	165 - 185
Weight (kg)	69.5 +/- 9.9	54.5 - 79.0
VO_2 max (ml/kg/min)	38.3 +/- 4.5	29.5 - 45.1
Work max (kg-m/min)	1080 +/- 270	600 - 1400
Arm Circumference (cm)	25.6 +/- 2.4	21.3 - 29.5

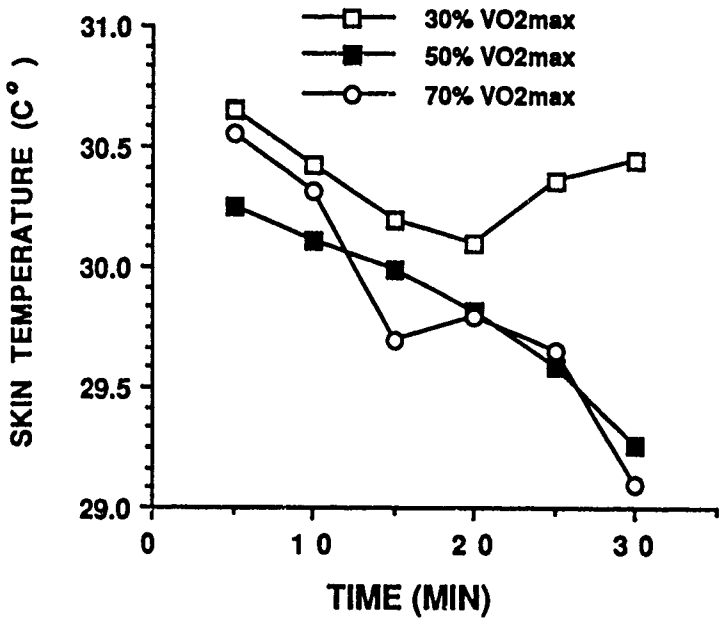
All data expressed as Mean +/- SD (Range given below).

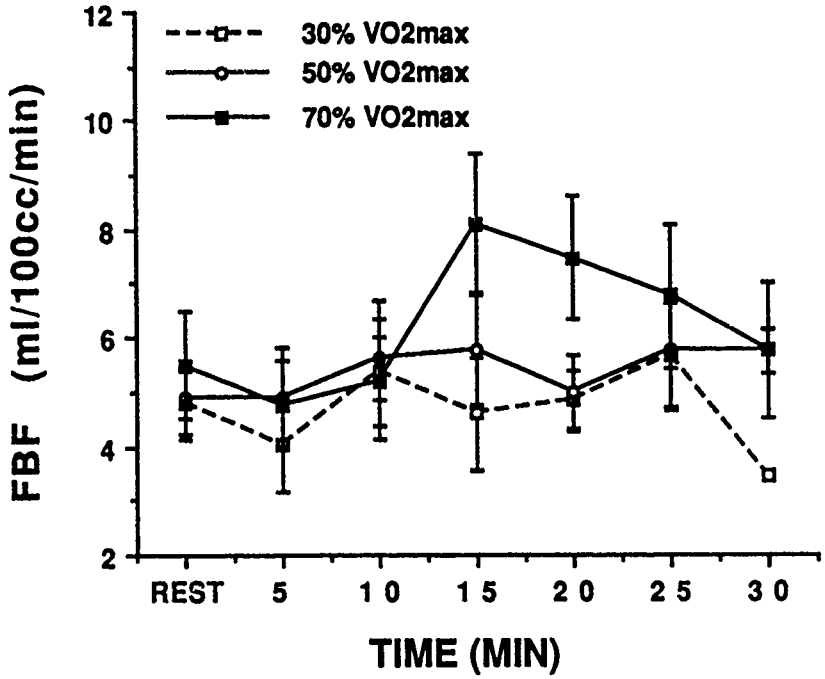
Table 2. Metabolic and Thermoregulatory Responses to Exercise at 30, 50, and 70% VO_2 max.

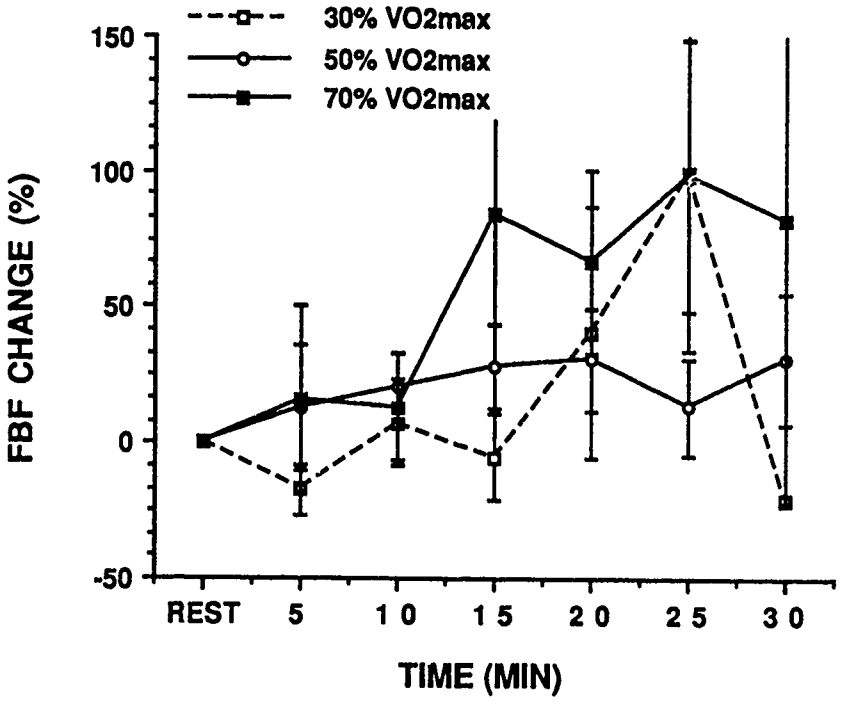
VARIABLE	30% (n=10)	50% (n=10)	70% (n= 9)
Work Rate (kg-m/min)	300 +/- 66 (200 - 400)	570 +/- 125 (400 - 800)	822 +/- 192 (500 - 1100)
VO_2 (ml/kg/min)	13.5 +/- 2.6 (10.2 - 17.5)	21.0 +/- 2.4 (17.2 - 24.5)	27.7 +/- 3.5 (21.3 - 32.0)
Heart Rate (beats/min)	107 +/- 14 (89 - 126)	132 +/- 9 (113 - 146)	151 +/- 13 (127 - 170)
Blood Pressure (mm Hg)	119/73 +/- 10/6 (104/64 - 136/82)	154/67 +/- 19/9 (127/51 - 175/79)	162/68 +/- 28/8 (125/57 - 200/81)
Tes (C)	36.3 +/- 0.6 (35.0 - 36.9)	36.8 +/- 0.5 (36.0 - 37.4)	37.0 +/- 0.2 (36.8 - 37.4)
Tsk (C)	30.2 +/- 0.9 (28.7 - 31.7)	29.5 +/- 1.7 (27.1 - 31.3)	29.6 +/- 1.8 (27.0 - 32.4)

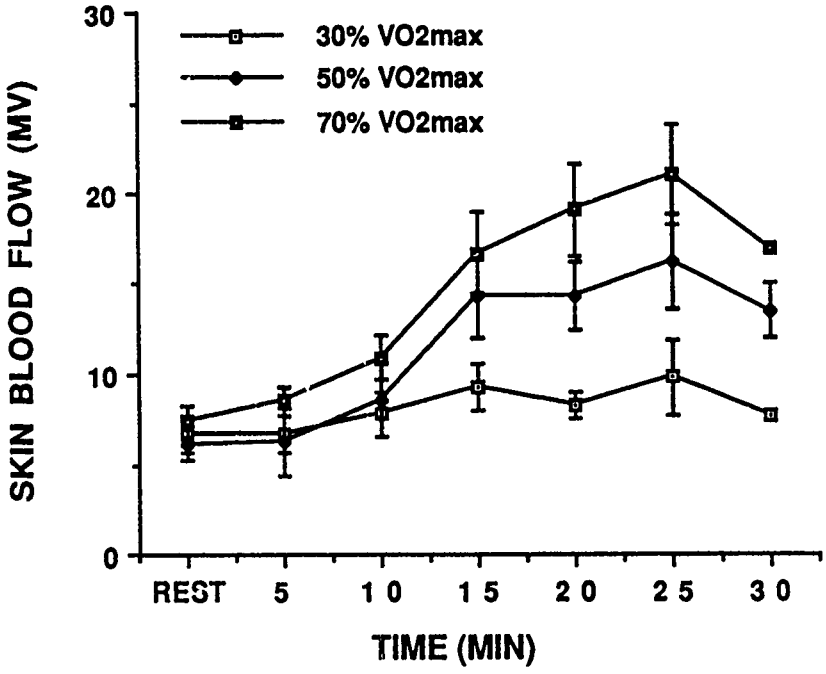
All data expressed as Mean +/- SD (Range given below).


30% VO2max
50% VO2max
70%VO2max
37.4
37.2
37.0
36.8
36.6
36.4
36.2
36.0
35.8
T-esophageal (C)
REST 5 10 15 20 25 30
TIME (MIN)

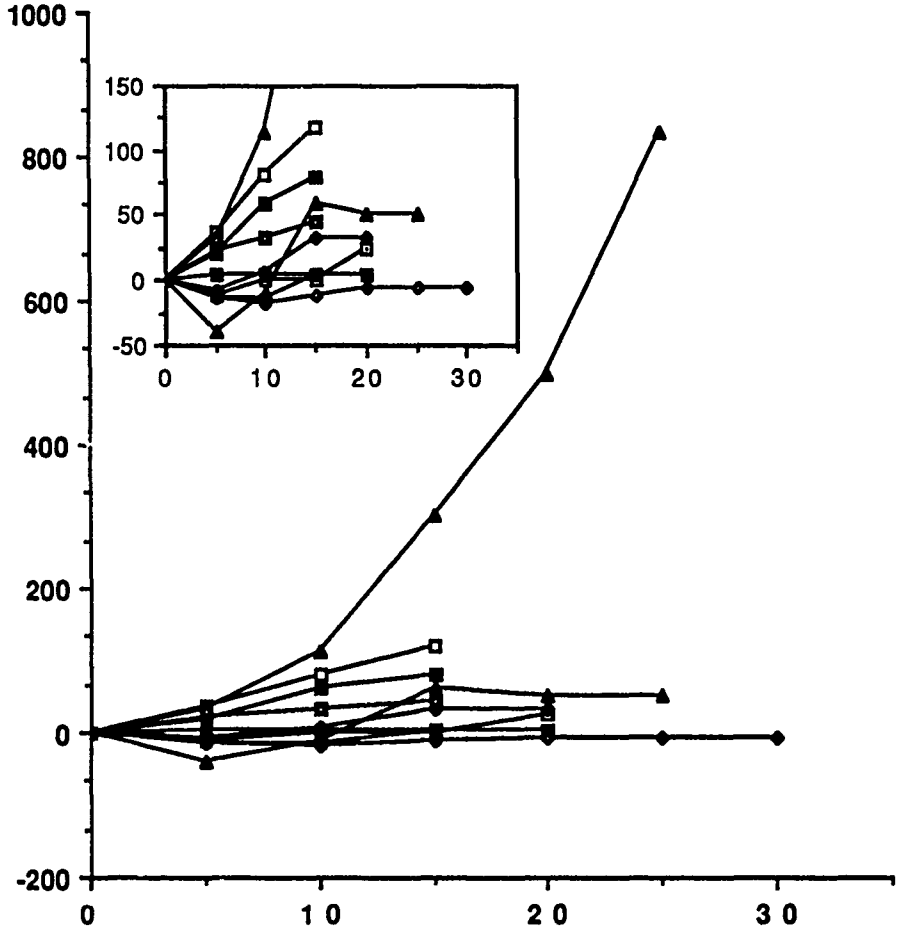

30% VO2max
50% VO2max
70% VO2max
SKIN TEMPERATURE (C°)
31.0
30.5
30.0
29.5
29.0
0
10
20
30
TIME (MIN)


30% VO2max
50% VO2max
70% VO2max
FBF (ml/100cc/min)
12
10
8
6
4
2
REST
5
10
15
20
25
30
TIME (MIN)

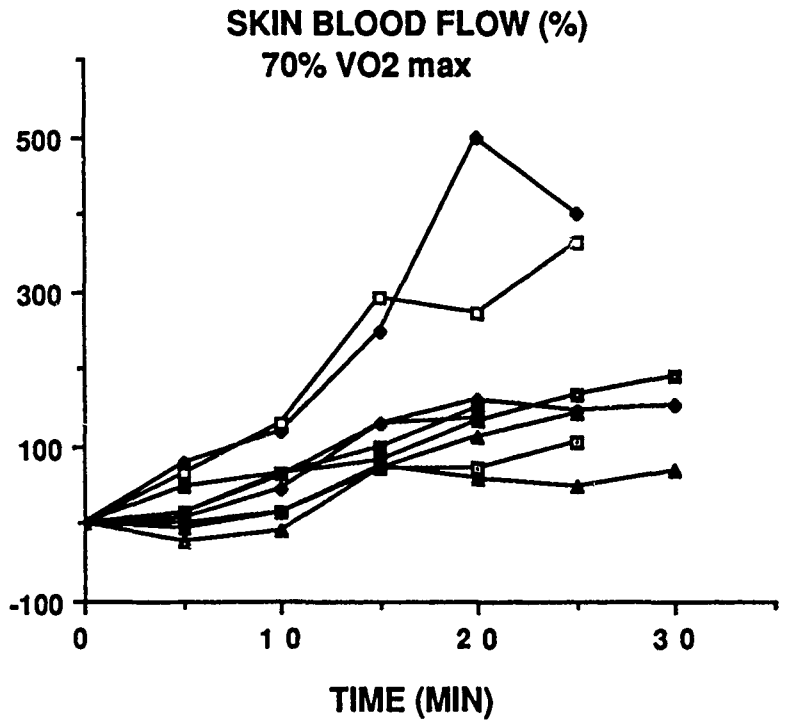

30% VO2max
50% VO2max
70% VO2max
150
100
50
0
-50
FBF CHANGE (%)
REST
5
10
15
20
25
30
TIME (MIN)

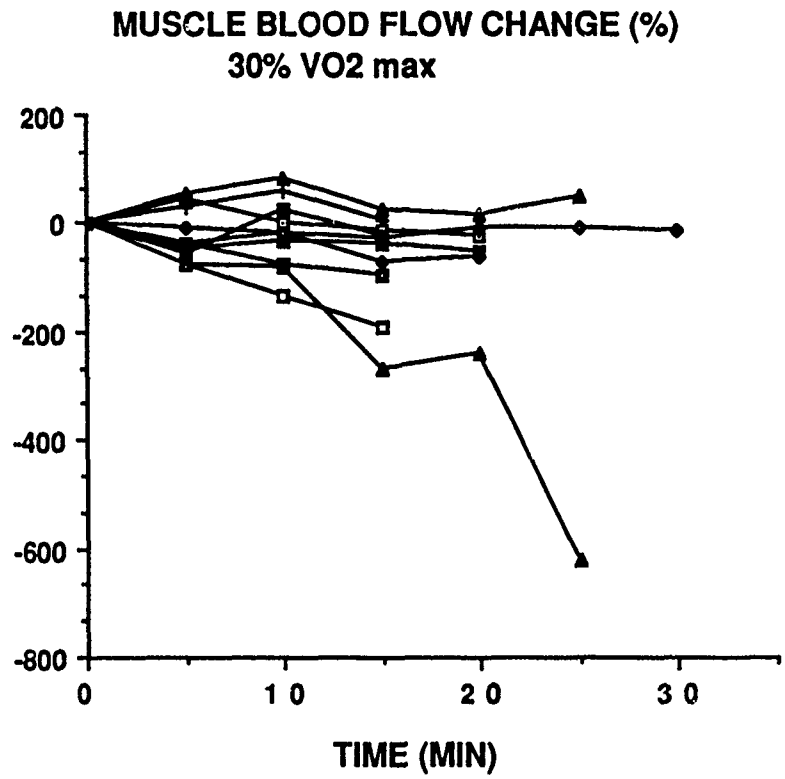

30% VO2max
50% VO2max
70% VO2max
SKIN BLOOD FLOW (MV)
30
20
10
0
REST
5
10
15
20
25
30
TIME (MIN)

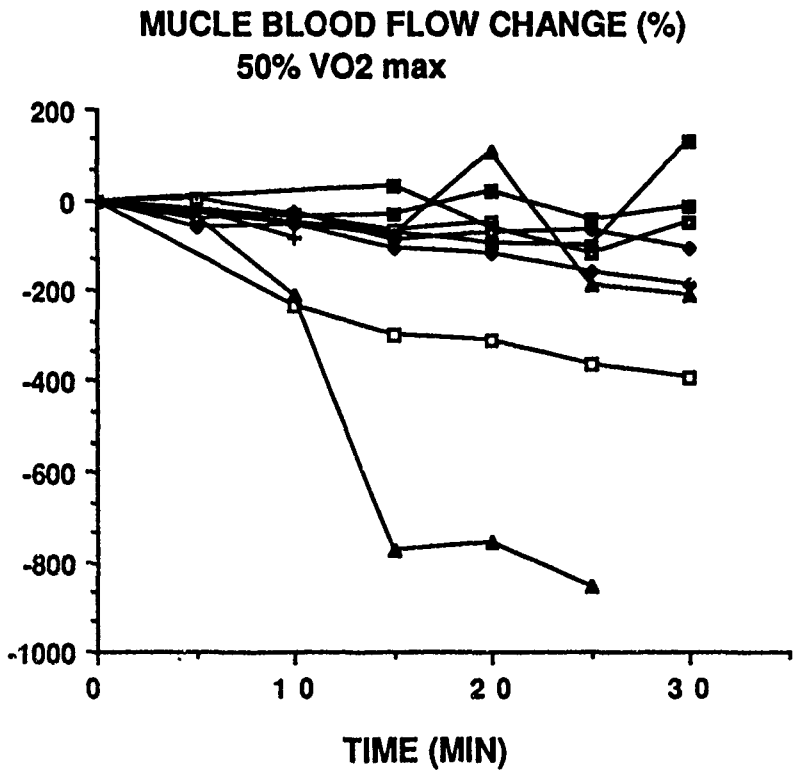
SKIN BLOOD FLOW CHANGE (%)
30% VO2 max

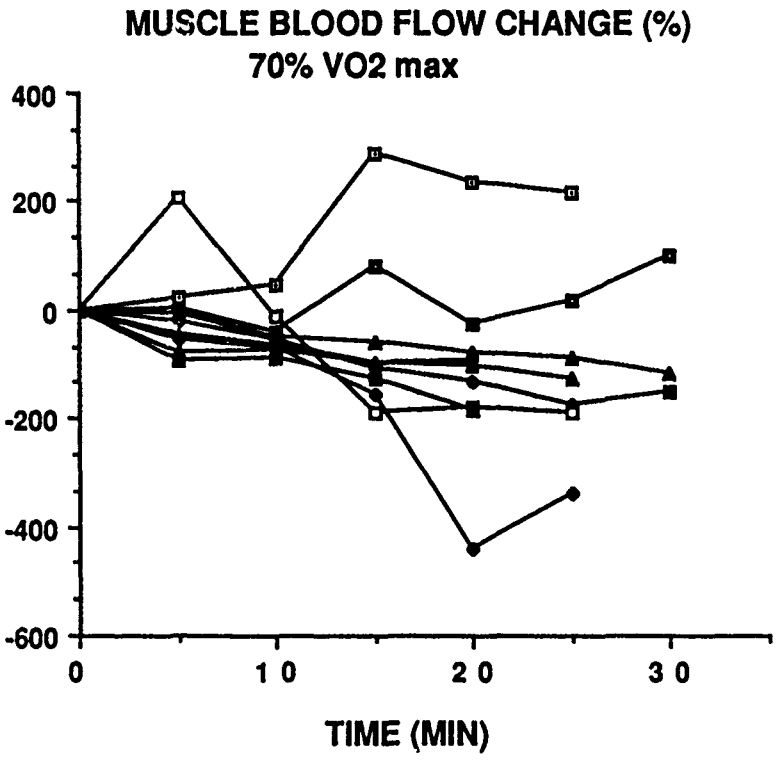


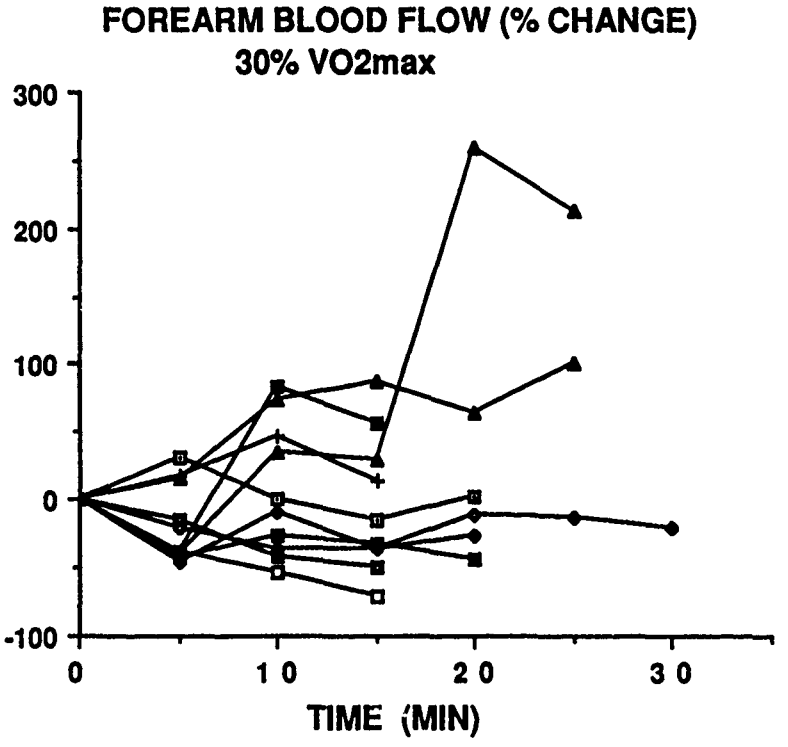
TIME (MIN)

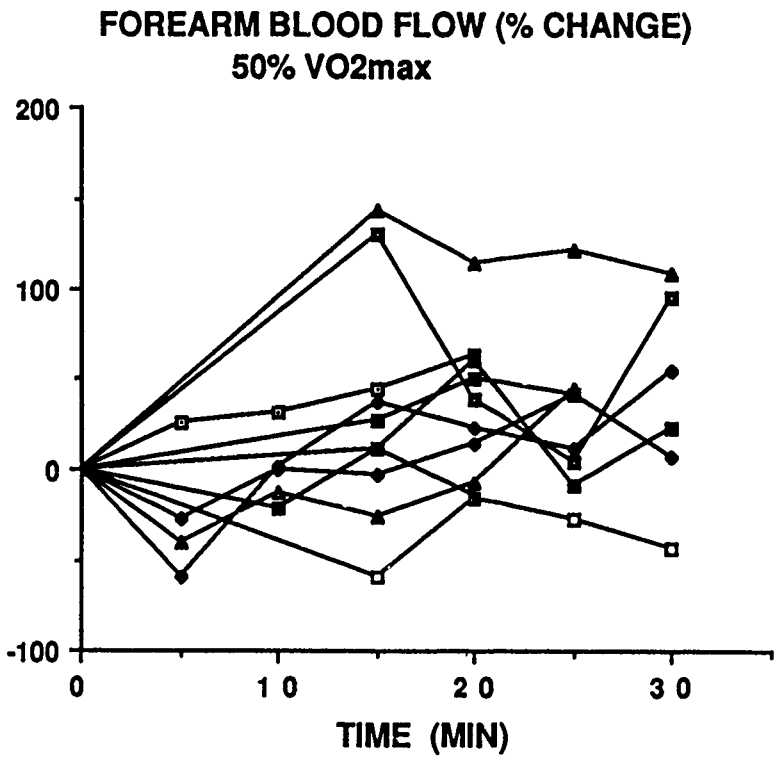

SKIN BLOOD FLOW (%)
70% VO2 max
500
300
100
-100
0
10
20
30
TIME (MIN)

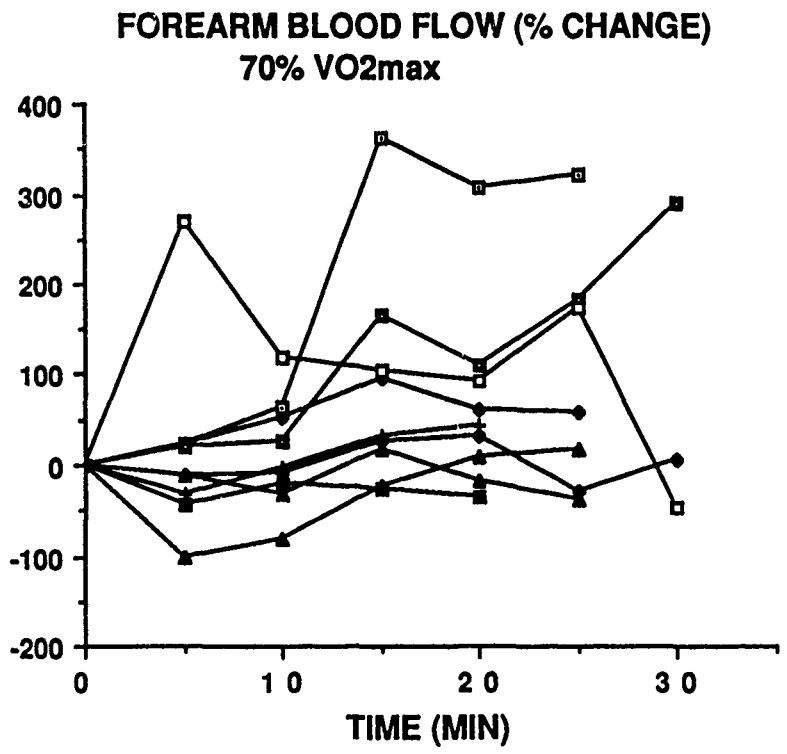

MUSCLE BLOOD FLOW CHANGE (%)
30% VO2 max
200
0
-200
-400
-600
-800
0
10
20
30
TIME (MIN)

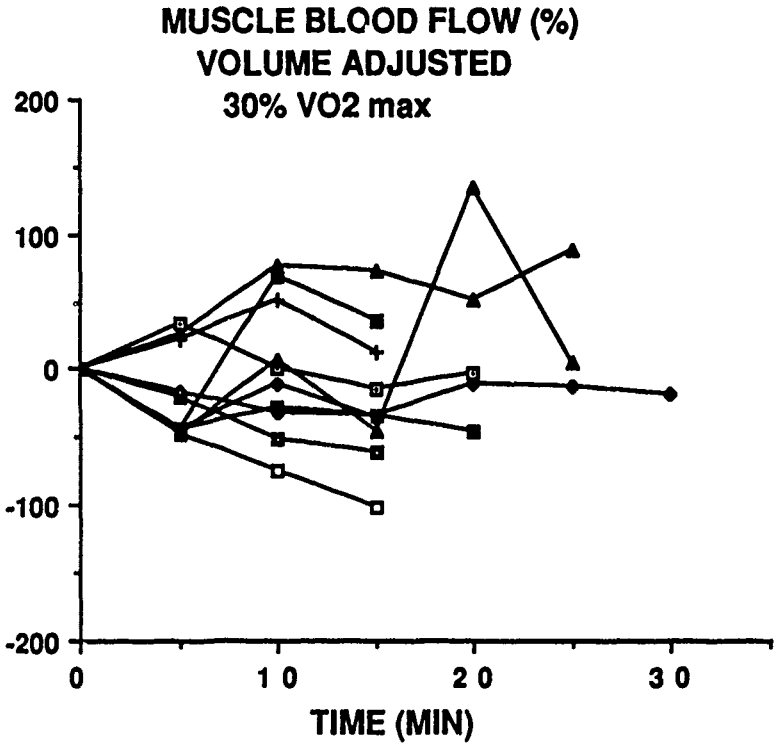

MUCLE BLOOD FLOW CHANGE (%)
50% VO2 max
200
0
-200
-400
-600
-800
-1000
0
10
20
30
TIME (MIN)

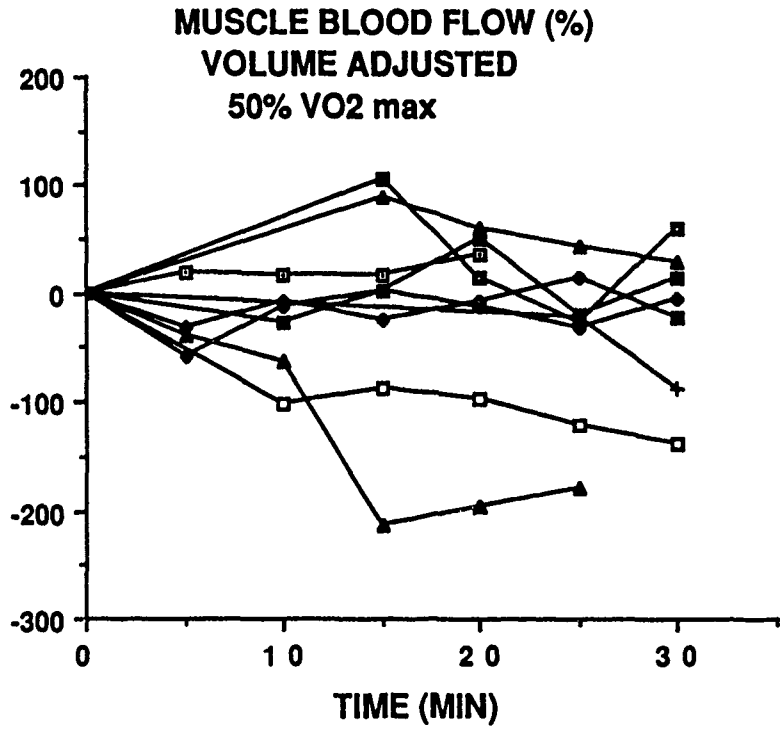

MUSCLE BLOOD FLOW CHANGE (%)
70% VO2 max
400
200
0
-200
-400
-600
0
10
20
30
TIME (MIN)

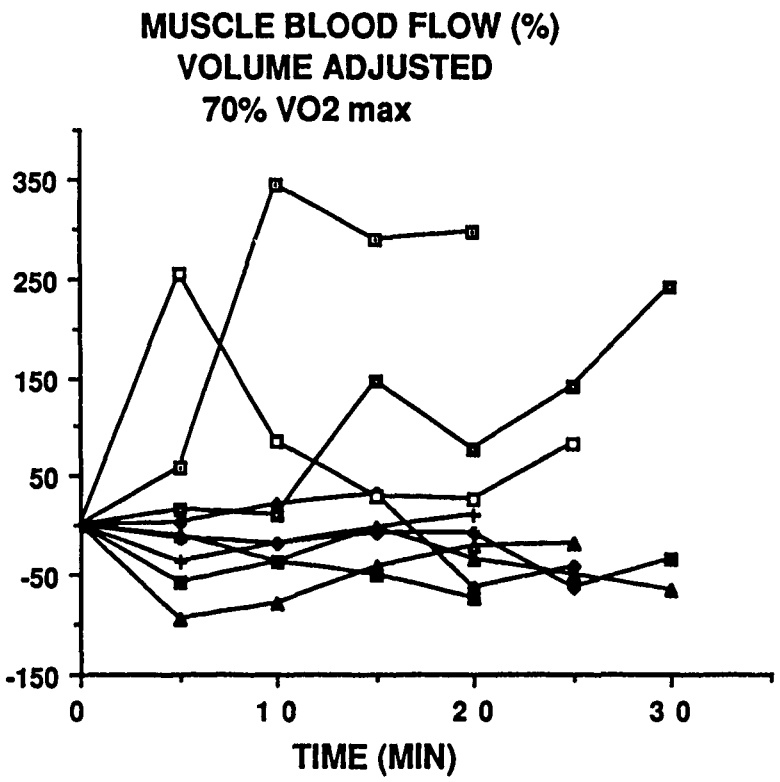

FOREARM BLOOD FLOW (% CHANGE)
30% VO2max
300
200
100
0
-100
0
10
20
30
TIME (MIN)


FOREARM BLOOD FLOW (% CHANGE)
50% VO2max
200
100
0
-100
0
10
20
30
TIME (MIN)


FOREARM BLOOD FLOW (% CHANGE)
70% VO2max
400
300
200
100
0
-100
-200
0
10
20
30
TIME (MIN)


MUSCLE BLOOD FLOW (%)
VOLUME ADJUSTED
30% VO2 max
200
100
0
-100
-200
0
10
20
30
TIME (MIN)


MUSCLE BLOOD FLOW (%)
VOLUME ADJUSTED
50% VO2 max
200
100
0
-100
-200
-300
0
10
20
30
TIME (MIN)


MUSCLE BLOOD FLOW (%)
VOLUME ADJUSTED
70% VO2 max
350
250
150
50
-50
-150
0
10
20
30
TIME (MIN)

1988 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

PHOTOPHYSICS AND PHOTOCHEMISTRY OF TRANSITION METAL COMPLEXES

Prepared by:	Dr. John A. Burke, Jr.
Academic Rank:	Professor
Department:	Chemistry
University:	Trinity University
	715 Stadium Drive
	San Antonio, TX 78212
Research Location:	Trinity University
USAF Researcher:	John Taboada, Ph. D.
	USAFSAM/NGOV
	Brooks AFB
	San Antonio, TX 78235-5301
Date:	December 27, 1989
Contract No:	F49620-87-004

PHOTOPHYSICS AND PHOTOCHEMISTRY OF TRANSITION METAL COMPLEXES

by

Dr. John A. Burke, Jr.

ABSTRACT

Synthesis of a series of new complexes of 8-[(pyridine-2-methylene)amino]quinoline, PMAQ, and methyl substituted PMAQ is reported. The oxidation of $Co(PMQA)_2^{2+}$ with H_2O_2 produces bis[N-8-(5,7-dichloroquinolyl)picolinamido]-cobalt(III), $[Co(DCQPA)_2]^+$, when isolated with HCl. The unchloronated analogue is produced when NH_4PF_6 is used as the precipitating agent instead of HCl. These materials have been characterized by elemental analyses, ir, uv/vis, magnetic moments, proton nmr and cyclic voltammetry. An x-ray molecular structure of $[Co(DCQPA)_2]Cl$ confirms an octahedral geometry for this complex and a similar geometry is expected for the rest of these complexes. Laser flash photolysis of $[Co(DCQPA)_2]^+$ in methanol solution reveals an excited state lifetime of 2.5 s when irradiated with a 355 nm excitation beam and probed with a colinear 633 nm beam.

ACKNOWLEDGEMENTS

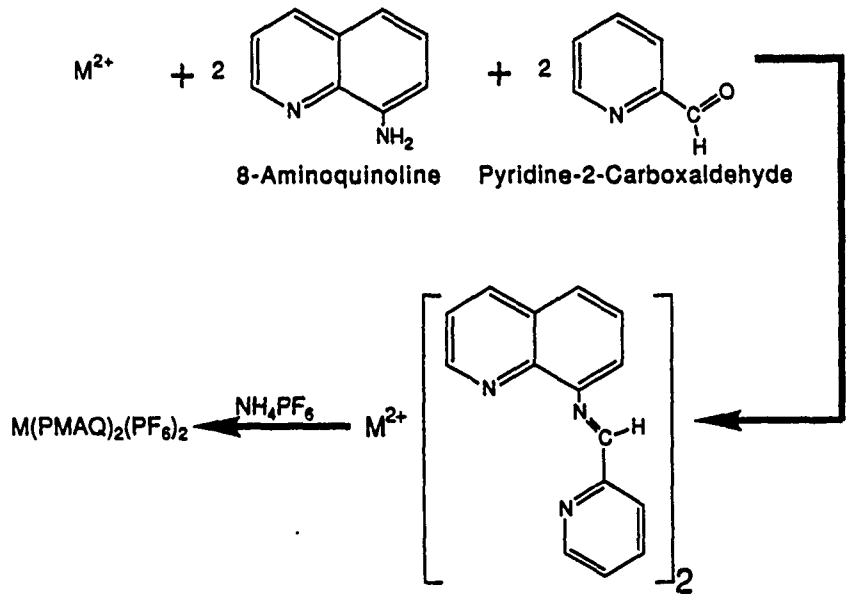
Initial work on the chemistry reported here was done at Brooks AFB during the summer of 1988 under the supervision of Dr. John Taboada, Ph. D. His support and encouragement to investigate these new materials has made the project both productive and enjoyable. Other collaborators must also be noted for the work they accomplished and other funding agencies also contributed to support students engaged in the project. Karin M. Keller is responsible for the synthetic work on $Co(DCQPA)_2Cl$ and the ruthenium compounds. During the summer of 1989 The Merck Company Foundation provided her research stipend. Bryan W. Hambric and Timothy A. Barckholtz are jointly responsible for the synthetic work on the other complexes and Timothy is primarily responsible for all the physical measurements on the iron and cobalt PMAQ complexes. Dr. Raymond E. Davis and his student, John Feagins, University of Texas at Austin, made an excellent contribution to our understanding by their structure determination of $Co(DCQPA)_2Cl$. Neil Leatherbury's stipend was provided by the NSF/SURE program on our campus. He developed the technique for the cyclic voltammetry and obtained the electrochemical data. Neil is a senior at Rice University. Dr. Patrick Holt and his students, Justin Glass and Michael Hobaugh, obtained the laser measurements.

I. INTRODUCTION

Coordination compounds which contain aromatic nitrogen donor ligands bound to transition metal ions exhibit a wide range of interesting energy related properties. Complexes of the inert metals ions, iron(II), cobalt(III), chromium-(III) and ruthenium(II), and polypyridine ligands such as 2,2′-bipyridine and 1,10-phenanthroline are the subjects of much of the early research in this area (1-8). Photosubstitution and photosolvation of the compounds are accessible because the dark reactions are very slow. These materials have been shown to store and transfer energy from the excited state that results on absorption of a photon of light in the visible region. Artificial photosynthesis using compounds of ruthenium(II) is actively being pursued (9). Our interest is in the synthesis and investigation of compounds that are similar to those of the polypyridines. This report describes the research on a class of transition metal complexes formed by the reaction of pyridine-2-carboxaldehyde with 8-aminoquinoline in the presence of the metal ion.

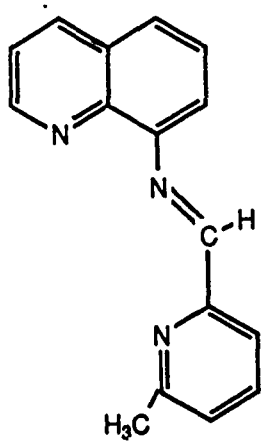
Reaction of 8-aminoquinoline bound to cobalt(II) and iron(II) with pyridine-2-carboxaldehyde under mild conditions in alcohol leads to the isolation of the metal ion complexes of 8-[(pyridine-2-methylene)amino]quinoline, here after referred to as PMAQ, by precipitation with an appropriate counter ion (equation 1). The methyl substitu-

ted PMAQ complexes have also been synthesized for the first time. The structures and abbreviations for these ligands are summarized in Figure 1.



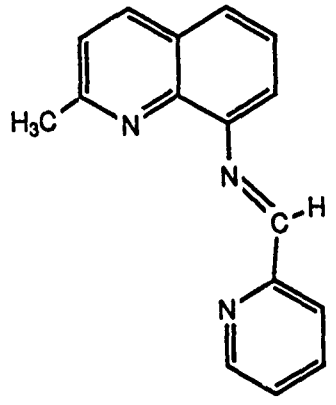
reaction 1)

The reaction of $Co(PMAQ)_2^{2+}$ with H_2O_2 has also been elucidated. The product of this reaction has been shown to be a cobalt(III) complex of oxidized PMAQ. Photophysical determination of the excited state lifetime for this material has been accomplished using laser spectroscopic techniques. In order to elucidate the structure of the ligand, a single crystal X-ray defraction structural analysis has been obtained.



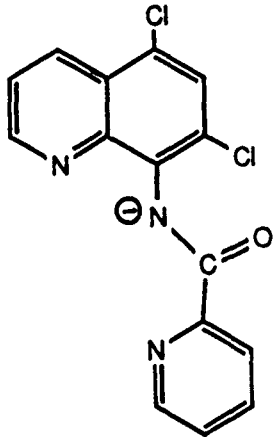
8-[(6-methylpyridine-2-methylene)amino]quinoline

Me-PMAQ



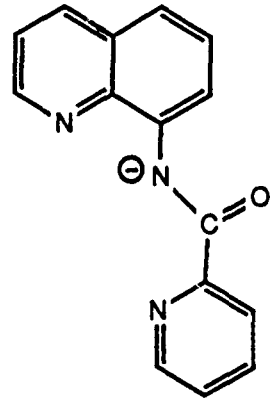
8-[(pyridine-2-methylene)-amino]quinaldine

PMAQ-Me



8-(5,7-dichloroquinolyl)-picolinamido

DCQPA



8-(quinolyl)-picolinamido

QPA

Figure 1. Structures of the ligands and the abbreviations used in the text to describe the complexes synthesized.

II. OBJECTIVES OF THE RESEARCH EFFORT

The oxidized cobalt complex has been shown to have an excited state absorption. This observation was made during the work reported in August, 1988. At that time the complex had not been characterized and was given the symbol CoSB1. When an methanol solution of CoSB1 was irradiated with 532 nm light from a YAG laser, the intensity of an orthogonal beam of 633 nm light from a He-Ne laser exiting the sample was observed to be a function of the power density of the 532 nm beam. Duplication of the synthesis of CoSB1 and its characterization by infrared and visible spectra, elemental analysis, cyclic voltammetry, conductivity and magnetic susceptibility were highest on the priority list for this project. The nature of the excited state of this material was also to be explored using laser techniques. The excited state lifetime and the spectral features of the phenomenon were of highest interest to evaluate the material's potential for application in laser devices.

An impure sample of a similar ruthenium complex had also exhibited this effect on orthogonal laser beams. Attempts to improve the synthesis and isolate and characterize this material had the next highest priority. In addition to these major objectives, tris(2,2'-bipyridine)-ruthenium(II) chloride was to be used to test and evaluate the optical system used in the laser beam experiments.

III. OXIDIZED COBALT COMPLEX

Synthesis of bis[N-8-(5,7-dichloroquinolyl)picolinamido]cobalt(III) chloride, $Co(DCQPA)_2Cl$. All materials used in the preparation of the compounds were obtained from Aldrich Chemical Company and were used without further purification. A solution of cobalt(II) choloride (5.31 g, 47 mmol) and 8-aminoquinoline (6.79 g, 47 mmol) in 85% aqueous methanol is stirred for 2-3 minutes. Pyridine-2-carboxaldehyde (10 ml, 100 mmol) is added to form a dark brown solution. Activated charcoal is added in catalytic amounts (approximately .01 g) before oxidation. Hydrogen peroxide (250 mL, 3% aqueous solution) is added in small portions and the solution is heated with stirring for 4 hr to complete the reaction and decompose the excess H_2O_2. The solution is filtered to remove the charcoal followed by the addition of 250 ml concentrated HCl. The mixture is evaporated at elevated temperatures. Some trials yield a deep red tar at this point and the volume is increased with water before the evaporation process is repeated. After isolation of the crystals by filtration they are purified by recrystallization from 5% aqueous isopropanol or isopropanol-ethylacetate mixtures. Yields of the recrystallized product range from 30 to 45%. Elemental analysis (Texas Analytical Laboratories, Inc.) calcd for $C_{30}H_{16}Cl_5CoN_6O_2$: C, 49.44; H, 2.22; N, 11.54; Co, 8.09; Cl, 24.32. Found: C, 48.94, H, 2.22; N, 11.44; Co, 7.34; Cl, 23.38.

Principal bands in the infrared spectrum of the product

obtained using a Perkin-Elmer 1600 series FTIR are found at 1654, 1604, 1563, 1508, 1489, 1386, 974 and 682 cm^{-1}. Molar absorptivities in the uv/vis spectrum measured in methanol solution using a Varian Cary 2315 Spectrophotometer are given in parenthesis after the wavelength maxima in nm: 486 (1220), 384 (19,100), 258 (53,900). Molar conductances (cm^2 $omh^{-1}mol^{-1}$) of 0.001 M solutions in two solvents were found as 91 in water and 98 in CH_3CN compared with ranges of 110-140 in water and 120-160 in CH_3CN normally found for 1:1 electrolytes in these solvents (10). Proton nmr spectra measured with a Varian VXR-300 MHz Spectrometer exhibit shifts for 8 protons in the aromatic region: 7.65 (1, t); 7.74 (1, d of d); 8.00 (1, s); 8.07 (1, s); 8.20 (2, m); 8.28 (1, d); 8.64 (1, d). The abbreviations s, d, t and m are used for singlet, doublet, triplet and multiplet, respectively. The samples are diamagnetic from measurement on a Johnson-Matthey Magnetic Susceptibility Appartus when corrected for the diamagnetism of the ligand. A single cyclic voltammetry wave is measured at -250 mv vs Ag/Ag^+ using a Pt disk working electrode attached to a BAS 100A Electroanalytical System. Voltammograms at scanning rates of 20 to 200 mv/s were measured on 4 mM solutions in CH_3CN with 0.1 M tetrabutylammoniumhexafluorophosphate as the electrolyte buffer.

Single crystals of sufficient size and quality for x-ray defraction analysis were grown from ethyl acetate-

ethyl ether mixtures. The molecular structure of the complex obtained working in collaboration with R. E. Davis, University of Texas - Austin, is illustrated in Figure 2. The structure is refined to an R value of 5%.

Laser flash photolysis was used to study the photophysical and photochemical behavior of $Co(DCQPA)_2Cl$. The third harmonic of a Nd:YAG laser (355 nm) excited the sample in solution and a nitrogen-pumped dye laser probed the excitation induced changes in the sample asorption. Both lasers delivered nanosecond pulses at 10 Hz and the overall temporal resolution of the apparatus is approximately 30 ns. The performance of the appartus was evaluated using the fluorecence of tris(bipyridine)ruthenium(II) prepared in our laboratories.

Experiments were conducted on 10^{-3} to 10^{-4} M solutions of $Co(DCQPA)_2Cl$ in methanol placed in a 5 cm pathlength sample cell. The optical path was oriented so that the excitation and probe beams were collinear but counterpropagating through the cell. Previous studies had shown that the excited system absorbed at 633 nm so probing was begun in this region of the spectrum. A rapid, instrument-limited rise in signal that coincided with the onset of excitation was observed. The signal showed no unusual early-time changes and it persisted for the maximum length measurable by the apparatus (1 ms).

Further investigation required the modification of the

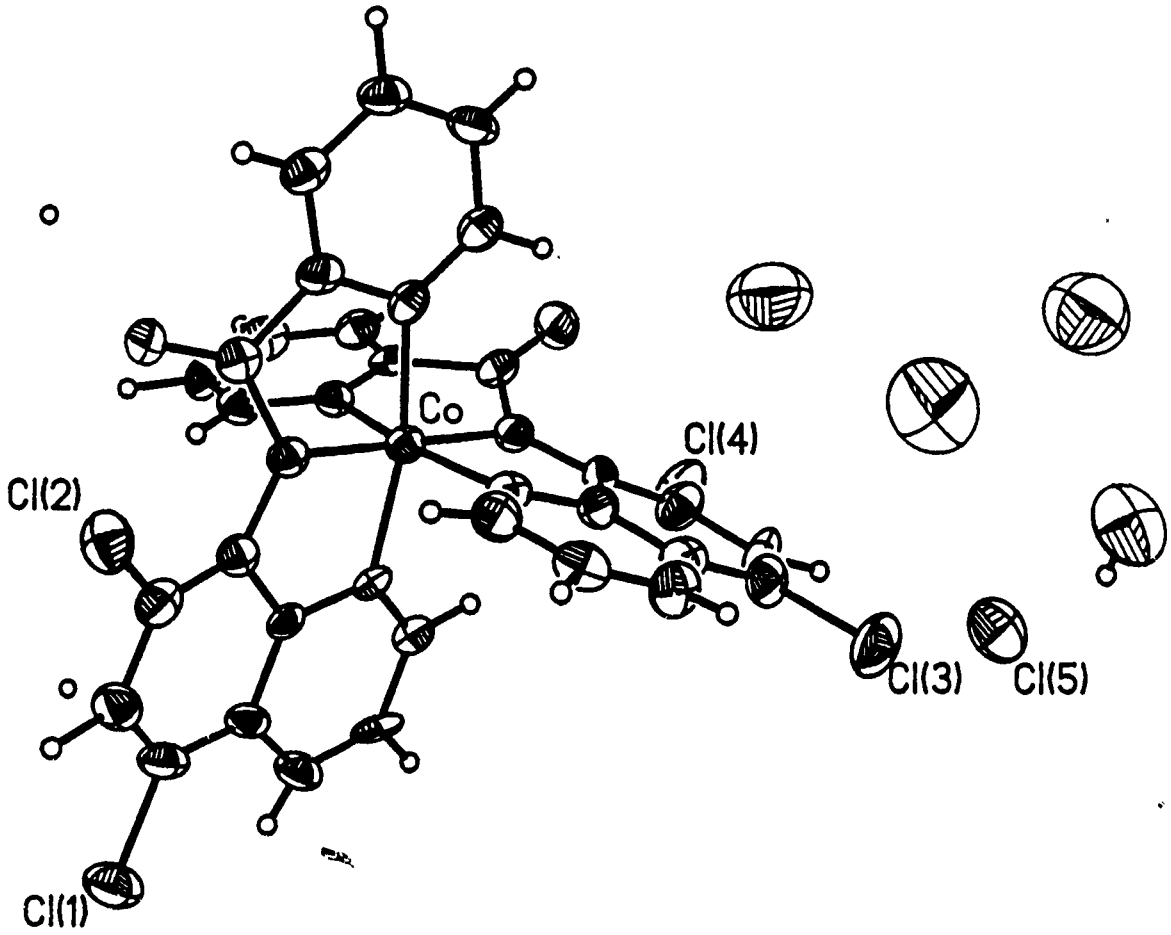


Figure 2. X-ray molecular structure of bis[8-(5,7-dichloroquinolyl)picolinamido]cobalt(III) chloride, $Co(DCQPA)_2Cl$.

apparatus illustrated in Figure 3. The dye laser was replaced with a 1 milliwatt, continuous-wave He-Ne laser as well as software modification to permit continuous data acquisition after each YAG laser shot. The output of the probe photodiode was also measured at the output of the sample cell using a storage oscilloscope. This arrangement allowed the intensity of the He-Ne beam exiting the sample to be measured as a function of time for periods on the order of several seconds.

Synthesis of bis[N-8-(quinolyl)picolinamido]cobalt(III) hexafluorophosphate, $Co(QPA)_2PF_6$. An analogous procedure to that for $Co(DCQPA)_2Cl$ produces the unchloronated complex if NH_4PF_6 is added to precipitate the product after oxidation with H_2O_2. Starting with 5 mmol of cobalt(II) chloride hexahydrate and 10 mmol of both 8-aminoquinoline and pyridine-2-carboxaldehyde gives $Co(QPA)_2PF_6$ in 74% yield. Significant peaks in the ir are observed at the following frequencies (cm^{-1}): 2364, 1644, 1604, 1503, 1468, 1397, 1359, 844, 763, 558. Shifts in the proton nmr are consistent with 10 protons on each ligand and are located in the aromatic region: 7.24-7.37 (2, m); 7.69 (1, d); 8.00-8.03 (2, m); 8.11-8.17 (3, m); 8.29 (1,d); 9.58 (1,d). The magnetic moment is 0.3 BM confirming that the cobalt is in the +3 oxidation state and the molar conductance in acetone is 118 $cm^2ohm^{-1}mol^{-1}$ as expected for a 1:1 electrolyte in this solvent (10). Cobalt analysis by atomic absorption on

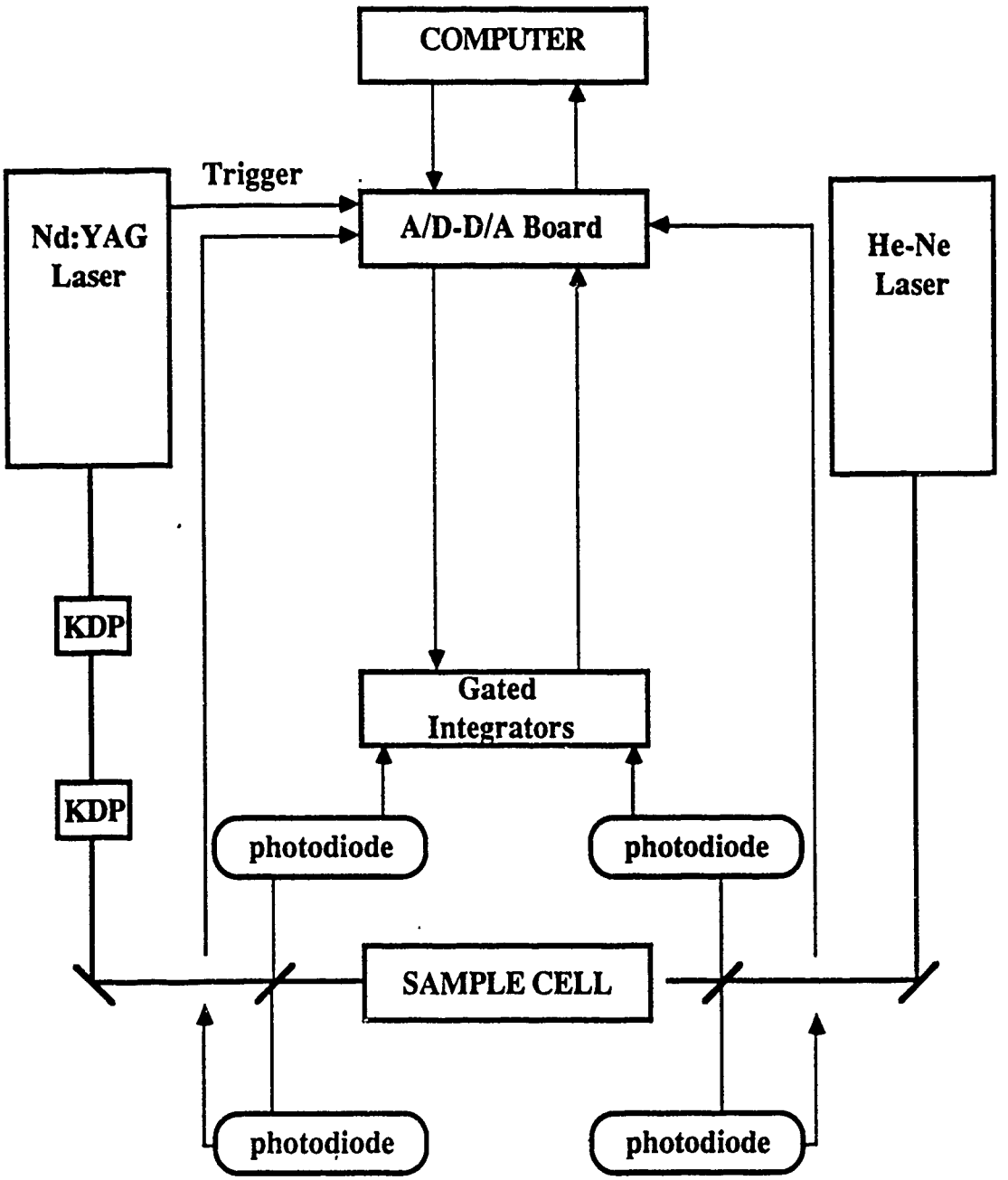


Figure 3. Diagram of the laser photolysis apparatus and data acquisition system.

a Varian AA Spectrometer is 8.4% Co compared with a calculated value of 8.41% for $C_{30}H_{16}CoN_6O_2PF_6$. The cyclic voltammogram exhibits one wave at -521 mv vs a Ag/Ag^+ reference electrode.

IV. METHYL SUBTITUTED PMAQ - IRON AND COBALT COMPLEXES

All of these materials were prepared in a manner diagrammed in equation 1. Early work with PMAQ was reported by Dwyer and co-workers (11). Our work modifies Dwyer's procedure in that the complex is produced using the metal ion as a template to produce the ligand. This procedure avoids the low yields observed when the ligand is prepared separately and then complexed to the metal ion. In the case of the iron(II) species iron filings were added to the reaction mixture to prevent oxidation of the metal and were filtered off just prior to precipitation with NH_4PF_6. These materials were characterized by elemental analyses, infrared spectra, proton nmr of the diamagnetic iron complexes, molar conductance and magnetic moments. The yields of all the materials were at least 60% of the theoretical. The uv/vis and cyclic voltammetry waves were measured in solution. Data for each complex is summarized below. All of this information is consistent with the formulation $ML_2(PF_6)_2$. An octahedral geometry around the metal ions by 6 nitrogen donor atoms is the expected structure for each of the complexes.

For $Fe(PMAQ)_2(PF_6)_2$ and the other compounds in this

series the elemental analyses were performed by Galbraith Laboratories. The other measurements were obtained as described above. Calcd. for $C_{30}H_{22}FeF_{12}N_6P_2$: C, 44.36; H, 2.73; N, 10.35; Fe, 6.87. Found: C, 44.35, H, 2.96; N, 10.15; Fe, 7.00. Principal bands in the ir (cm^{-1}): 1594, 1537, 1462, 1398, 1299, 1253, 1214, 840, 765, 558. Chemical shifts in the proton nmr: 7.18-7.11 (2,m), 7.65 (1,d), 7.90 (1,dd), 8.03 (1,d), 8.19-8.08 (2,m), 8.23 (1,d), 8.36 (1,d), 9.11 (1,d), 10.66 (1,s). Molar conductance measured in acetone is 184 and is within the range of 160-200 for 2:1 electrolytes in acetone. Three 1 electron, pseudo-reversible, cv waves occur at -930, -676, and 1158 mv vs Ag/Ag^+. UV/vis spectrum in methanol wavelength maxima in nm (molar absorptivity): 658 (5360), 563 (3390), 420-470 (4710), 372 (34,800), 352 (34,900).

For $Fe(Me\text{-}PMAQ)_2(PF_6)_2$ elemental analyses calcd. for $C_{32}H_{26}FeF_{12}N_6P_2$: C, 45.74; H, 3.12; N, 10.00; Fe, 6.65. Found: C, 46.45, H, 3.24; N, 10.15; Fe, 6.68. Principal bands in the ir (cm^{-1}): 1597, 1544, 1467, 1402, 1319, 1242, 1221, 843, 778, 558. Chemical shifts in the proton nmr: 2.05 (3,s), 7.49-7.58 (2,m), 7.92 (1,dd), 8.14 (1,d), 8.21 (1,dd), 8.33 (1,d), 8.58 (1,d), 8.85 (1,s), 9.23 (1,d), 12.57 (1,s). Molar conductance measured in acetone is 203. Two 1 electron, pseudo-reversible, cv waves occur at -1222, -990 mv vs Ag/Ag^+. UV/vis spectrum in methanol wavelength maxima in nm (molar absorptivity): 654 (3970), 558 (2880),

454-472 (3500), 373 (28,800), 355 (31,400).

For $Fe(PMAQ\text{-}Me)_2(PF_6)_2$ elemental analyses calcd. for $C_{32}H_{26}FeF_{12}N_6P_2$: C, 45.74; H, 3.12; N, 10.00; Fe, 6.65. Found: C, 46.25, H, 3.15; N, 9.97; Fe, 6.49. Principal bands in the ir (cm^{-1}): 1608, 1561, 1461, 1378, 1308, 1208, 838, 771, 556. Chemical shifts in the proton nmr: 2.08 (3,s), 7.78-1.84 (2,m), 8.07 (1,dd), 8.20 (1,d), 8.43 (1,dd), 8.65 (2,m), 9.47 (1,d), 10.17 (1,s), 13.88 (1,s). Molar conductance measured in acetone is 190. One 1 electron, pseudo-reversible, cv wave occurs at -992 mv vs Ag/Ag^+. UV/vis spectrum in methanol wavelength maxima in nm (molar absorptivity): 658 (6600), 562 (3850), 456 sh (3760), 372 (37,600), 356 (38,400).

For $Co(PMAQ)_2(PF_6)_2$ elemental analyses calcd. for $C_{30}H_{22}CoF_{12}N_6P_2$: C, 44.19; H, 2.72; N, 10.31; Co, 7.23. Found: C, 44.43, H, 2.79; N, 10.31; Co, 7.94. Principal bands in the ir (cm^{-1}): 1618, 1599, 1506, 1474, 1397, 1223, 841, 770, 558. Molar conductance measured in acetone is 200. The magnetic moment is 2.85 BM. Three 1 electron, pseudo-reversible, cv waves occur at -1182, -837, and -48 mv vs Ag/Ag^+. UV/vis spectrum in methanol wavelength maxima in nm (molar absorptivity): 416 sh (5860), 394 sh (8870), 355 (14,200), 312 sh (9030).

For $Co(Me\text{-}PMAQ)_2(PF_6)_2$ elemental analyses calcd. for $C_{32}H_{26}CoF_{12}N_6P_2$: C, 45.57; H, 3.11; N, 9.96; Co, 6.99. Found: C, 46.04, H, 3.08; N, 10.05; Co, 6.46. Principal

bands in the ir (cm^{-1}): 1627, 1599, 1504, 1467, 1399, 1219, 841, 764, 558. Molar conductance measured in acetone is 211. The magnetic moment is 4.69 BM. Two 1 electron, pseudo-reversible, cv waves occur at -1207 and -852 mv vs Ag/Ag^{+}. UV/vis spectrum in methanol wavelength maxima in nm (molar absorptivity): 355 (24,700), 262 (31,700).

For $Co(PMAQ\text{-}Me)_2(PF_6)_2$ elemental analyses calcd. for $C_{32}H_{26}CoF_{12}N_6P_2$: C, 45.57; H, 3.11; N, 9.96; Co, 6.99. Found: C, 45.88, H, 3.09; N, 9.96; Co, 6.95. Principal bands in the ir (cm^{-1}): 1622, 1596, 1508, 1479, 1374, 834, 770, 558. Molar conductance measured in acetone is 185. The magnetic moment is 5.00 BM. Three 1 electron, pseudo-reversible, cv waves occur at -1231, -814 and -346 mv vs Ag/Ag^{+}. UV/vis spectrum in methanol wavelength maxima in nm (molar absorptivity): 358 (17,300), 315 sh (11,70), 266 (28,500).

V. RUTHENIUM COMPLEXES

All attempts to prepare pure samples of the ruthenium complexes to date have only yielded complicated mixtures. Up to thirteen column chromatography bands have been observed in purification steps using alumina packed columns. As a result of this complication no additional information is available concerning these potentially interesting materials.

The preparations of $Ru(PMAQ)_2(PF_6)_2$ have been attempted in the same manner as those that are successful for the iron

and cobalt analogues. Ruthenium(III) chloride is used as the starting material and the hypophosphrous acid or oxalic acid are employed as reducing agents for the metal. Reaction times of as long as 72 hr are no more successful than shorter times in yielding tractible crude products.

VI. RECOMMENDATIONS

In as much as the materials reported here are all new, there is much work to be done to explore their properties and to evaluate their potential for useful applications. Similarities between these complexes and those of the polypyridine ligands can only suggest that the new substances will have the capability to store and transfer energy absorbed in the ultraviolet and visible region to other molecules and thus promote novel and important photochemistry. The scope of the work to date has shown that one of these substances, $Co(DCQPA)_2Cl$, is being promoted to an excited state by absorbing 355 nm laser light and that the excited state has a lifetime of 2.5 s. This compound is stable to irradiation under the conditions to which it has been subjected in the experiments described in this report. More experiments are planned after the conclusion of this project to elucidate the behavior of the excited $Co(DCQPA)_2Cl$ in methanol solution. We plan to study the transient signal as a function of the probe wavelength and to investigate the early time kinetics more throughly to search for additional, shorter-lived transients.

Our quest for a viable route to prepare the ruthenium analogues will continue. Reaction conditions that produce fewer by-products must be found before suitably pure ruthenium complexes can be successfully isolated and characterized.

Photophysical and photochemical evaluation of the iron and cobalt complexes of PMAQ, Me-PMAQ, and PMAQ-Me is also planned. The cobalt(II) compounds will be oxidized to ascertain what effect the methyl substituent has on the lifetime of the excited state. The iron compounds will be subjected to laser induced studies as well. Photochemical solvation and substitution in the coordination sphere for these materials will have to be investigated to determine their stability under irradiation.

REFERENCES

1. Ferraudi, G. J. Elements of Inorganic Photochemistry; John Wiley and Sons: New York, 1988; Chapter 6.

2. Meyer, T. J. Prog. Inorg. Chem. **1984**, 30, 389.

3. Vleck, A. A. Coord. Chem. Rev **1982**, 43, 39.

4. Crosby, G. A. J. Chem. Ed. **1983**, 60, 843.

6. Baker, B. R. and Mehta, B. D. Inorg. Chem. **1965**, 4, 848

7. Krause, R. A. Struct. Bonding **1987**, 67, 1.

8. Kumar, C. V., Barton, J. K., Gould, I. R., Turro, N. J. and Van Houten, J. Inorg. Chem. **1988**, 27, 648.

9. Meyer, T. J. Acc. Chem. Res. **1989**, 22, 163.

10. Angelici, R. J. Synthesis and Technique in Inorganic Chemistry, 2nd ed.; Saunders: Philadelphia, 1977, p213.

11. Dwyer, F.P.; Gill, N.S.; Gyarfas, E. C. and Lions, F. J. Am. Chem. Soc. **1953**, 75, 3334.

USAF-UES MINI GRANT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

SERUM SQUALENE AND CHOLESTEROL RATIO AS RISK PREDICTOR FOR CORONARY ARTERY DISEASE

Principal Investigator:	Dr. Hoffman H. Chen
Academic Rank:	Associate Professor
Department and University:	Department of Chemistry Grambling State University Grambling, Louisiana 71245
USAF Research Collaborator:	Dr. Harvey A. Schwertner
Research Location of Collaborator:	USAF SAM/NGP, Brooks AFB, San Antonio, Texas 78235-5301

Date:	June 30, 1988
Contract No.:	F49620-85-C-0013/SB5851-0360
Subcontract No.:	5-760-6MG-118
Duration:	One year (January 1, 1987 - December 31, 1987)

SERUM SQUALENE AND CHOLESTEROL RATIO AS RISK PREDICTOR OF CORONARY ARTERY DISEASE

ABSTRACT

The research reported herein was undertaken as a "Follow on" project which is designed to continue the effort of the Summer Faculty Research Program initiated during the Summer of 1986 at USAF School of Aerospace Medicine, Brooks AFB, San Antonio, Texas.

In order to improve the method for predicting heart disease, an analytical method has been devel[illegible]r the analysis of squalene in serums. The method consists of three fundamental steps: lipid extraction, separation of hydrocarbons by column chromatography, and squalene quantitation by gas chromatography. Gas chromatographic profiles of hydrocarbons in serums were analysed for 52 specimens collected from patients with and without coronary artery disease. The concentration of serum squalene was not depended on age and sex. The amount of squalene in serum (21-968 ng/ml) is relatively small comparing to that of serum cholesterol (111-369 mg/dl) and triglycerides (61-568 mg/dl). During preliminary studies, ratio of squalene to cholesterol showed no significant pattern as a risk Predictor. However, in some cases, the squalene concentrations were more related to the concentration of triglycerides in serum.

A. Introduction

The medical evaluations for predicting coronary disease (CAD) are required for the relative young Air Force flying personnel who are referred to as the Aeromedical Consultation Service. Abnormalities of lipoproteins have been found in patients with CAD, as well as in patients who have had a myocardial infraction (1). The ratio of total cholesterol to high density lipoprotein (HDC) cholesterol has been reported as a significantly useful laboratory predictors of CAD than total cholesterol or HDL - cholesterol alone (2). Recently Messeri et al reported that determinations of the amniotic fluid squalene/cholesterol ratio is a very useful diagnostic tool in clinical practice (3,4). From the analysis of serum samples during the 1980 summer research, squalene and other hydrocarbons were believed to be presented in the non-polar hydrocarbon fraction of lipoproteins after thin-layer chromatographic separation step (5). Since squalene is known as the direct precusor of cholesterol through a series of enzymatic reactions (6,7). The squalene concentration and its ratio to cholesterol in lipoprotein could conceivably be a new better risk predictor than the total/HDL cholesterol ratio.

The proposed study is a continuation of a research effort to explore the possible usefulness of new risk factor to enhance the evaluation of flying personnel seen by the Aeromedical Consultation Service. The primary mission of this study is to develop reliable laboratory analytical method for squalene and cholesterol in serum and to evaluate the diagnostic value of squalene and its ratio to cholesterol as new predictor of coronary artery disease. We hope this research will improve the specificity and sensitivity of current clinical methods for predicting heart disease in the flying personnel referred to this medical activity. Early detection would allow corrective action to be taken at an earlier time.

B. Research Objectives

the specific objectives of the proposed research are as follows:

1. Develop method for the quantitative analysis of squalene and cholesterol in serum lipoproteins.
2. Perform squalene and cholesterol analysis in serum from approximately twenty-five patients with and without coronary artery disease respectively.
3. Perform appropriate statistical analysis of results to determine if there are any significant differences in serum squalene and its ratio to cholesterol in the two populations. Determine the specificity and sensitivity of the squalene and squalene/cholesterol results. Establish and rank squalene and squalene/cholesterol values in importance of predictors of coronary artery disease.

C. EXPERIMENTAL METHOD

1. Chemicals

Squalene, n-eicosane and sodium dihydrogen phosphate were products of Pfaltz & Bauer, Inc., St. Waterbury, CT. Silica gel 40 (30-70 mesh ASTM) was purchased from MC/B Manufacturing Chemists, Inc. Spice silica gel sample preparation cartridge was obtained from Analtech, Inc. Other solvents were of analytical grade and were purchased from Aldrich Chemical Company, Inc.

2. Serum Specimens

Blood samples were drawn from patients after an overnight fast of approximately 12 hours. Blood samples are then permitted to clot, and serum was separated by centrifugation.

3. Lipid Extraction

Extraction of non-polar lipids in lipoprotein was by a modification of a method described by Slayback et al (10). About 1.0 ml of serum was used for the quantitative analysis of squalene and cholesterol. With a 1.0 ml graduated pipet, transfer the serum into labeled 25 ml centrifuge tube with a Teflon-lined screw cap. Record the volume of the serum sample. An aliquot of an internal standard solution of n-eicosane equivalent to 2 g was added to the serum. After vortexing the tube briefly, with a 1 ml graduated pipet, add 1 ml of 0.2M sodium dihydrogen phosphate buffer and swirl the tube. With a repipet, add 3 ml of ethyl acetate. Extract lipid by vortex and then shaking for 10 minutes on an oscillation shaker, then centrifuge for 5 minutes to separate the organic and aqueous phases (3/4 full speed). With a 5 ml graduated pipet, withdraw carefully the ethyl acetate phase into labeled 12 x 125 mm screw top culture tubes. Evaporate ethyl acetate under nitrogen gas at 40°C (water bath). The dry residue

was dissolved by the addition of 1 ml of hexane and vortex. The hydrocarbons were eluted with 3 ml of hexane on a column (3 cm long and 0.5 cm in diameter) silica gel 40 (30-70 mesh ASTM). The hexane solution was evaporated to dryness under a stream of nitrogen at 40°C. The residue was redissolved in 50 $\mu\ell$ of hexane immediately before injection into the gas chromatograph. Care was taken to maintain a nitrogen atmosphere in tube at all steps in the extraction procedure. The extraction procedure was verified by taking standard solutions of n-eicosane, squalene and cholesterol through the method and both n-eicosane and squalene were found to be quantitatively recovered.

4. Gas Chromatograph

A portion (1 to 2 $\mu\ell$) of the hexane extract was chromatographed on the glass column (6 feet x ¼ inch OD, 3% SE-30 on 80/100 Supelcoport glass column and fitted in a Tracor 560 gas chromatography. The flow rate of helium as 60 ml/minute and the temperature program was from 150° to 240°C at the increasing rate of 4°/minute. Injector and FID detector temperatures were 280°C and 240°C respectively. Areas of peaks were calculated by triangulation. There were slight differences in relative detector responses for compounds relative weight responses 1.05 and 1.00 for squalene and n-eicosane respectively and these were taken into account in calculations.

The identities of peaks designated as squalen were verified by co-injection method on three different columns (6 feet x ¼ inch OD, 3% OV-1 on 80/100 Supelcoport, 3% SE-30 on 80/100 Supelcoport, and 3% OV-1 on 80/100 chromosorb W.H.P.). The amount of squalene in 1 ml serum was too small to be verified by Dupont 321 mass spectrometer system coupled to the gas chromatograph with a 3% SE-30 on chromosorb W (6 feet x 2 mm ID).

5. Cholesterol and Triglycerides Analysis

Both cholesterol and triglycerides of serums were analyzed by KODAK EKTACHEM 700 Analyzer using KODAK EKTACHEM Clinical Chemistry slides for calorimetric test. 10 $\mu\ell$ serum was collected by standard venipuncture technique from patients fasting for at least 12, preferably 16 hours prior to drawing the samples. Serum was removed promptly from the clot or cells and analyze as soon as possible. Avoid repeated freeze/ thaw cycles.

D. Result and Discussion

Because of the relative concentrations of serum squalene (21-968 ng/ml), cholesterol (111-369 mg/dl), and triglycerides (61-568 mg/dl) were quite different, the low concentration of squalene and other hydrocarbons were separated by column chromatograph (eluted with hexane on silica gel) from cholesterol, triglycerides, and other high molecular weight lipids to make it easier for hydrocarbons analysis by gas chromatograph. The amounts of cholesterol and triglycerides were analysed independly by colorimetric method with Kodak EKTACHEM 700 analyzer. The direct injection of hydrocarbon extract onto a non-polar column in resolution of hydrocarbons into different classes on the basis of molecular weight. The identity of squalene peak was assigned on the basis of retention time data and co-chromatography with standards. There were slight differences in relative flame ionization detector responses for compounds (relative weight responses 1.05 and 1.00 for squalene and n-eicosane respectively) and these were taken into account in calculations. The extraction procedure was verified by taking standard solution of squalene and n-eicosane through the method and both hydrocarbons were found to be quantitatively recovered (100 = 5 percent). In general, serum squalene with the second highest retention time in the hydrocarbons extract was the minor compound. Mean value for squalene was 169 ng/ml among 52 serum specimens. The reproducibility of squalene analysis was investigated by taking 10 x 1.0 ml aliquots of a pooled serum through the method. The data obtained was squalene 103 ± 5 (mean ± SD), CV 4.85.

From the preliminary studies, the concentrations of serum squalene were not dependent on patients' age (18 to 96 years) and sex. However, in some cases, the squalene concentrations were related to the concentrations of triglycerides in serum. Probably due to the fact that both squalene and triglycerides are

believed to be derived from sebaceous glands (13, 14). It was noticeable for sample No. 29 that patient had not only the unusual high concentration of triglycerides (568 mg/dl) but also high concentration of squalene (752 ng/ml). Ratio of serum squalene to cholesterol showed no significant pattern in all ages.

The data presented here was preliminary in nature because of the small number of unscreened patients in the series. In addition, medication of patients could have an effect on skin lipids in serum. Therefore, it is recommended to screen the serum specimens more carefully to see the role of serum squalene and other hydrocarbons in relating to coronary heart disease.

Table 1. The Contents of Squalene, Cholesterol, and Triglyceride in Serums

Sample No.	Age	Sex	Squalene (ng/ml)	Cholesterol (mg/dl)	Triglyceride (mg/dl)	MI
1	18	M	247	175	95	No
2	20	M	53	240	98	No
3	21	F	42	195	52	No
4	21	M	196	142	233	No
5	24	M	50	138	107	No
6	26	M	307	136	62	No
7	27	F	63	111	93	No
8	28	F	340	202	118	No
9	28	F	286	191	152	No
10	31	F	103	213	181	No
11	32	F	32	163	112	No
12	32	M	62	172	162	No
13	33	M	190	196	74	No
14	34	F	288	164	182	No
15	39	F	41	193	77	No
16	39	M	54	226	188	No
17	40	F	212	248	73	No
18	42	F	129	179	72	No
19	43	F	86	208	67	No
20	45	F	93	211	166	No
21	46	M	90	201	122	No
22	46	M	96	219	123	No
23	51	F	125	202	35	No
24	54	M	84	242	170	No
25	55	M	150	182	240	No
26	58	M	69	304	251	No

Table 1. The Contents of Squalene, Cholesterol, and Triglyceride in Serums (continued)

Sample No.	Age	Sex	Squalene (ng/ml)	Cholesterol (mg/dl)	Triglyceride (mg/dl)	MI
27	58	F	968	222	144	Yes
28	59	F	206	369	195	No
29	59	M	752	223	568	No
30	62	F	139	173	78	No
31	60	F	91	189	117	No
32	65	F	85	195	138	No
33	65	F	65	237	90	No
34	69	M	168	224	75	Yes
35	69	M	344	183	172	No
36	71	F	233	151	203	No
37	72	F	150	315	168	Yes
38	73	F	70	260	281	Yes
39	75	F	93	275	129	No
40	75	M	218	153	88	Yes
41	76	F	89	206	115	No
42	78	M	155	176	139	No
43	76	F	85	183	97	Yes
44	79	M	183	170	127	No
45	81	M	125	131	86	No
46	82	F	124	217	105	Yes
47	83	M	172	188	218	No
48	83	M	156	147	259	No
49	82	M	21	202	90	Yes
50	84	F	176	206	194	No
51	93	F	252	116	54	No
52	96	F	213	213	61	No

E. Suggestions for Further Research

1. From the preliminary study on serum squalene concentrations of 52 patient samples with and without coronary artery disease, it was observed that the relative small amounts of squalene was the minor compound among other unidentified hydrocarbons. I would like to recommend to analyze and identify the detailed hydrocarbons profile in order to see if there is any other significant relationship of these hydrocarbons to the predictions of coronary artery disease in the flying personnel.
2. Although from the preliminary information of the limited number of patient samples, no significant pattern was observed that ratio of squalene/cholesterol could be a risk predictor for heart disease. However, by analysing the change of serum squalene and other hydrocarbons of the individual for a period of time might give a better picture to see any significant relations of squalene concentrations to the developing of heart disease.

ACKNOWLEDGEMENTS

Research sponsored by the Air Force Office of Scientific Research, Boilling AFB, DC, under Contract No. F49620-85-C-0013/5B5851-0360, Sub-contract No. 5-760-6MG-118.

The writer wishes to express his appreciation to the School of Aerospace Medicine, Brooks AFB, San Antonio, Texas for having selected him as a participant in its Summer Faculty Research Program (Summer 1986) during which time the research was initiated. My sincere gratitude goes to the research collaborator, Dr. Harvey A. Schwertner, and to members of the staff of the Technical Library at USAF/SAM for their invaluable assistance in conducting literature searches. I also wish to express my appreciation and gratitude to Dr. James R. Brown and his staff at Lincoln General Hospital, Ruston, Louisiana for their generous assistance in providing serum samples.

I am greatly indebted to those associated with the proposal review process at Universal Energy Systems, Inc., for selecting my Mini Grant Proposal among the ones which were funded as follow-on research projects.

F. References

1. L. Wallentin and B. Moberg, Lecithin-cholesterol Acyl Transferase Rate and High Density Lipoprotein in Coronary Artery Disease. Atherosclerosis 41:155, 1982.

2. G. S. Uhl, R. G. Troxler, J. R. Hickman, and D. Clark. Relation Between High Density Lipoprotein Cholesterol and Coronary Artery Disease in Asymptomatic Men. The American Journal of Cardiology. Vol. 48, 903 (1981).

3. G. Messeri and D. Billi, Amniotic Fluid Squalene and Gestational Age. Clinical Chemistry, Vol. 28, No. 8, 1810 (1982).

4. S. J. Wysocki, R. Hahnel, M. J. Millward, and D. T. Jenkins, Amniotic Fluid Squalene and Fetal Maturity. Br. J. Obstet. Gynaecol. 86, 854-860 (1979).

5. Hoffman Chen, Serum Phospholipid and Cholesterol Ester Fatty Acids As Risk Predictors for Coronary Artery Disease. The Final Report for 1986 USAF-UES Summer Faculty Research Program sponsored by the Air Force Office of Scientific Research conducted by the Universal Energy System, Inc.

6. K. S. Block, the Biological Synthesis of Cholesterol. Science, 150: 19-28 (1965).

7. J. W. Cornforth, R. H. Cornforth and K. K. Mathew, J. Chem. Soc., 2539 (1959).

8. R. G. D. Steel and J. H. Torrie, in "Principles and Procedures of Statistics." McGraw-Hill Book Co., New York, 1960, p. 194.

9. D. A. Clark, N. J. Garcia, P. R. Razell, and E. L. Mosser, Studies of the Accuracy of Measurement of Serum High Density Lipoprotein Cholesterol Levels. Aviat. Space Environ. Med. 1984; 55:941.

10. J. R. B. Slayback, L. W. Y. Cheung, and R. P. Geyer, Analytical Biochemistry, 83, 372 (1977).

11. W. B. Kannel, M. J. Garcia, P. M. McNamara, and G. Pearson, Serum lipid precursors of coronary heart disease. Hum. Pathol. 1971; 2: 129-51.

12. S. Yaari, S. Even-Zohar, U. Goldbourt and H. N. Neufeld, Associations of serum high density lipoprotein and total cholesterol with total, cholesterol with total, cardiovascular, and cancer mortality in a 7-year prospective study of 10,000 men. The Lancet, May 9, 1981, 1011-1015.

13. N. Nicolaides and S. Rothman, Journal of Investigative Dermatology, 24, 125 (1955).

14. R. S. Greene, D. T. Downing, P. E. Pochi, and J. S. Strauss, Journal of Investigative Dermatology, 54, 240 (1970).

Edit Systems

For Normal, Resting ECG Data Compression

Final Report - Preliminary Draft

for AFSOR Mini-Grant : A Feasibility Study for a Computerized ECG Database

FRANK HADLOCK

Computer Science Department, 615-372-3687

Tennessee Technological University

I. Introduction

The objectives of this mini-grant were to consider techniques for converting the microfiche ECG database at Brooks School of Aerospace Medicine to a digitized representation for computer storage and retrieval, to investigate techniques for data compression, and to investigate techniques for screeing abnormal from normal beats with very low probability of missing abnormal beats.

At the beginning of the contract period, Brooks School of Aerospace Medicine contracted with Quest Systems to do the conversion from microfiche to digitized format, using a PC-based system with optical frame grabber. The focus of this grant was on the development of techniques for data compression and for discriminating between normal and abnormal beats, to be applied to the computerized data base resulting from the conversion by Quest Systems.

One approach to data compression for ECG data is to use a template for the QRS complex (and possibly the T wave) formed through some statistical averaging, either over the population or over the individual. The difference between a QRS complex and the template is encoded and stored (to recover it is only necessary to decode the difference and add the template).

The approach taken in this study was motivated by the idea that an integrated approach to compression and interpretation should be taken, to allow for data base queries to extract all heart beats exhibiting a specified morphology. If the compression technique is one for which morphology can be determined directly from the compressed form, then retrieval will be much faster than if each beat in the data base must first be decompressed before determining whether it has the specified morphology. The basic approach investigated in this study is as follows:

1]. The ECG waveform is represented as a sequence of first differences (so that the waveform can be recovered from the initial value and the addition of first differences).

2]. First differences are quantized into $2n+1$ intervals of width δ. The $(n+1)st$ interval corresponds to differences:

$$|d| \leq \frac{\delta}{2}$$

the jth interval corresponding to differences:

$$-(n-j+1)\frac{\delta}{2} \leq d < -j\frac{\delta}{2}, \quad 1 \leq j \leq n$$

and the jth interval corresponding to differences:

$$(j-n)\frac{\delta}{2} < d \leq (j+1-n)\frac{\delta}{2}, \quad n < j \leq 2n+1$$

3]. The signal change is additively decomposed into it's *quotient* and *remainder* module δ. That is, a first difference d is represented as $d = j\delta + r$, with the result that a QRS complex and T-wave H may be represented by a pair of strings $\langle j, r \rangle$ with j being the string of interval indices or quotients, $j_1 j_2 \ldots j_w$ and r being the string of remainders $r_1 r_2 \ldots r_w$ where w is the width of H. Assuming that n and δ are chosen to accomodate the maximum slope occuring in an R-wave and the maximum negative slope in an S-wave, H can be recovered from $\langle j, r \rangle$ and the initial value.

4]. Assuming a resting ECG from a presumably healthy patient, the interval index strings j will cluster about a suitably chosen template string j^* and storing some string transformation $T : r \rightarrow r^*$ will result in a data compression of r from which the morphology of the corresponding heart beat can be determined.

II. String Transformations

Various notions of string transformations have be defined as a basis for the distance between two strings strings or sequences. The simplest is *Hamming distance*, which is applicable to equal length strings and has as basis the transformation of one string into another by *equal* and *unequal* character *substitution*. The *Damerau-Levenshtein* [6] metric can be viewed as an extension of Hamming distance, adding single character insertion/deletion operations to the substitution operations, and making possible the comparison of unequal length strings. For the purpose of this study, a more general notion of string transformation was developed [1] termed an *edit system*.

An *edit system* is defined as a 4-tuple $= < A, N, E, c >$ composed of an alphabet A, a set N of symbols used as character variables, a *finite* set of edit rules E, and a cost function c. Each edit rule e has an associated, non-negative cost $c(e)$ and is of the form $x \rightarrow y$ where x, y are strings belonging to $(A \cup N)^*$. A basic requirement is that any character variable occurring in y must have first occurred in x. We will use W, W' and W_i for character variables so that W can assume any value in A. Multiple occurrences of the same character variable such as W or W_i are related by the requirement that later occurrences are bound to the value assumed by a previous occurrence. The character

variable W' is related to that of W by requiring that W' assume any value not assumed by a previous occurrence of W.

Further definitions are needed before the concept of an edit system is complete. If x is a string over $A \cup N$ and u is a string over A with $|u| = |x|$, then u is an *instance* of x if $u_i = x_i$ when x_i is in A and if, for x_i, x_j in N, $x_j = x_i \Rightarrow u_j = u_i$. Also, $x_j = x_i' \Rightarrow u_j \neq u_i$. This ensures that multiple occurrences of a character variable are bound to the same value while W' cannot assume the same value as W. As examples, 00 and 11 are instances of WW and W_1W_2 but not WW' while 01 and 10 are instances of WW' and W_1W_2 but not WW.

If u is an instance of x, then the pair x/u defines a partial function $p_{x/u}$ or p from $N' = \{W|W = x_i \text{ for some } i\}$ into A by $p(x_i) = u_i$. As an example $p(W) = 0$ and $p(W') = 1$ for the partial function defined by $W1W'/011$. If y is a string over $A \cup N'$ and p is a partial function from N' into A, we extend function p to strings by defining $p(y)$ as the string obtained from y by taking characters from A as they occur and replacing character variables by their values as assigned by p. The partial function of the last example would thus produce 110 when applied to $W'1W$. The *application* of an edit rule $e : x \rightarrow y$ to a string u is possible *only if* u is an instance of x in which case the result is the string $e(u) = p_{x/u}(y)$.

To transform strings other than those which are an instance of the left hand side of an edit rule, we can adopt alternative conventions, the tradeoff being between the power of an edit system and the complexity of edit distance computation. For this study, the convention was adopted which has the effect that any characters produced by edit operations cannot be further edited. This may not seem like much of a restriction, but it precludes adjacent interchange, $W_1W_2 \rightarrow W_2W_1$, from being used more than once to move a character more than one character position.

We define an *item* to be a marked string of the form $w.z$ where w and z are strings over A. Rule $e : x \rightarrow y$ is *applicable* to item $w.z$ if $z = uv$ where u is an instance of x; in this case the result is the item $e(w.z) = wp_{x/u}(y).v$ with the dot moved to ensure that the next rule is applied to the first part of v. In this way we ensure that rules are applied in left to right fashion to a previously unedited portion of the string and that every part of the string is edited. A *transformation* $T : s \rightarrow t$ of a string s into a string t (both in A^*) is a sequence of edit rules $e_1, \ldots, e_k$ such that e_1 is applicable to item $h_1 = \Lambda.s$, e_2 is applicable to item $h_2 = e_1(h_1)$, etc., with e_i being applicable to item h_i and producing $h_{i+1} = e_i(h_i)$ and h_{k+1} being equal to $t.\Lambda$ (here Λ is the empty string).

Finally we turn to the edit costs as a basis for defining *edit distance*. The cost of an edit rule may be any non-negative number. As in simple Levenshtein distance, the cost $c(T)$ of a transformation T in a general edit system is defined to be the sum of the individual edit costs and the *edit distance* with respect to edit system S is defined to be $d_S(s,t) = \min\{c(T) \mid t \text{ is an edit sequence with respect to } S \text{ for which } T : s \rightarrow t\}$.

In this study, only rule costs were considerd which *conform* to Levenshtein distance in the following sense. Simple Levenshtein distance is based on edit costs that tend to conserve characters: a cost of 1 is associated with insertion, deletion, and unequal substitution. If $d_L(x,y)$ denotes the simple Levenshtein distance between strings x and y over A, then $d_L(x,y)$ is the minimum transformation cost required to transform x into y, a transformation being a sequence of edit operations (over insertion, deletion, equal or

unequal substitution) and having cost equal to the sum of the individual edit costs. Rule costs *conform* to Levenshtein distance if $c(e) \geq d_L(u, e(u))$ for rule $e : x \rightarrow y$ and for any string u an instance of x (in which case $e(u) = p_{x/u}(y)$). For example, $WW \rightarrow W$ and $e : W \rightarrow WW$ must have cost ≥ 1 while $W_1W_2 \rightarrow W_2W_1$ may have cost 0.

Of particular interest were *edit systems* which consist of some combinations of the operations of insertion ($\Lambda \rightarrow W$), deletion ($W \rightarrow \Lambda$), equal substitution ($W \rightarrow W$), unequal substitution ($W \rightarrow W'$), compression ($WW \rightarrow W$), projection ($W_1W \rightarrow W_1$) and ($W_1 \rightarrow W_1W$), and expansion ($W \rightarrow WW$).

An *expansion system* will be one for which strings may only be transformed into strings of equal length or longer. A *compression system* will be one for which strings may only be transformed into strings of equal length or shorter, while a *complete system* is one for which any string can be transformed into any other. Some specific systems are shown to have applications to *data compression*. An efficient algorithm for computing the least cost transformation $T : r \rightarrow r^*$ of the interval index string r for a heartbeat into a template string r^* was developed and appears in [1].

III. Applications of Compression/Expansion Systems to Data Compression

It is anticipated that general edit systems will have applications to the same sorts of problems to which Levenshtein distance has been applied, and perhaps more accurately model primitive operations comprising string transformation. These applications are numerous; they include pattern recognition, spelling correction, data compression and error correction in variable length codes. A survey of applications is contained in [7].

In this study, the interest is in the application of edit systems to data compression and pattern recognition. First is given a sketch of an approach to using edit systems for data compression. In [4] the author gives an algorithm for constructing the longest common subsequence and the shortest common supersequence of a set of strings. For a set of very similar strings, such as the heartbeats in a resting ECG strip of a healthy patient, an effective compression technique is to use the LCS or SCS as template and store the transformations of each string in the set into the template j^*. Now each string can be transformed into the the longest common subsequence using only the operations of deletion and equal substitution, and can be transformed into the shortest common supersequence using only the operations of insertion and equal substitution. This technique can be extended to a larger set of operations. For sake of example, consider the simple expansion system:

$$E = \langle e_1 : W \rightarrow WW,\ e_2 : \Lambda \rightarrow W,\ e_3 : W \rightarrow W \rangle$$

To illustrate the compression process, consider the two strings j^* and j with transformation $T : j \rightarrow j^*$:

$$j^* = bbeeeeeaaaaaaaddd$$

$$j = beeeeaaaaaaddd$$

$$T = e_1e_3 \ldots e_3e_1e_3 \ldots e_3e_1e_3e_3$$

Define the *inverse of an edit system* as being another edit system which provides the inverse to any transformation provided by the first system in the following sense: there is a 1-1

onto function $f : E \rightarrow E'$ such that if T is a transformation over E with $T : s \rightarrow t$, then $f(T)$ is the corresponding transformation over E' for which $f(T) : t \rightarrow s$. As an example, the inverse of the expansion system E above is:

$$E' = \langle e'_1 : WW \rightarrow W,\ e'_2 : W \rightarrow \Lambda,\ e'_3 : W \rightarrow W \rangle$$

with $f(e_i) = e'_i$.

Now if Y is a set of strings all transformable into a string X under some edit system with an inverse edit system, then Y may be represented by string X along with the inverse transformations. Since the underlying assumption is that the strings, along with the template, are very similar, the *equal substitution* operation will be the most prevalent. For this example, the remaining inverse operations, e'_1, e'_2 of compression and deletion are encoded with a single bit. Then a run (of *equal substitutions*) length encoding scheme gives a storage requirement of $k \cdot (log_2(|j^*|) + 1)$ where k is the number of operations $\neq$ *equal substitution* required to perform the inverse transformation of the LCS j^* into j. For the above example, storage requirements are $3 \cdot (4 + 1) = 20$ bits versus $16 \cdot log_2(2n + 1)$ bits where $2n + 1$ is the number of quantiles or difference intervals.

IV. Summary

The support provided by the AFSOR mini-grant partially supported the work detailed in [1] ,[2], and [3], copies of which are attached. The objectives of the contract have only been partly achieved, due to the unavailability of digitized resting ECG data for healthy patients. When this data becomes available, the compression techniques will be applied and the results incorporated into the final draft of this report.

Although there are many techniques for data compression, the technique described in this report has the property that it is based on transformations of compressed beats into a template string, with the basic operations employed by the transformation corresponding to small, local changes in width, slope and morphology in general. For example, a notch due to block could be modeled by a rule/rules in the edit system allowing the substitution of a notch for a peak, which would have the obvious counterpart/counterparts in an inverse edit system. Since the compressed beats are the inverse transformations (to transform the template into the original beat), all beats with block can be retrieved by direct search for these transformations.

REFERENCES

1. HADLOCK, F. O. (1988), General Edit Systems for String or Sequence Transformation, *Proceedings of SE Region ACM Conference, 1988*
2. HADLOCK, F. O. (1988), Minimum Detour Methods for String or Sequence Comparison, *Congressus Numerantium.*
3. HADLOCK, F. O., (1988), An Efficient Algorithm for Pattern Detection and Classification, *Proceedings of The First International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems IEA / AIE -88*

4. HADLOCK, F. O., Algorithms for Longest Common Subsequence and Shortest Common Supersequence, *Congressus Numerantium, No.*.
5. HART, P. E., NELSON, N. G., and RAPHAEL, B., (1978), A formal basis for the heuristic determination of minimum cost paths. *IEEE Trans. Syst. Sci. Cybern*, Vol. SSC-4, 100-107.
6. LEVENSHTEIN, V. I. (1966), Binary codes capable of correcting deletions, insertions, and reversals. *Cybernetics and Control Theory 10*, 707-710.
7. SANKOFF, D. AND KRUSKAL, G. B. (Eds.), (1983), Time warps, string edits, and macromolecules: The theory and practice of sequence comparison. Addison Wesley, Inc.
8. WAGNER, R. A. AND FISCHER, M. G., (1974), The string to string correction problem. *JACM 21*, 168-173.

General Edit Systems

For String or Sequence Transformation

Frank Hadlock

Tennessee Technological University

I. Introduction

Various notions of edit distance between two strings have been defined, the simplest being that of *Hamming distance*, which can be thought of as being the number of unequal substitutions needed to transform one binary string into another. The *Damerau-Levenshtein* [6] metric can be viewed as an extension of Hamming distance, adding single character insertion/deletion operations to the substitution operations, and making possible the comparison of unequal length strings. *General Levenshtein distance* [8] employs arbitrary, nonnegative costs as a function of the edit operation and the character operands. In this paper, we consider a more general notion of string transformation which will be termed an *edit system.*

We define an *edit system* as a 4-tuple $= < A, N, E, c >$ composed of an alphabet A, a set N of symbols used as character variables, a *finite* set of edit rules E, and a cost function c. Each edit rule e has an associated, non-negative cost $c(e)$ and is of the form $x \rightarrow y$ where x, y are strings belonging to $(A \cup N)^*$. A basic requirement is that any character variable occurring in y must have first occurred in x. We will use W, W' and W_i for character variables so that W can assume any value in A. Multiple occurrences of the same character variable such as W or W_i are related by the requirement that later occurrences are bound to the value assumed by a previous occurrence. The character variable W' is related to that of W by requiring that W' assume any value not assumed by a previous occurrence of W.

Further definitions are needed before the concept of an edit system is complete. If x is a string over $A \cup N$ and u is a string over A with $|u| = |x|$, then u is an *instance* of x if $u_i = x_i$ when x_i is in A and if, for x_i, x_j in N, $x_j = x_i \Rightarrow u_j = u_i$. Also, $x_j = x_i' \Rightarrow u_j \neq u_i$. This ensures that multiple occurrences of a character variable are bound to the same value while W' cannot assume the same value as W. As examples, 00 and 11 are instances of WW and W_1W_2 but not WW' while 01 and 10 are instances of WW' and W_1W_2 but not WW.

If u is an instance of x, then the pair x/u defines a partial function $p_{x/u}$ or p from $N' = \{W | W = x_i \text{ for some } i\}$ into A by $p(x_i) = u_i$. As an example $p(W) = 0$ and $p(W') = 1$ for the partial function defined by $W1W'/011$. If y is a string over $A \cup N'$ and p is a partial function from N' into A, we extend function p to strings by defining $p(y)$ as the

string obtained from y by taking characters from A as they occur and replacing character variables by their values as assigned by p. The partial function of the last example would thus produce 110 when applied to $W'1W$. The *application* of an edit rule $e : x \rightarrow y$ to a string u is possible *only if* u is an instance of x in which case the result is the string $e(u) = p_{x/u}(y)$.

To transform strings other than those which are an instance of the left hand side of an edit rule, we can adopt alternative conventions, the tradeoff being between the power of an edit system and the complexity of edit distance computation. In this paper, the convention adopted has the effect that any characters produced by edit operations cannot be further edited. This may not seem like much of a restriction, but it precludes adjacent interchange, $W_1W_2 \rightarrow W_2W_1$, from being used more than once to move a character more than one character position.

We define an *item* to be a marked string of the form $w.z$ where w and z are strings over A. Rule $e : x \rightarrow y$ is *applicable* to item $w.z$ if $z = uv$ where u is an instance of x; in this case the result is the item $e(w.z) = wp_{x/u}(y).v$ with the dot moved to ensure that the next rule is applied to the first part of v. In this way we ensure that rules are applied in left to right fashion to a previously unedited portion of the string and that every part of the string is edited. A *transformation* $T : s \rightarrow t$ of a string s into a string t (both in A^*) is a sequence of edit rules $e_1, \ldots, e_k$ such that e_1 is applicable to item $h_1 = \Lambda.s$, e_2 is applicable to item $h_2 = e_1(h_1)$, etc., with e_i being applicable to item h_i and producing $h_{i+1} = e_i(h_i)$ and h_{k+1} being equal to $t.\Lambda$ (here Λ is the empty string).

Finally we turn to the edit costs as a basis for defining *edit distance*. The cost of an edit rule may be any non-negative number. As in simple Levenshtein distance, the cost $c(T)$ of a transformation T in a general edit system is defined to be the sum of the individual edit costs and the *edit distance* with respect to edit system S is defined to be $d_S(s,t) = \min\{c(T) \mid t \text{ is an edit sequence with respect to } S \text{ for which } T : s \rightarrow t\}$.

In this paper we will only consider rule costs which *conform* to Levenshtein distance in the following sense. Simple Levenshtein distance is based on edit costs that tend to conserve characters: a cost of 1 is associated with insertion, deletion, and unequal substitution. If $d_L(x,y)$ denotes the simple Levenshtein distance between strings x and y over A, then $d_L(x,y)$ is the minimum transformation cost required to transform x into y, a transformation being a sequence of edit operations (over insertion, deletion, equal or unequal substitution) and having cost equal to the sum of the individual edit costs. Rule costs *conform* to Levenshtein distance if $c(e) \geq d_L(u, e(u))$ for rule $e : x \rightarrow y$ and for any string u an instance of x (in which case $e(u) = p_{x/u}(y)$). For example, $WW \rightarrow W$ and $e : W \rightarrow WW$ must have cost ≥ 1 while $W_1W_2 \rightarrow W_2W_1$ may have cost 0.

Of particular interest will be *edit systems* which consist of some combinations of the operations of insertion ($\Lambda \rightarrow W$), deletion ($W \rightarrow \Lambda$), equal substitution ($W \rightarrow W$), unequal substitution ($W \rightarrow W'$), adjacent interchange ($W_1W_2 \rightarrow W_2W_1$), compression ($\bar{W}\bar{W} \rightarrow \bar{W}$), projection ($W_1W \rightarrow W_1$) and ($W_1 \rightarrow W_1W$), and expansion ($W \rightarrow WW$).

An *expansion system* will be one for which strings may only be transformed into strings of equal length or longer. A *compression system* will be one for which strings may only be transformed into strings of equal length or shorter, while a *complete system* is one for which any string can be transformed into any other. Some specific systems are shown to have applications to *data compression*.

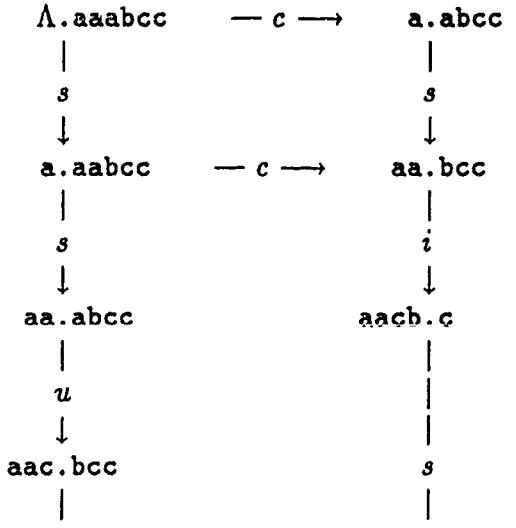
A transformation graph is defined for any edit system, and the edit distance is defined to be the least cost transformation sequence, corresponding to a shortest path from initial vertex to goal/final vertex. A shortest path algorithm, referred to by the author as the Minimum Detour algorithm, is adapted to the task of computing the edit distance. It is a 'tuned version' of the A^* algorithm, assuming a monotonic bound on the remaining distance, and a small set of edge costs. It is shown that for some edit systems, the algorithm is $O(dn)$ while for others it is $O(mn)$ where d is the distance and m and n are the string lengths with $m \leq n$.

II. Transformation Graphs and Distance Computation

In order to model all transformation sequences which can potentially transform source string s into target string t, a *transformation graph* G is defined with vertices corresponding to items and edges corresponding to edit operations. The *initial vertex* of G is the item $\Lambda.s$ while the *goal vertex* is $t.\Lambda$. In general, a vertex corresponds to an item $w.z$ where w is a prefix of t and z a suffix of s. If $s = s'z$ and $t = wt'$, then s' has been transformed into w and it remains to transform z into t'. Application of rule $e : x \rightarrow y$ to item $w.z$ is possible *only* if $z = uv$ such that u is an instance of x in which case the result is item $wp_{x/u}(y).v$ so that $p_{x/u}(y)$ must be a prefix of the remainder of the target t'. If so, there is an edge from the vertex for $w.z = w.uv$ to the vertex $wp_{x/u}(y).v$. It is clear that a transformation of s into t corresponds to a path in G from initial to goal vertex. Figure 1 provides an illustration.

$s = \mathtt{aaabcc},\ t = \mathtt{aacbc}$

$E = \{WW \rightarrow W,\ W_1W_2 \rightarrow W_2W_1,\ W \rightarrow W',\ W \rightarrow W\}$



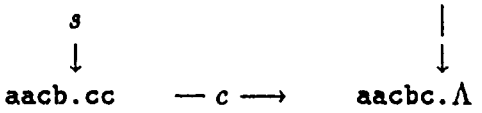


Figure 1. A Transformation Graph

For this source–target pair, there are three possible transformations with respect to the edit system, corresponding to the three paths from initial to goal vertex. The edit operations have been labeled c, i, u and s for compression, interchange, unequal substitution and substitution. The labels on the edges indicate the operations represented. If a cost system is used which assigns 1 to each operation except $W \rightarrow W$ which has cost 0, then any path from initial to goal vertex has cost 2. It is also clear that *path length* equals *transformation cost* and that the edit distance $d_S(s,t) = 2$.

III. The Minimum Detour Shortest Path Algorithm

In order to determine edit distance, it suffices to find the shortest path in the transformation graph from initial to goal vertex. For this purpose, we may use the Minimum Detour shortest path algorithm, first introduced by the author for finding shortest paths in grid graphs [1] and then applied to finding Levenshtein distance [2]. It may be regarded as a 'tuned' version of the heuristic search A^* algorithm [5]. The basis for selecting it over Dijkstra's algorithm is that a *lower bound function* $h(u)$ is easily computed for the transformation graph; the function is then used by the Minimum Detour algorithm to establish a sense of direction, guiding the search so that there are *never more* vertices expanded than are expanded by Dijkstra's algorithm.

The basic definitions are repeated here for the sake of completeness. Let $h(u)$ be a lower bound to the length of the shortest path from u to a goal vertex with the property that the change in h, $h(u) - h(v)$, observed in moving along an edge from u to v, never exceeds the edge cost $c(u,v)$ (monotonic condition). On a local level, it is desired to reduce the lower bound by the maximum amount so the *detour*, d_{uv}, of v with respect to u, is defined to be

$$d_{uv} = c(u,v) - (h(u) - h(v)) \text{ , or}$$
$$c(u,v) + h(v) - h(v).$$

The detour t_p of a path P is defined to be the sum of the detours of all but the initial vertex:

$$d_p = \sum d_{u_i u_{i+1}} \text{ where } P = u_1, u_2 \ldots u_n.$$

It is easily established [2] that path length equals path detour plus a constant $(h(u_1) - h(u_n))$ so that, since monotonicity implies path detour is non-negative, we have:

Theorem. *A minimum detour path from initial to goal vertex in the transformation graph*

is a shortest path and has length equal to the edit distance.

Assuming for the moment a lower bound function, the following algorithm can then be used to determine edit distance. The algorithm is the Minimum Detour algorithm [1], [2] and [3], tailored for finding minimum detour paths in transformation graphs for general edit systems. The vertex set may be coordinatized using integer coordinate pairs (i,j) where i ranges from 0 to $|t|$ (i is the length of the target prefix w of t produced already by the edit sequence) and j from 0 to $|s|$ (j is the length of the source prefix being the complement of z and which has been edited to produce w).

An array Label, indexed by these coordinates and initialized to False at compile time, will be used to indicate whether or not a vertex has been previously expanded. To store the unexpanded vertices on the frontier list, the algorithm employs a series of stacks indexed by the detour of the path from the initial vertex to the frontier vertex. Assuming a lower bound function $h(i,j)$ to the edit distance from vertex (i,j) to goal vertex (m,n) (edit distance between source and target suffixes), the following algorithm computes the minimum detour path (and thus the edit distance):

A. Minimum Detour Algorithm

```
{Let m = |t| and n = |s| so that the initial
vertex Λ.s is represented by (0,0) and the
goal vertex t.Λ by (m,n)}

Let c = h(0,0) be lower bound of distance
from initial vertex to goal vertex;
Let Complete = False;
Let path detour pd = 0; D_S(s,t) =∞;
Push (0,0) on stack S_pd;
Repeat      {processes stacks in increasing
              order of detour}
   Repeat     {processes stack of
                detour pd}
        Pop (i,j) from S_pd;
        If Not Label(i,j) Then
           If i = m and j = n
              Then Complete = True;
                   D_S(s,t) = c + pd End
              Else
                 Expand (i,j)
     Until Complete or S_pd empty;
     If Not Complete Then
        pd = min{i | S_i nonempty}
```

```
Until Complete or All S_i empty
End.
```

Where Expand(i,j) is the procedure

```
Expand (i,j);
  Label(i,j) = True;
  For each edit rule e : x → y Do
    If Not Label (i+|y|,j+|x|)
      Then
        Let u = s_{j+1}..s_{j+|x|+1};
        If u is an instance of x
          Then If p_{x/u}(y) = t_{i+1}..t_{i+|y|+1}
            Then
              {compute detour d of
               neighbor and push on
               stack indexed by detour}
              d = c(e) + h(i+|y|,j+|x|)
                       - h(i,j);
              Push (i+|y|,j+|x|)
               on S_{pd+d}
              End
      End
End
```

The algorithm halts when it has found a least cost edit sequence with the cost of that edit sequence, or else has determined that the source string cannot be transformed into the target string (all stacks S_i empty and $D_S(s,t) = \infty$).

IV. Lower Bound Function for Edit Distance

In this paper we will only consider cost systems which conform to simple Levenshtein distance so that any lower bound function which is valid for simple Levenshtein distance will serve for edit distance. In [2], the author introduced a bound function for simple Levenshtein distance based on character frequency.

For any vertex in the transformation graph (i,j) a source suffix $s_{i+1}..s_m$ must be edited to produce a target suffix $t_{j+1}..t_n$. To establish a lower bound to the edit distance between suffixes, define character frequency functions $f(s,i,\mathbf{a})$ giving the frequency of character **a** in the source suffix beginning with character s_{i+1} and $f(t,j,\mathbf{a})$ similarly defined for the target suffix. Clearly these functions can be computed in linear time in a single right to left pass of the source and target strings. Now define the *excess* functions E (of source over target or vise versa) as:

$$E[s,i,j] = \sum_{\substack{a \in A \\ f(s,i,a) > f(t,j,a)}} [f(s,i,a) - f(t,j,a)]$$

and the frequency bound as:

$$h(i,j) = max\{E[s,i,j], E[t,i,j]\}.$$

It is shown in [2] that this is indeed a lower bound and that it is monotonic

The following example illustrates the use of the Minimum Detour algorithm with the frequency lower bound function. For the edit system shown, assume that all edit rules have cost 1 except equal substitution, $W \rightarrow W$, which has cost 0. The lower bounds are as shown. For the initial vertex, Λ.abaaa, the frequency functions for source suffix abaaa and target suffix aaaa are $f(s,0,\mathtt{a}) = 4$, $f(s,0,\mathtt{b}) = 1$, $f(t,0,\mathtt{a}) = 4$, $f(t,0,\mathtt{b}) = 0$ giving $E[s,0,0] = 1, E[t,0,0] = 0$ and $h(0,0) = 1$.

$$s = \mathtt{abaaa},\ t = \mathtt{aaaa}$$

$E = \{WW \rightarrow W,\ WW' \rightarrow W,\ W \rightarrow W,\ W \rightarrow W'\}$

$\Lambda.\mathtt{abaaa}^{h=1} - 0 \longrightarrow \mathtt{a.abcc}^{h=1} - 1 \longrightarrow \mathtt{aa.aaa}^{h=1}$

$\quad|$

$\quad 1$

$\quad\downarrow$

$\mathtt{a.aaa}^{h=0}$

Figure 2. Example of Lower Bound Function

In expanding the initial vertex, it is seen that both neighbors **a.aaa** and **a.baaa**, obtained by applying edit rules $WW' \rightarrow W$ and $W \rightarrow W$ respectively, are of detour 0. However, applying $W \rightarrow W'$ to **a.baaa** to obtain **aa.aaa** produces a detour of 1 so that Minimum Detour would find the distance of 1 by expanding **a.aaa** and successor items through equal substitution in linear time.

V. Applications of General Edit Systems

It is anticipated that general edit systems will have applications to the same sorts of problems to which Levenshtein distance has been applied, and perhaps more accurately model primitive operations comprising string transformation. These applications are numerous; they include pattern recognition, spelling correction, data compression and error correction in variable length codes. A survey of applications is contained in [7].

Here we sketch an approach to using edit systems for data compression. In [4] the author gives an algorithm for constructing the longest common subsequence and the shortest common supersequence of a set of strings. For a set of very similar strings, such as the heartbeats in a resting ECG strip of a healthy patient, an effective compression technique is to store the transformations of either the LCS or SCS into each string in the set. Now the longest common subsequence is transformable into each string using only the operations

of insertion and equal substitution, the shortest common supersequence requiring only the operations of deletion and equal substitution. This technique can be extended to a larger set of operations.

Define the *inverse of an edit system* as being another edit system which provides the inverse to any transformation provided by the first system in the following sense: there is a 1-1 onto function $f : E \rightarrow E'$ such that if T is a transformation over E with $T : s \rightarrow t$, then $f(T)$ is the corresponding transformation over E' for which $f(T) : t \rightarrow s$. As an example, the inverse, for a binary alphabet, of the system in Figure 2 is $E = \{W \rightarrow WW,\ W \rightarrow WW',\ W \rightarrow W,\ W \rightarrow W'\}$ with $f(WW \rightarrow W) = W \rightarrow WW$, $f(WW' \rightarrow W) = W \rightarrow WW'$, etc. Now if Y is a set of strings all transformable into a string X under some edit system with an inverse edit system, then Y may be represented by string X along with the inverse transformations.

VI. Summary

In this paper, the notion of edit distance, introduced in [6] and [8] has been extended and the Minimum Detour algorithm first introduced in [1] and applied to Levenshtein distance computation [2] and string classification in [3] has been adapted to finding edit distance when edit costs conform to simple Levenshtein distance. For special cases, the algorithm will produce the distance in $O(dn)$ time where d is the distance. This was shown to be the case in [2] for simple Levenshtein distance (which is an edit system with edit costs conforming to simple Levenshtein distance). In any case, by using the Label array, no more than mn vertices are expanded so distance computation is $O(mn)$ in general.

REFERENCES

1. HADLOCK, F. O. (1977), A shortest path algorithm for grid graphs, *Networks* 7:323-334.
2. HADLOCK, F. O. (1988), Minimum Detour Methods for String or Sequence Comparison, *Congressus Numerantium.*
3. HADLOCK, F. O., Minimum Detour Methods for Levenshtein Based String Classification, (submitted).
4. HADLOCK, F. O., Algorithms for Longest Common Subsequence and Shortest Common Supersequence, (submitted).
5. HART, P. E., NELSON, N. G., and RAPHAEL, B., (1978), A formal basis for the heuristic determination of minimum cost paths. *IEEE Trans. Syst. Sci. Cybern,* Vol. SSC-4, 100-107.
6. LEVENSHTEIN, V. I. (1966), Binary codes capable of correcting deletions, insertions, and reversals. *Cybernetics and Control Theory 10,* 707-710.
7. SANKOFF, D. AND KRUSKAL, G. B. (Eds.), (1983), Time warps, string edits, and macromolecules: The theory and practice of sequence comparison. Addison Wesley, Inc.

8. WAGNER, R. A. AND FISCHER, M. G., (1974), The string to string correction problem. *JACM 21*, 168-173.

1989 RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Development of a New Ultrasensitive Cholesterol Assay System for the Determination of Free Cholesterol in Biological Fluids

Prepared by:	Eric R. Johnson, Ph.D.
Academic Rank:	Professor
Department and University:	Department of Chemistry Ball State University
Research Location during 1988 FSRP:	USAF SAM/NGIL Brooks AFB, TX 78235
USAF Researcher:	Harvey A. Schwertner
Date:	15 January 1990
Contract No:	F49620-87-R-0004

ABSTRACT

Development of a New Ultrasensitive Cholesterol Assay System for the Determination of Free Cholesterol in Biological Fluids

A sensitive method suitable for the analysis of subnanogram amounts of cholesterol in urine has been developed. Free cholesterol in urine is isolated by ultrafiltration. The free cholesterol in the ultrafiltrate and the added internal standard, epicoprostanol (5-beta- cholestan-3-alpha-ol) are then derivatized with 2,3,4,5,6-pentafluorobenzoyl chloride to form their corresponding pentafluorobenzoyl (PFB-) esters. PFB-cholesterol is then separated from PFB-epicoprostanol and other by-products by electron capture gas chromatography. The method has a lower limit of sensitivity of approximately 100 picograms of cholesterol injected, which corresponds to a concentration of 50 nanograms free cholesterol per mL urine.

The search for esterases and sulfatases specific for the corresponding cholesterol derivatives was initiated. Several commercially available esterases and sulfatases were evaluated for their ability to catalyze the hydrolysis of cholesteryl palmitate and cholesterol sulfatase.

Acknowledgements

I am grateful to the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research project. We also thank Universal Energy Systems, Inc., for their administrative support of the 1989 Research Initiation Program.

In the accomplishment of research objectives during this productive and rewarding project, I am indebted to a number of people. Dr. Harvey A. Schwertner, USAFSAM/NGIL, provided enthusiastic support and encouragement, expert technical advice, and excellent ideas to this project. A number of people from Ball State University contributed to the success of this project. My student research assistants, Ted Ashburn and Randi Rieman, performed admirably in their assigned research duties. Dr. Patricia Lang and Mr. Ron Mendenhall provided infrared spectral data and Mr. David Bir obtained mass spectral data for the characterization of cholesterol derivatives synthesized in this assay. Dr. Scott Pattison, Chair of the Ball State University Department of Chemistry, and Dr. James Pyle and the staff in the Office of Research and Sponsored Programs provided excellent administrative and cost sharing support.

1. INTRODUCTION

Coronary artery disease (CAD) is a leading cause of non-accidental death in the United States. In 1985, over 500,000 deaths were attributable to CAD (1). In addition to the immeasurable pain and suffering of the individual victims and their families, CAD is very costly to society with an estimated annual cost of over $60 billion (1,2).

A portion of this financial burden falls upon the Air Force as CAD is also the leading cause of non-accidental deaths among active duty Air Force personnel. A 1981 study (3) estimated that the cost to the Air Force in lost training, hospitalization, and disability retirement of a rated pilot affected by CAD could exceed $500,000, which translates to an annual cost of over $40 million.

Although still a serious problem, the incidence of CAD is declining, due at least in part to the identification and popularization of risk factors associated with the development of CAD (4). While these results are encouraging, further decline in CAD incidence depends on the earlier detection of CAD in people at risk. At present, the earliest CAD symptoms appear too late for many individuals, leaving coronary bypass surgery or balloon angioplasty as the only alternatives to prolong and improve the quality of life. Sensitive and specific methods must be developed to identify those persons at risk at an earlier stage, when dietary and lifestyle modifications can be used to better advantage.

Several potential markers could prove to be useful in predicting risk levels in CAD. One such marker is the fraction of cholesterol in biological fluids that is not bound to protein, which is commonly

referred to as free cholesterol. The small but important pool of free cholesterol that likely exists in biological fluids could mirror physiological events occurring in the cardiovascular system that reflect the extent and severity of CAD. A method to detect the admittedly low concentration of free cholesterol in biological fluids would be valuable for the evaluation of this marker as a predictor of CAD risk.

In addition to free cholesterol, simple conjugates of cholesterol such as cholesterol esters, sulfates, and glucuronides are present in the various fluid pools of the human body. Potential correlations of the levels of these substances in serum, saliva, or urine to CAD risk have not been investigated due to a lack of reliable methods to specifically detect such small quantities.

During the 1988 FSRP, a cholesterol assay was developed that can be used for the measurement of subnanogram amounts of the sterol. This assay was further optimized and a manuscript describing this assay was prepared with Dr. Harvey Schwertner during the early part of the RIP support period. This manuscript has been accepted for publication in Clinical Chemistry. This research describes the development of a method that incorporates this ultrasensitive assay in the measurement of the levels of free cholesterol in urine. The development of these methods is a necessary prerequisite to the evaluation of the potential of free cholesterol and cholesterol conjugates in predicting CAD risk.

2. OBJECTIVES

As stated above, the ultimate goal of this research, in collaboration with Dr. Harvey Schwertner (USAFSAM/NGIL), is to evaluate the effectiveness of free cholesterol and cholesterol conjugates in

serum, saliva, and urine as predictors of CAD risk. The first step in this process, the development of an ultrasensitive cholesterol assay, was accomplished during the 1988 FSRP and the beginning of the 1989 RIP support period. The specific objectives are outlined and follows:

I. Develop an ultrafiltration system for the isolation of free (non-protein-bound) cholesterol and cholesterol conjugates from biological fluids.

II. Enhance the sensitivity and minimize interferences from reaction by-products in the ultrasensitive cholesterol assay system by optimizing assay conditions.

III. Using commercially available esterases, sulfatases, and glucuronidases, develop a sample treatment protocol that can be used with the ultrasensitive cholesterol assay to determine the amounts of cholesterol and cholesterol conjugates (esters, sulfates, and glucuronides) present in the protein-free fluid fractions obtained from objective 1.

IV. Modify this sample treatment protocol to permit the assay of cholesterol and cholesterol conjugates in biological fluids.

V. Determine the reproducibility for the sample treatment and assay processes.

3. METHODS AND RESULTS

I. Separation of free cholesterol.

Approach - An Amicon MPS-1 ultrafiltration device with 30,000 molecular weight cutoff membrane (Amicon YMT membrane) was evaluated for use in the separation of non-protein-bound cholesterol and hydrolyzable cholesterol esters. This system was chosen as a result of its routine

use in the analysis of non-protein-bound testosterone (6), a compound which is similar to cholesterol in structure. These membranes are designed to pass free testosterone with minimal adsorption. To determine whether or not free cholesterol could pass through the YMT membrane, samples of ultrafiltered urine were spiked with various detectable concentrations of free cholesterol and subsequently passed through a fresh YMT ultrafiltration membrane. Samples were assayed for cholesterol using the assay method that was initially developed in the 1988 FSRP and enhanced in this project.

Results - Initial experiments with the YMT membrane indicated that some, but not all, of the cholesterol in whole urine passed through the membrane. As shown in Figure 1, about 23% of the total cholesterol and hydrolyzable conjugates in a whole urine sample passed through the ultrafiltration membrane, indicating that the remaining 77% was likely bound to macromolecular complexes in excess of 30,000 molecular weight.

In order to assess the adsorption of free cholesterol to the YMT membranes, a known amount of cholesterol was added to a sample of ultrafiltered urine. A 500 microliter aliquot of this sample was assayed for cholesterol and hydrolyzable conjugates using the enhanced ultrasensitive assay described below. The remainder of this sample was passed through a fresh YMT membrane until the volumes of the filtrate and retentate were approximately equal. Aliquots (500 microliters) of the filtrate and retentate were also assayed for cholesterol and hydrolyzable conjugates. Figure 2 illustrates that the cholesterol concentrations found in these three samples were nearly the same. Thus the YMT membrane does not appear to adsorb free cholesterol from urine.

Figure 1
Ultrafiltration of Whole Urine

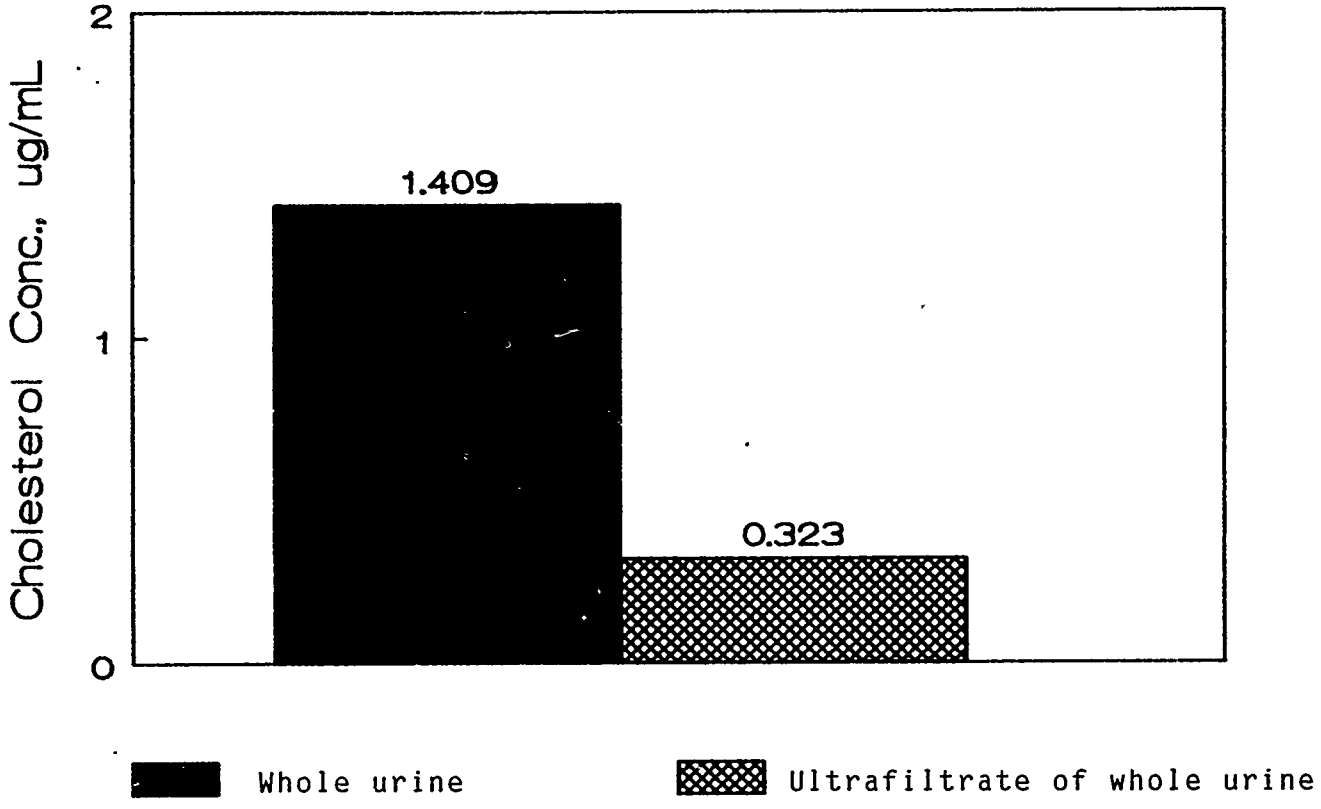
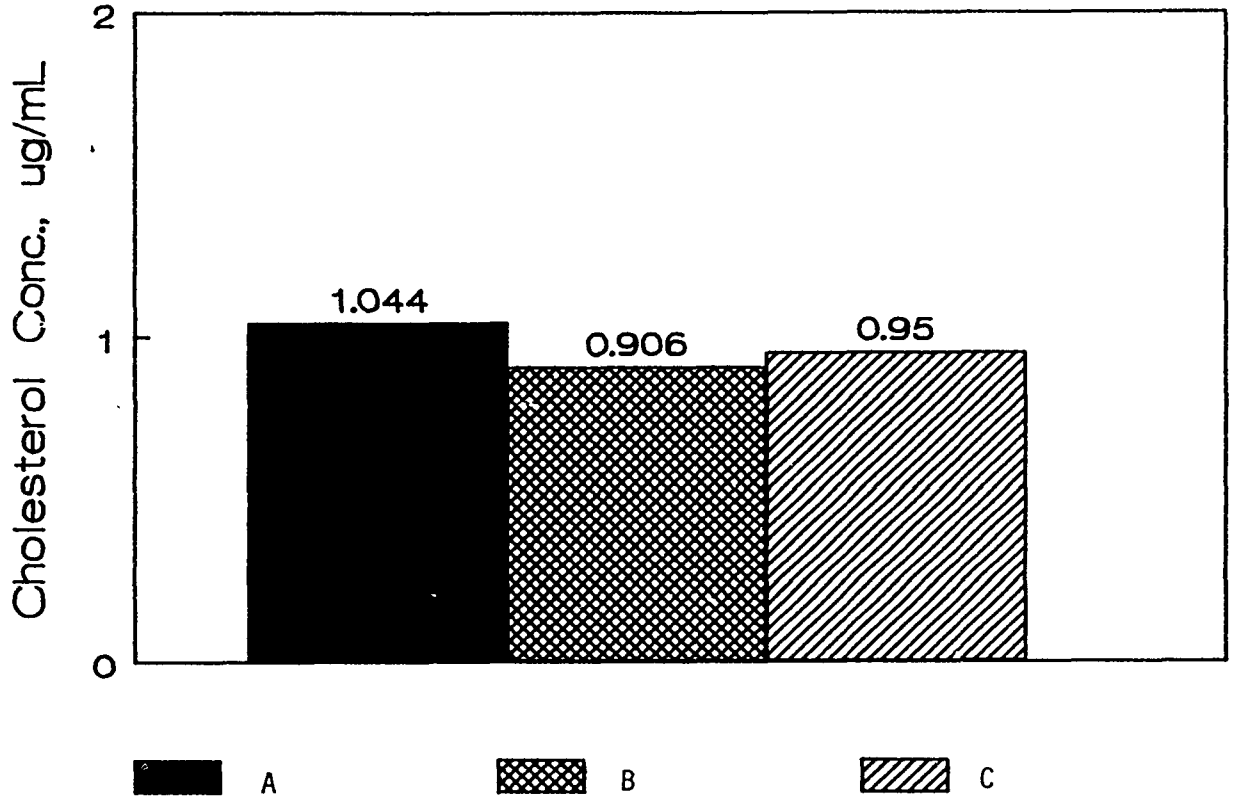


Figure 2
Cholesterol Passage through Membrane



A: Ultrafiltered urine, to which 1 ug/mL cholesterol was added.

B: Retentate of Sample A after ultrafiltration through fresh membrane.

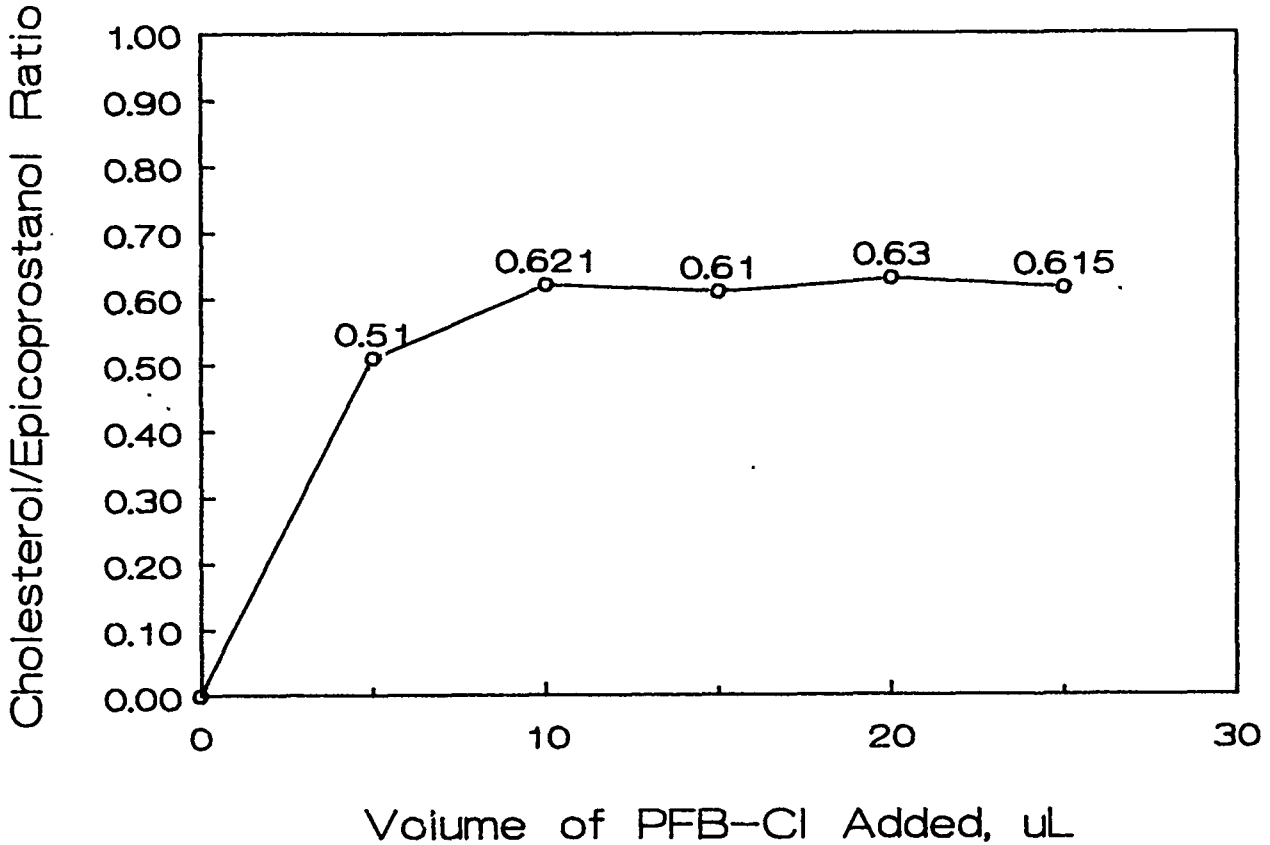
C: Filtrate of Sample A after ultrafiltration through fresh membrane.

II. Optimization of the cholesterol assay system.

Approach - Since the values for free cholesterol in biological fluids were likely to be toward the lower limits of sensitivity of the assay originally designed during the 1988 FSRP, a study to increase the sensitivity of the assay was undertaken. In order to minimize chromatographic peaks arising from by-products of the reaction of 2,3,4,5,6-pentafluorobenzoyl chloride (PFB-Cl) with water and other hydroxylic components of the system, lower amounts of PFB-Cl were added. The volume of the final solvent (ethylacetate) was adjusted downward in relation to the amount of urine to give a concentrating effect on the resulting cholesterol concentration.

Results - As indicated in Figure 3, as low as 10 microliters of PFB-Cl gave essentially the same yield of PFB-cholesterol as the original 25 microliter volume. At the same time, the amount of signal detected by the electron capture system in the early part of the chromatograms was observed to decrease substantially. This tended to minimize interference of these by-products with the elution of PFB-ester of epicoprostanol, the internal standard. Since the original assay called for 1.0 mL of biological fluid to be extracted, derivatized, and dissolved in 1.0 mL of ethylacetate for injection into the gas chromatograph. With the cleaner chromatogram resulting from the lower amount of PFB-Cl in the derivatization step, only 0.5 mL ethylacetate was used. Thus the sensitivity was increased by a factor of two.

Figure 3. Effect of Amount of PFB-Cl on Yield of PFB-Cholesterol



The most sensitive version of this assay is performed as follows:

Preparation of Urine Samples for Cholesterol Analysis by EC-GC

1. Centrifuge freshly collected urine or saliva at 3500 x g for 20 min. Place 1.0 mL of the urine or saliva supernatant and 1.0 mL distilled water in a 16 mm x 125 mm screw-capped culture tube.

2. Add 100 microliters of the internal standard solution (epicoprostanol, 0.03 mg/mL in 100% ethanol).

3. Add 7 mL chloroform-methanol (2:1, v/v), cap, and shake for 10 min on an Eberbach shaker at 4 linear cycles per sec. Centrifuge the sample for 5 min at 1500 x g to separate the phases.

4. Transfer approximately 4-5 mL of the chloroform layer (lower phase) to a 16 mm x 125 mm screw-capped culture tube. Evaporate the chloroform with a gentle stream of nitrogen at 60 ^{o}C in a heating block.

5. For the determination of total cholesterol, add 1.0 mL 0.5 M KOH in methanol. Cap the tubes and heat for 20 min at 95 ^{o}C in a heating block. Vortex the sample for 5 sec at 0, 10, and 20 minutes during the heating period. For the determination of non-esterified cholesterol, add 2.0 mL methanol and proceed to step 6 without heating.

6. Cool the samples to room temperature and add 1.0 mL distilled water and 4.0 mL hexane. Shake for 10 min on an Eberbach shaker at 4 cycles per sec. Centrifuge at 600 x g for 15 seconds to separate the phases. Transfer approximately 4 mL of the hexane layer (top phase) to a clean 13 mm x 100 mm screw-capped culture tube. Evaporate the hexane with a gentle stream of nitrogen at 60 ^{o}C in a heating block.

7. Add 10 microliters of PFB-chloride to the dried extract, cap the tube, and heat for 60 min at 95 ^{o}C in a heating block. Evaporate the

excess PFB-chloride in the vacuum centrifuge with heating on for 15 minutes. Dissolve the dried sample in 0.5 mL ethylacetate for injection into the gas chromatograph using the chromatographic conditions described in Figure 4.

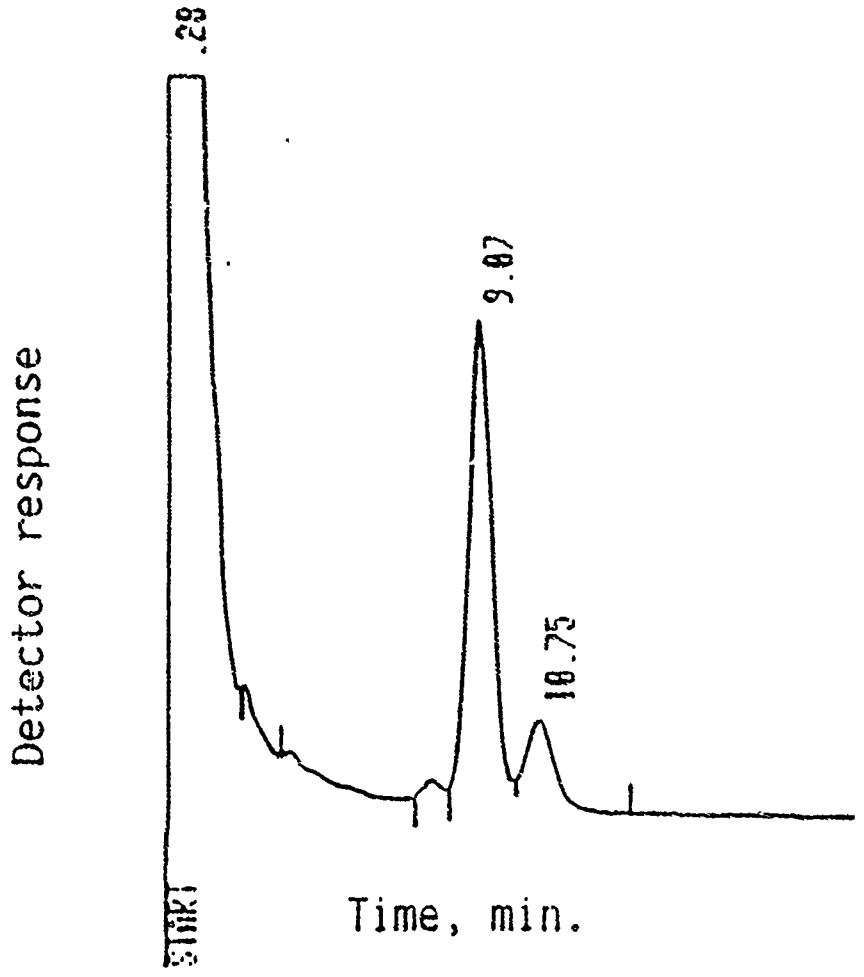
In addition to the above changes in the assay, the identities of the PFB-esters of cholesterol and epicoprostanol standards were confirmed using infrared spectroscopy and mass spectrometry. Chemical ionization mass spectrometric analysis of the PFB-derivative of cholesterol (mol. wt. 580.7 daltons) showed a parent ion peak at a m/e ratio of 579.9. In comparison the infrared spectrum of a cholesterol standard, the infrared spectrum of PFB-cholesterol showed the appearance of a characteristic aromatic ester absorption band at 1233 cm^{-1} and the complete disappearance of the broad oxygen-hydrogen stretching band at 3400 cm^{-1}. These data support the observation that cholesterol was converted to its PFB-ester in high yield by reaction with 2,3,4,5,6-pentafluorobenzoyl chloride. Similar results were obtained for the comparison of the infrared spectra of the internal standard, epicoprostanol, and its PFB-ester.

The derivative characterization and many of the improvements described above have been incorporated in the manuscript that has been accepted for publication in Clinical Chemistry (7).

III. Evaluation of enzymes for the hydrolysis of cholesterol conjugates.

Approach - In addition to free cholesterol, simple cholesterol conjugates such as cholesterol esters and sulfates may be present in the ultrafiltered samples of biological fluids. The levels of these compounds can be measured by adding specific esterases (8), sulfatases,

Figure 4. Electron Capture Chromatographic Analysis of Cholesterol in Urine



Instrument: Hewlett-Packard 5890A Gas Chromatograph equipped with a ^{63}Ni electron capture detector

Column: 0.53 mm x 5 m HP-1
Injector Temperature: 280 ^{o}C
Oven Temperature: 270 ^{o}C
Detector Temperature: 280 ^{o}C
Carrier Gas Flow: 5% methane in argon, 6.0 mL/min
Sample Injected: 1 uL

and glucuronidases (9) to the ultrafiltrates to release cholesterol from its conjugated form for derivatization to PFB-cholesterol and subsequent quantitation by electron capture gas chromatography as outlined above. For each enzyme treatment, the result is the sum of the free cholesterol initially present and the cholesterol released by specific enzyme action. The difference between this value and the value obtained for the free cholesterol present in the sample prior to enzyme treatment is the amount of cholesterol existing as the particular conjugate in the ultrafiltered sample. For each specific enzyme, the extent of hydrolysis and recovery of cholesterol from the substrate conjugate can be assessed by addition of a known amount of the conjugate or unconjugated cholesterol to different aliquots of an ultrafiltered fluid sample.

Results - An esterase (esterase from porcine liver, Catalog Number E-3128 from Sigma Chemical Company, St. Louis, MO) and a sulfatase (Type VI from Aerobacter aerogenes, Catalog Number S-1629 from Sigma Chemical Company, St. Louis, MO) were evaluated for use in this procedure by assessing their extent of hydrolysis on cholesteryl palmitate or cholesterol sulfate, respectively. In the case of the esterase, only 26% of the total cholesteryl palmitate was hydrolyzed after 30 minutes at $37_{o}C$ in 0.2 M Tris-HCl buffer, pH 7.4. The sulfatase, acting on cholesterol sulfate, hydrolyzed only 18% of the substrate under the same conditions. More optimal conditions or more efficient enzymes must be found in order to quantitatively determine the amounts of these cholesterol conjugates in biological fluids.

IV. Modification of sample treatment protocol to measure cholesterol conjugates

Approach - The procedure for the analysis of unconjugated cholesterol in whole saliva and urine was developed during the 1988 FSRP and further modified during this project (7). The method involves the extraction of saliva or urine with chloroform-methanol followed by derivatization with PFB-chloride for subsequent gas chromatographic analysis as indicated above. Based on previous work (7), assay interferences from other compounds in these fluids are largely eliminated by the extraction process. Both unconjugated cholesterol and base-hydrolyzable cholesterol conjugates are easily measured with this process. The large molecular weight components present in whole saliva or urine apparently did not interfere with the cholesterol assay of these samples.

The determination of cholesterol conjugates can be accomplished by specific enzyme-catalyzed hydrolysis prior to the extraction step. Small aliquots of a saliva or urine sample are treated with a specific esterase or sulfatase enzyme followed by chloroform-methanol extraction and the remainder of the assay procedure as indicated above. The extent of hydrolysis and recovery of the individual cholesterol conjugates will be determined by re-assaying samples to which a known amount of a given cholesterol conjugate had been added.

Results - Initial experiments with the esterase and the sulfatase chosen for evaluation did not provide the high degree of substrate hydrolysis necessary for incorporation into this assay.

Additional experiments are required to increase the amount of substrate hydrolysis or find more efficient enzymes to use in this assay.

V. Evaluation of reproducibility of the assay.

Approach - Assay procedures were evaluated for reproducibility by determining the coefficient of variation on fifteen aliquots of the same fluid sample.

Results - The coefficient of variation was determined for the new version of the cholesterol assay described in objective II. Fifteen aliquots of a urine sample were extracted, hydrolyzed, and derivatized as indicated in the assay procedure. The results are shown in Table 1. The within-run coefficient of variation for these samples was determined to be 3.00% at a cholesterol concentration of 3.06 ug/mL urine.

Table 1
Determination of the within-run coefficient of variation
for the revised cholesterol assay

Trial	Ratio (epicoprostanol/cholesterol)
1	0.7101
2	0.6970
3	0.7769
4	0.7126
5	0.7123
6	0.7349
7	0.7119
8	0.6927
9	0.7076
10	0.7359
11	0.7262
12	0.7176
13	0.7417
14	0.7443
15	0.7369

Mean ratio = $0.724 \pm .022$

Coefficient of variation = 3.00 %

4. RECOMMENDATIONS

I. The Amicon MPS-1 ultrafiltration system should be evaluated for use in the separation of free cholesterol and cholesterol conjugates in serum and saliva. The small but important pool of free cholesterol that likely exists in serum could mirror physiological events occurring in the cardiovascular system that reflect the extent and severity of CAD. A method to detect the admittedly low concentration of free cholesterol in serum would be valuable for the evaluation of this marker as a predictor of CAD risk.

Since salivary fluid has been suggested to be a membrane ultrafiltrate of serum (5), cholesterol levels in salivary fluid may correlate with the free cholesterol levels in serum. Measurement of salivary cholesterol levels is attractive in that samples can be taken from patients by non-invasive means. With some ingenuity in the design of a sampling device, it may be possible to use saliva samples to follow cholesterol levels during periods of G-force-induced stress in centrifuge studies and, ultimately, during flight. Salivary cholesterol levels could prove useful in predicting CAD risk. Given the success of the MPS-1 system in the separation and quantitation of free cholesterol in urine, application of this technique to saliva and serum would likely also be successful.

II. The fraction of cholesterol that is not bound to serum proteins likely exists in forms other than the free sterol. Since a more rapid exchange rate between lipoproteins and membrane systems has been observed for cholesterol than other lipids, a water-soluble form has been proposed to rationalize this transfer (10). The water-soluble form could be a

glucuronide, a sulfate, or could simply be bound to a relatively small protein such as albumin (11). The development of enzyme-based systems utilizing the new cholesterol assay described in this report to quantitate these forms in ultrafiltered serum will facilitate future studies to evaluate these cholesterol conjugates as potential CAD markers. Other commercially available sulfatases, glucuronidases, and esterases should be evaluated for their ability to cleave these cholesterol conjugates more efficiently and completely. Conditions should be found that will give an essentially complete cleavage of the specific conjugates in a reasonable hydrolysis time. If this is successful, then the cholesterol assay protocol can be modified to take advantage of the hydrolytic specificity of these enzymes to quantitate the amount of individual cholesterol conjugates present in the samples.

REFERENCES

1. NIH Consensus Statement, "Lowering Blood Cholesterol to Prevent Heart Disease", Arteriosclerosis, 1985, Vol. 5, pp. 404-412.
2. Levy, R.I., and Moskowitz, J., "Cardiovascular Research: Decades of Progress, a Decade of Promise", Science, 1982, Vol. 217, pp. 121-129.
3. Petersen, C.C., "Cost of Cardiovascular Disease and Planning for Future Directions of the U.S. Air Force HEART Program", Prev. Med., 1981, Vol. 10, pp. 277-286.
4. American College of Cardiology, "Eleventh Bethesda Conference: Prevention of Coronary Heart Disease", Am. J. Card., 1981, Vol. 47, pp. 713-776.

5. Lochner, A., Welsner, S., and Zlatkis, A., "Gas Chromatographic-Mass Spectrometric Analysis of Volatile Constituents in Saliva", J. Chromatography, 1986, Vol. 378, pp. 267-282.

6. Green, P.J., and Yucis, M.J., "Free Testosterone Determination by Ultrafiltration and Comparison with Dialysis", Clin. Chem., 1982, Vol. 28, pp. 1237-1238.

7. Schwertner, H.A., Johnson, E.R., and Lane, T.E., "Electron Capture Gas Chromatography as a Sensitive Method for Measuring Subnanogram Amounts of Cholesterol", Clin. Chem., 1990, in press.

8. Wiebe, D.A., and Bernert, J.T., Jr., "Influence of Incomplete Cholesteryl Ester Hydrolysis on Enzymic Measurements of Cholesterol", Clin. Chem., 1984, Vol. 30, p. 252.

9. Setchell, K.D.R., Alme, B., Axelson, M., and Sjovall, J., "The Multicomponent Analysis of Conjugates of Neutral Steroids in Urine by Lipophilic Ion Exchange Chromatography and Computerised Gas Chromatography-Mass Spectrometry", J. Steroid Biochemistry, 1976, Vol. 7, pp. 615-629.

10. Goldstein, J.L., Hegelson, J.A.S., and Brown, M.S., "Inhibition of Cholesterol Synthesis with Compactin Renders Growth of Cultured Cells Dependent on the Low Density Lipoprotein Receptor", J. Biol. Chem., 1979, Vol. 254, pp. 5403.

11. Deliconstantinos, G., Tsopanakis, C., Karayiannakos, P., and Skalkeas, G., "Evidence for the Existence of Non-esterified Cholesterol Carried by Albumin in Rat Serum", Atherosclerosis, 1986, Vol. 61, pp. 67-75.

1989 REPORT ON 1988 RIP

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by
Universal Energy Systems, Inc.

Prepared by:	Harold G. Longbotham
Academic Rank:	Assistant Professor of Electrical Engineering
Department and University:	Electrical Engineering, University of Texas at San Antonio
Research Location:	USAFSAM/RZV Brooks AFB San Antonio, Texas 78235
USAF Researcher:	Captain Norman Barsalou

Date: 21 December, 1989

Application of Nonlinear Filters to VEP Data

by

Harold Longbotham

Abstract

This research continued the research started in the summer of 1988 for the Radiation Lab at Brooks Air Force Base. It was determined that the analysis of VEP data could be improved through the use of symmetric OS (order statistic) filters, for the data sets examined. This research extended the previous research by providing a software package and user's guide that can be used for elimination of additive noise in analysis of VEP data. It also provides the software needed for determining the need of an adaptive algorithm, and the factors on which the adaptation will depend. Some of the factors that may be considered are timing and dosage of anesthesia, amplitude, phase, and muscle artifacts. We assume periodic signals (such as EKG and respiratory) have been eliminated.

As a result of this work the above software was delivered to BAFB/RZV (see the appendix), adjusted to run on their system, and two papers were written (as outlined in the following sections). One of these papers was published in the proceedings of the International Circuits and Systems Conference in Nanjing China. In it we discussed implimentations of positive boolean filters and their usage in the processing of VEP data. The second paper has been selected to be delivered at the International Symposium on Circuits and Systems at New Orleans in May of 1990. In it we outline the proceedures used in the research covered under this grant and demonstrate how the are applicable to periodic data.

APPLICATION OF NONLINEAR FILTERS TO VEP DATA

HAROLD LONGBOTHAM

INTRODUCTION

Historical Perspective

D. Regan [1] has defined a steady state evoked potential (SSEP) to be "scalp potentials that are locked in time (or in phase) to the sensory stimulus which evokes them". We also know the system is nonlinear due to the following observations by J.E. Desmedt [2], "The fact that the VEP (visual evoked potential) to sinusoidally modulated light contains higher harmonics indicates directly that the human VEP system is nonlinear". The above statements lead us to consider steady state VEP data collection as consisting of inputting a periodic signal into a noisy nonlinear system whose parameters are unknown, but with a response one could assume in the noiseless case to be of the same period as the stimulus. This yields the classical approach (outlined below) to raw data noise elimination that has been used to date, and the modifications we consider in this paper.

The signal to noise ratio in VEP data varies from fair (in anesthetized monkeys with implanted probes) to poor (in alert humans with scalp electrodes) necessitating filtering of the data to improve the signal to noise ratio (a graph of raw SSVEP data from an anesthetized monkey in Appendix 1). The previous definition of SSEP allows the current processing of raw data to be done by two methods which we will now describe, for the improvement of these two methods is the focus of our application.

Let us assume the stimulus and response are sampled simultaneously, at N sample points in each period of the stimulus, yielding a sequence of points $\{x_i\}$, $1 \leq i \leq L$, where the number of complete periods of sample data is the lagest integer less than L/N. We will assume the jitter is constant, i.e. the stimulus is sampled at exactly the same points

in each period. The first method is to reduce all information to the time frame encompassed by one period of the stimulus indicated by the N different averages of the data sets $\{x_{J + jN}\}$, where J is a fixed integer such that $1 \leq J \leq N$ and j is a nonnegative integer such that $J + jN \leq L$. We note the result is a set of N means of responses, where each mean is assumed to be the mean of a constant value perturbed by noise. This technique is justified by noting the mean is the MLE (maximum likelihood estimate) of an i.i.d. (independent and identically distributed) sample from a Normal distribution[3]. The second method is to use an FIR filter of length M that windows each of M consecutive values in the data set $\{x_i\}$ and replaces the central value with the average M values in the window. This method is justified by noting that the FIR averager filter is the BLUE among all linear filters for a constant signal that is corrupted by additive Gaussian noise[4]. We note the first method makes the assumption of the noise being independent from point to point and identically (normally) distributed at each point and the second method is only optimal over linear combinations of the sample values. Since we are assuming the system is nonlinear, it seems reasonable to believe a new nonlinear approach may improve on the linear (averager) technique presently used in each of the above methods.

The Median Filter

The median of a data set with an odd number of values is the data value in the set for which at least half of the values in the set are greater than or equal to it and at least half of the values in the data set are less than or equal to it. The median filter of length 2N+1 is defined as the filter which considers each consecutive set of 2N+1 values and replaces the central value with the median of the 2N+1 values. For an even number of values there are other definitions, however from now on we will restrict ourselves to data sets/windows with an odd number of values since the theoretical results on which we will base our discussion are only derived for odd length filters [6, 8].

Properties of the Median Filter

Tukey [5] was the first to suggest the use of the median in the place of the averager because of its ability to totally eliminate impulses and maintain monotone regions. The reason for the recent surge of interest in the median filter lies not only these facts, but also from recent theoretical work that has accurately described the pass band of the median. Gallagher and Wise [6] have shown that a median filter of length 2N+l will pass a finite length signal unaltered (such a signal is called a root) if and only if each set of N+2 consecutive values is monotone. This assumes the signal is initially padded with N values equal to the first value. They have also shown that any signal is reduced to a root by a finite number of passes (depending on the signal length) of the median filter. Lately the median filter has proven successful in a variety of applications involving impulsive noise (including estimation of sonar signals [l0] and real-time TV signal processing with impulsive noise [ll]).

The median filter is applicable in our situation since a constant signal is definitely monotone and filter length may be adjusted to compensate for the varying lengths of artifacts. As an example assume the stimulus is at l0 cycles per second, and the response is of the same period. Assume our subjects are alert humans with muscle artifacts on the order of one second. This would produce a burst of impulsive noise our "assumed" constant samples of duration l0. Therefore we would require a median filter of length 2l to completely remove the artifact.

The OS (Order Statistic) Filter

An OS filter [7] is very similar to an Flr filter except that when a set of consecutive values is windowed from the data sequence, the values in the window are ordered before weighting. The median is a simple example of an OS filter where there is an odd number of weights and the weights are all zero except for the central weight, which is one. It is also noted that the averager is both an OS filter and an FIR filter, since if all data values are equally weighted ordering has no effect on the outcome of the filtering.

Applications of OS Filters

Optimal OS filters (using the MSE, mean square error, criterion) have been found for the estimation of a constant signal with additive i.i.d. noise, where the noise is from a variety of distributions [7]. It has been shown that in order for the OS filter to provide an unbiased estimate (of a constant signal with additive noise) the coefficients must sum to one, and if the noise distribution is symmetric, so wil be the weights. One of the more popular of the OS filters is the a-trimmed inner mean, where the coefficients are symmetric, zeros on the outside and M values of l/M in the center. This allows one to use editing of outliers (due to outer zeros) and smoothing (due to the averaging of the inner ordered values) simultaneously.

Measure to be Used in the Comparison of Various Filters

The basic assumption in the analysis of SSVEP data is that the response is periodic, in the same period as the stimulus. This assumption, in the absence of noise and jitter will result in constant sequences. A common measure of how much a sequence deviates from an assumed signal is the MSE or mean square error. Therefore if the assumed signal is constant and estimated by the average of the sequence values, we need only calculate the sum of squares of the deviations of the sequence values from the average. We note this is just the variance of the signal. If the raw data has N points in each period, there will be N of these variances.

DELIVERABLES RESULTING FROM THIS RESEARCH

Software Package and User's Manual

A software package has been developed that will enable the investigator to use any OS filter (i.e. one may specify the filter length and coefficient set) to filter SSVEP data. This package enables the PI to look at the variability of the noise from level to level, consider nonsymetric filters, and examine the need for adaptive filters (as a function of phase or as a function of the dosage of anesthesia). A laser print out is then generated by the graphics package that will graph the raw data, the reconstructed data (the data with assumed constant sequences filtered with the entered OS filter and then reconstructed), and a graph of the filtered data (the reconstructed data passed through an user selected OS filter). Care must be taken that the coefficients of the OS filter sum to one, so as not to bias the measure of repeatability. One must also be sure the data segment to be evaluated is padded with enough data points at the front end for there to be an output for the first and last values in the segment. This is described in more detail in the user's manual.

Papers

The following is the abstract of a paper published in the proceedings of the International Conference of Circuits and Systems in Nanjing, China, 1989. The full body of the paper is included in the appendix.

"A Generalized Stack Filter"

Harold Longbotham and Parimal Patel
Division of Engineering
University of Texas at San Antonio
San Antonio, Texas 78285

Current research [l, 2] into the area of stack filters has demonstrated that when identical positive boolean functions are "stacked", the input signals obey the threshold decomposition and output signals obey the stacking principle. It has been shown that when the boolean function at each level, B, is a rank order filter, the stack filter, S_B, is in turn the corresponding rank order filter. Root structures for these specific filters have also been discussed. The importance of this past research cannot be underestimated, in that it leads to a specific implementation of rank order filters in real time [3, 4], whereas before the ordering operation involved in order statistic filtering incurred an extra cost in the time delay.

This paper seeks to extend these results in several directions. First we will show how an explicit formulation for the stack filter S_B can be written for any positive boolean function B. A root for a rank order filter is a signal that passes through the filter unchanged, i.e. a signal in the passband of the filter. The root structure for rank order filters when the class of signals is restricted to finite length signals has previously been discussed. We extend this to the class of two-sided infinite length signals and stack filters other than rank order.

Next we point out it is unnecessary to restrict oneself to the commonly defined stack filter definition (in which the same boolean function is used at each level), in order to develop digital filters that obey the threshold

decomposition/stacking principles. We then provide necessary and sufficient conditions on the variation of the boolean filters from level to level in order for stacking and threshold decomposition to be upheld. The resulting stack filters are then discussed along with their root structures.

We now address the question as to the importance of the stacking property in the scheme of boolean implemented digital filters. We all realize the result of output stacking on the boolean levels allows us to use a simple tree search instead of summing to determine the output. However, as we will have demonstrated, this severely restricts one as to the functions that can be implemented. Therefore we now attack the inverse problem of determining boolean filters on each level, to implement a given stack filter S_B, where the stacking property for outputs may or not be obeyed. The effect of this progress on the implementation of stack filter and general stack filters will be discussed.

The following is the abstract of a paper accepted to be delivered at abd published in the proceedings of the International Conference of Circuits and Systems, New Orleans, 1990.

Analysis of Periodic Signals via Order Statistic Filters

Harold Longbotham
Jerry Keating
The Nonlinear Signal Processing Group
The University of Texas at San Antonio

Randy Glickman
Opthomology Department
The University of Texas Health Science Center

Periodic signals are filtered according to frequency content when possible. However, it often occurs the frequency content of a response and/or noise is unknown. In such situations one may sometimes assume the output to have the same common period as the stimulus. In general, such a situation may arise anytime one is operating on a periodic signal with a linear time variant since the output is easily shown to be of the same period as the input, even though nothing may be said about the frequency content of the output. We will discuss the application of OS and other nonlinear filters in such situations. An example we will use to illustrate the optimality properties of such filters is the filtering of VEP (visual evoked potential) data.

Assume one has a data sequence x{n} with period N and length L. Assuming no knowledge of the frequency content, this type of data may be filtered by breaking the data sequence up into N sequences $\{y^i k\}$, $1 \le i \le N$, where $k = i+jN$ and j is such that $1 \le i+jN \le L$. Now each of the N sequences can be assumed to be constant in the absence

of noise and in the presence of noise we may filter them accordingly. Prior to our studies, most filtering was done using a sliding average on each constant sequence, which would be optimal for normally distributed i.i.d. noise. In this paper, we demonstrate the optimality of OS (order statistic) and other nonlinear filters for additive i.i.d. noise, bursty impulsive noise, and signed bursty noise, with specific applications.

Investigators have been guided in this filtering by two criterion; first to obtain the "most periodic" signal possible, using as short a window as possible and only one pass. We develop a measure we call the repeatability measure that estimates how periodic a signal is. We then use this measure to compare different filters of the same length in finding the optimal filter. We demonstrated the consistency of optimal OS filters for different window lengths and that a definite cutoff exists beyond which one only gets asymptotically better results by extending the window width.

BIBLIOGRAPHY

[1] D. REAGAN, Evoked Potentials in Psychology, Sensory Physiology, and Clinical Medicine, Chapman and Hall, London, England, 1972.

[2] J.E. DESMEDT, Visual Evoked Potentials In Man, Oxford University Press, Oxford, England, 1977.

[3] V.K. ROHATGI, An Introduction To Probability Theory and Mathmatical Statistics, John Wiley and Sons, New York, 1976.

[4] S.R. SEARLE, Linear Models, John Wiley and Sons, New York, 1971.

[5] J.W. TUKEY, Exploratory Data Analysis, Addison-Wesley, Reading, Massachusetts, 1971 (preliminary edition).

[6] N.C. GALLAGHER, JR., AND G.L. WISE, A Theoretical Analysis of Median Filters, IEEE Trans. Acoust., Speech, Signal Process., 29 (1981), pp. 1136-1141.

[7] A.C. BOVIK, T.S. HUANG, and D.C. MUNSON, A Generalizaton of Median Filtering Using Linear Combinations of Order Statistics, IEEE Trans. Acoust., Speech, Signal Process., 31 (1983), pp. 1342-1350.

[8] H.G. LONGBOTHAM and A.C. BOVIK, Paper accepted to be published in the IEEE Trans. Acoust, Speech, Signal Process., 1988.

[9] P.J. ROUSSEEUW, Least Median Of Squares Regression, Journal of the American Statistical Association, vol.79, number 388 (1984), Theory and Methods Section.

[10] K.M. WONG, S. CHEN, "Detection of Narrow-Band Sonar Signals Using Order S Statistical Filters", IEEE Trans. Acoust., Speech, Signal Process., 35 (1987), pp. 597-612.

[11] S.S. PEARLMAN, S. EISENHANDLER, P.W. LYONS, and M.J. SHUMILA, "Adaptive Median Filtering For Noise Elimination in Real Time TV Signals", IEEE Trans. on Communications, 35 (1987), pp. 646-652.

1988 USAF-UES RESEARCH INITIATION PROGRAM (RIP)

FOR SUMMER FACULTY RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Effects of Low Dose Soman on CNS Neurotransmitters

Prepared by:	Mohammed A. Maleque, Ph.D.
Academic Rank:	Associate Professor
Department:	Pharmacology
University:	Meharry Medical College Nashville, Tennessee
Collaboration:	Lt. Col. Stan L. Hartgraves USAF-SAM/RZB Brooks AFB San Antonio, TX 78235
Date:	September 20, 1989
Contract No.	F49620-85-C-0013/SB 5851-0360

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. I must thank the Universal Energy Systems for their concern and help to me in all administrative and directional aspects of this program.

My experience was rewarding and enriching because of many different influences. I met many Air Force Scientists and learned a lot about the needs and goals of various Air Force Research Projects. Lt. Col. S. Hartgrave's interest in this project truly served as a constant source of stimulation. The help of Dr. M. Murphy, Lt. S. Kerenyi, SSGt. D. Evans, AlC 1. Albano, and Ms. S. Miller was greatly appreciated. The administrative support and kindly treatment of Rodney Darrah, Project Director and his staff are certainly praiseworthy. Mr. Walter Williams' help in typing this report is highly appreciated.

Effect of Low Dose Soman on CNS Neurotransmitters

by

Mohammed A. Maleque, Ph.D.

Abstract

The level of central neurotransmitters was measured in rats following exposure to a single dose of soman (pinacolyl-methyl-phosphonofluoridate, 75 ug/kg) or CBDP (2-(0-cresyl)-4H-1:3:2-benzo dioxa-phosporin-2-oxide, 4 mg/kg) by High Pressure Liquid Chromotagraphy-Electrochemical (HPLC-EC) detection techniques. The central neurotransmitters included biogenic amines, namely, norepinephrine, epinephrine, dopamine, 5-hydroxytryptamine and the metabolite, 5-hydroxyindoleacetic acid. Five brain regions i.e.. striatum, hippocampus, amygdala, pyriform cortex and frontal cortex were used in this study. Some changes were observed with the levels of biogenic amines following a single exposure to soman or CBDP. This suggests subsequent alterations in physiological and behavioral changes to occur with respect to each transmitter function. This study has provided important information regarding early indication of biochemical changes associated with a single exposure to a minimum dose of soman or CBDP.

Effects of Low Dose Soman on CNS Neurotransmitters

INTRODUCTION

Organophosphate (OP) nerve agents like soman (pinacolyl-methylphosphonofluoridate) are among the most toxic substances known. Soman doses on the order of 10^{-10} to 10^{-11} times body weight produce prostration, convulsions, and death due to respiratory failure in most mammalian species studied. The threat to military personnel of exposure to OP agents has led to extensive research aimed at better defining the threat and at examining the levels of central neurotransmitters following exposure to low level soman dosage. Since human subjects can not routinely be exposed to OP threat agents, the use of animal models is necessary in this research. Although monkeys generally are considered to be the most useful animal model of human performance, they are expensive and time consuming to train, test and maintain (1). Thus, a recent editorial in Science has emphasized the use of alternative animal models for research on chemical defense bioeffects (2). The most predominant animal model in biology is the albino rat (3). Its use in toxicology and physiology is particularly dominant (4, 5). There is a very large chemical defense data based on this species. However, the rat is at least 10 times less sensitive than are primates on a dosage by weight basis (6-10).

The OP agents are extremely toxic because they inhibit the activity of acetylcholinesterase (AChE), an enzyme that hydrolyzes acetylcholine (ACh). Agent-induced reductions in AChE activity result in an excess of ACh, which can overstimulate and eventually block both neural and neuromuscular transmission. In addition, OP agents are known to produce diverse effects on both the peripheral and central nervous system (6, 11, 12, 23). Most studies have focused on their acute and chronic effects on cholinergic function more than transmitter induced functions (11, 14-18). Neurochemical studies have found steadily increased brain ACh levels (13, 19), decreased ACh receptors (20), decreased dopamine levels (15) and increased serotonin levels (21) following their use.

Although many studies have focused on the bioeffects of acute and repeated doses of soman (13, 22-25), the effect of low dose soman on the biogenic amines is lacking. The intent of the project is to investigate the effects of a single low dose soman on the biogenic amines in five brain regions. The purpose is to measure the levels of norepinephrine (NE), epinephrine (EPi), dopamine (DA), 5-hydroxytryptamine (5-HT), and 5-Hydroxyindoleacetic acid (5-HIAA) in pyriform cortex, amygdala, hippocampus, striatum and frontal cortex 1 hour following a single exposure to soman (75 ug/kg).

Method

Study Plan: Rats (Sprague-Dawley, male, wt. 150-250g) were exposed to soman at Brooks AFB-SAM-RZB. Soman and soman-treated rats were handled by RZB personnels. The biochemical studies were done at the Department of Pharmacology, Meharry Medical College, Nashville, Tennessee.

Analysis of Biogenic Amines: High Pressure Liquid Chromatography-Electrochemical (HPLC-EC) detection techniques (27) were employed for the analysis of biogenic amines, eg., NE, EPi, DA, 5-HT and their metabolite 5-HIAA. HPLC-EC systems comprise units from Waters Associates and Bioanalytical Systems. The procedure involves separation of sample constituents by liquid chromatography prior to their direct oxidation at a carbon electrode in a thin layer electrochemical cell. The technique is reasonably selective since two requirements, retention time and redox activity at the selected potential must be met simultaneously. The sensitivity of the instrument permits detection limits ca. 0.2 pmoles.

Mobile Phases: A stock buffer is prepared as follows: 13.6 g sodium acetate, 2.0 g heptane salforic acid, 0.2 g of EDTA and 90.0 ml of acetic acid are dissolved in 2 liters of deionized distilled water. To prepare mobile phase, sufficient

stock buffer is added to 170 ml of acetonitrile to make 2 liters of solution, which is then filtered through a 0.20-um membrane filter (rainin). After filtering and removing vacuum, 360 ml of UV-grade tetrahydrofuran (THF) are mixed gently with the solution. This mobile phase is capped at all times to prevent THF evaporation. In some instances, a mobile phase of 160 ml of acetonitrile and stock buffer to make 2 liters is used. This mobile phase is filtered and degassed after mixing.

Reagents: Norepinephrine bitartrate (NE), epinephrine bitartrate (EPi), dopamine (DA), serotonin (5-HT), and 5-hydroxyindoleacetic acid (5-HIAA) were obtained and used as received. Standard solutions containing 50 ug/ml and 50 ng/ml of each solute (as a mixture) were made up in 0.5 M perchloric acid containing 0.1% cysteine as an antioxidant.

Sample Preparation: Adult Sprague-Dawley rats were divided into four groups: control (received saline), vehicle (received propyleneglycol and ethanol 5%), soman (75 ug/kg in saline), CBDP (4 mg/kg in vehicle). Rats weighing 150-250g were decapitated, and their brains were removed rapidly and placed on chilled glass plates over ice. The brain striatum, hippocampus, amygdala, pyriform cortex, and frontal cortex were dissected. Tissue parts were placed in appropriate volume of ice cold 0.05 M $HClO_4$ and

sonicated for 3 minutes at 0°C, then centrifuged at 15,000g for 20 minutes at 4°C. The supernatant was removed and stored in Eppendorf polypropylene tubes at -70°C.

Sample Analysis: Typically 20 ul of the sample was injected; the sensitivity was 10 nA f.s. For the parallel adjacent arrangement, the potentials are 800/650 or 800/700 mV vs. Ag/AgCl 3 M NaCl. the current ratio is expressed as the peak current at the low-potential electrode divided by that at the higher potential. The higher-potential electrode was used for quantitative calculations. Calculations of supernatant concentrations were performed by direct comparison of sample peak heights to those of a 20-ul injection of an external standard containing all desired neurochemicals, using the higher-potential electrode. For the series arrangement, the upstream electrode was held at +850mV while the downstream electrode is at -150mV. the collection efficiency was defined at the downstream current divided by the upstream current. Quantitation using both electrodes was used by comparing peak heights of unknowns to standards. In cases where interferences were present, the more selective downstream elecrtrode was used for quantitation. Data were analyzed by student "t" test.

Results

Rats exposed to a single dose of soman (75 ug/kg) or CBDP (4 mg/kg) showed variable changes in the levels of biogenic

amines in different brain regions (Table 1). In the saline treated rats, the levels of NE were 164.91 ± 30.35, 163.18 ± 3.22, 142.25 ± 12.84, 163.85 ± 4.67 and 116.14 ± 6.79 pmoles/mg of wet tissue in the striatum, hippocampus, amygdala, pyriform cortex and frontal cortex, respectively. The levels of NE following a single exposure to soman were 187.69 ± 37.90, 151.21 ± 3.60, 135.50 ± 7.37, 148.49 ± 5.58 and 111.95 ± 4.80 pmoles/mg of wet tissue in the striatum, hippocampus, amygdala, pyriform cortex and frontal cortex, respectively (Fig. 1). It appears from these data that following exposure to soman (75 ug/mg) there was no change in the levels of NE, however, a small increase (13%) shown in the striatum was not significant. The levels of NE following CBDP were 134.82 ± 25.73, 163.97 ± 3.76, 191.44 ± 12.32, 161. 63 ± 5.94 and 106.95 ± 5.86 pmoles/mg of wet tissue in the striatum, hippocampus, amygdala pyriform cortex and frontal cortex, respectively. There was no significant difference between control and CBDP treated in any brain regions (Table 1, Fig. 1). Soman significantly increased the levels of EPi in the amygdala ($p < 0.01$, Table 1).

The levels of DA was increased in the amygdala following exposure to soman. The levels of DA were 70.81 ± 5.18 pmoles/mg of wet tissue before treatment and 111.77 ± 10.01 pmoles/mg of wet tissue after soman and 105.38 ± 17.42 pmoles/mg of wet tissue after CBDP exposure (Fig. 3).

The levels of 5-HT in the striatum were 89.96 ± 8.49 pmoles/mg of wet tissue in the untreated samples (saline group) and 77.23 ± 12.41 pmoles/mg of wet tissue after a single exposure to soman (75 ug/kg). The level of 5-HT in the striatum after exposure to CBDP (4 mg/kg) was 75.67 ± 9.23 pmoles/mg of wet tissue. Potentiations of 5-HT levels in the amygdala and frontal cortex following soman exposure were statistically significant.Compared to control values the level of 5-HIAA were incresed in hippocampus, pyriform cortex, frontal cortex and amygdala after soman treatment (Table I).

Discusssion

It is evident from these data that the levels of biogenic amines eg. NE, EPi, DA, 5-HT and the metabolite 5-HIAA vary from one brain region to another. These differences in the amount of neurotransmitters justify the involvement of each region for a specialized function. Soman is one of the most toxic nerve gases known (28, 29). It causes inhibition of cholinesterase activity, delayed neuropathy (30, 31, 32, 33, 13, 29) behavioral changes (11, 13, 34), altered neurotransmitter levels (13, 15, 19, 21). These effects are evident from acute studies with relatively large doses or from chronic studies with different dose levels. However, the effect of a single exposure to a median dose which is close to the LD_{50} is not established. Similarly, the effect due to a single

exposure to a non-toxic cholinesterase inhibitor (CBDP)is lacking. Results of this study have provided valuable information regarding early changes in the levels of neurotransmitters and the brain regions involved. Data obtained in this study showed that somn (75ug/kg) increased the level of NE in the hippocampus, EPi, DA, 5-HT and 5-HIAA in the amygdala (Table I). It also increased the level of 5-HT and its metabolite in the frontal cortex. CBDP or vehicle (PG+E) induced changes were not significant.

Recommendation

This study outlines the necessity for further studies with soman to determine (a) the minimum dose levels, (b) exposure time (c) degree of changes in biogenic amine levels, (d) identification of physiological functions or behavioral modality expressed due to change in transmitter levels and (e) histopathological changes in different brain regions. These parameters will help to find or develop agents that will prevent the onset of these changes or prevent deterioration of CNS functions or cure abnormalities induced by toxic anticholinesterase agents.

Table 1. The levels of norepinephrine (NE), epinephrine (Epi), dopamine (DA), 5-hydroxyindoleacetic acid (5-HIAA), and 5-hydroxytryptamine (5-HT) in saline, Soman (75 ug/kg), CBDP (4mg/kg) and vehicle (PG+E, 1 ml/kg propylene glycol + 5% ethanol) treated rat brain regions.

Brain Regions	Nor	Epi	DA	5-HIAA	5-HT
Striatum					
Control	164.91 ± 30.35(7)	1358.37 ± 266.45(7)	600.55 ± 67.97(7)	283.05 ± 33.28(7)	89.96 ± 8.49(7)
Soman	187.69 ± 37.90(5)	1666.0 ± 325.22(5)	425.88 ± 73.69(5)	321.06 ± 32.74(5)	77.23 ± 12.41(5)
CBDP	134.82 ± 25.73(5)	973.39 ± 42.92(5)	437.74 ± 59.76(5)	217.75 ± 22.15(5)	75.67 ± 9.23(5)
PG+E	111.86 ± 24.16(5)	985.70 ± 22.10(5)	293.39 ± 49.01(5)	207.96 ± 23.46(5)	64.38 ± 6.00(5)
Hippocampus					
Control	163.18 ± 3.22(7)	N.D	N.D	320.41 ± 5.83(7)	72.28 ± 4.32(7)
Soman	151.21 ± 3.60(5)*	N.D	N.D	402.58 ± 11.62(5)***	70.01 ± 3.73(5)
CBDP	163.97 ± 3.76(5)	N.D	N.D	318.07 ± 12.12(5)	80.40 ± 5.54(5)
PG+E	171.43 ± 1.15(5)	N.D	N.D	333.91 ± 21.11(5)	100.88 ± 18.05(5)
Amygdala					
Control	142.25 ± 12.84(6)	149.51 ± 10.29(5)	70.81 ± 5.18(5)	289.81 ± 17.84(6)	77.37 ± 5.78(6)
Soman	135.50 ± 7.37(5)	282.02 ± 32.45(5)**	111.77 ± 10.01(5)**	374.37 ± 12.78(5)**	104.91 ± 5.52(5)**
CBDP	191.44 ± 12.32(4)	264.31 ± 51.69(4)	105.38 ± 17.42(4)	406.29 ± 36.75(4)	88.38 ± 10.52(4)
PG+E	182.19 ± 29.62(4)	191.50 ± 17.95(4)	93.35 ± 1.78(4)	375.33 ± 53.28(4)	102.37 ± 5.54(4)
Pyriform Cortex					
Control	163.85 ± 4.67(7)	N.D	48.02 ± 1.84(7)	235.27 ± 16.02(7)	74.92 ± 3.98(7)
Soman	148.49 ± 5.58(5)	N.D	53.16 ± 3.43(5)	304.97 ± 20.75(5)*	91.02 ± 5.05(5)
CBDP	161.63 ± 5.94(5)	N.D	51.44 ± 1.74(5)	249.38 ± 20.35(5)	81.87 ± 10.02(5)
PG+E	169.58 ± 11.26(5)	N.D	50.37 ± 2.63(5)	275.89 ± 20.58(5)	92.04 ± 5.55(5)
Frontal Cortex					
Control	116.14 ± 6.79(7)	124.09 ± 4.56(7)	44.14 ± 3.34(6)	215.18 ± 14.17(7)	70.20 ± 2.47(7)
Soman	111.95 ± 4.80(5)	127.32 ± 10.22(5)	50.79 ± 8.18(5)	298.13 ± 15.06(5)**	94.54 ± 4.94(5)**
CBDP	106.95 ± 5.86(5)	126.92 ± 6.25(5)	45.27 ± 1.33(5)	200.82 ± 21.65(5)	63.50 ± 5.94(5)
PG+E	109.08 ± 3.18(5)	128.61 ± 5.53(5)	43.98 ± 2.71(3)	210.36 ± 15.57(5)	71.73 ± 4.24(5)

Units = pmoles/mg of wet tissue; N.D = Non-detectable; * = P <0.05; ** = P <0.01; *** = P <0.001

Fig 1. Levels of norepinephrine (NE) in the rat striatum, hippocampus, pyriform cortex, frontal cortex and amygdala before (C) and after treatment with soman (SOM, 75 ug/Kg) or CBDP (4 mg/Kg). Values are mean ± S.E from Table 1. * = $P<0.05$

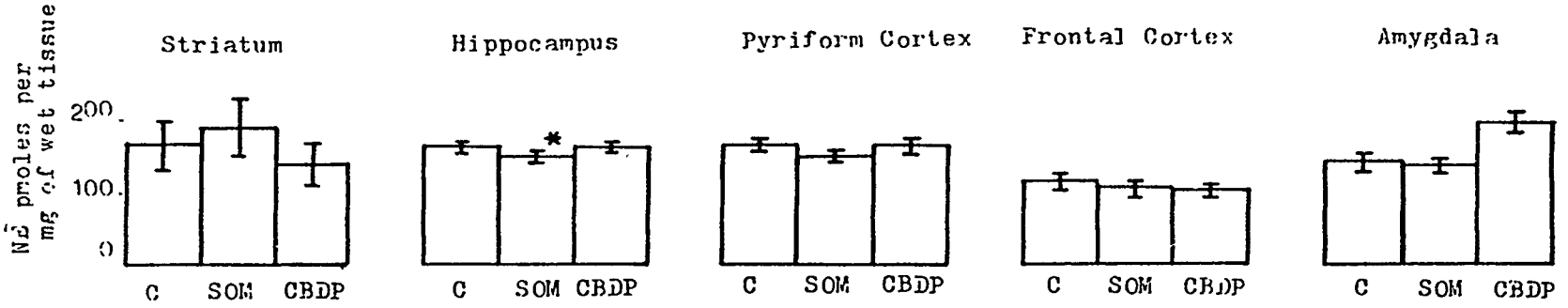


Fig 2. Levels of epinephrine (EPi) in the rat striatum, frontal cortex and amygdala before (C) and after treatment with soman (SOM, 75 ug/Kg) or CBDP (4 mg/Kg). Values are mean ± S.E from Table 1. ** = P<0.01

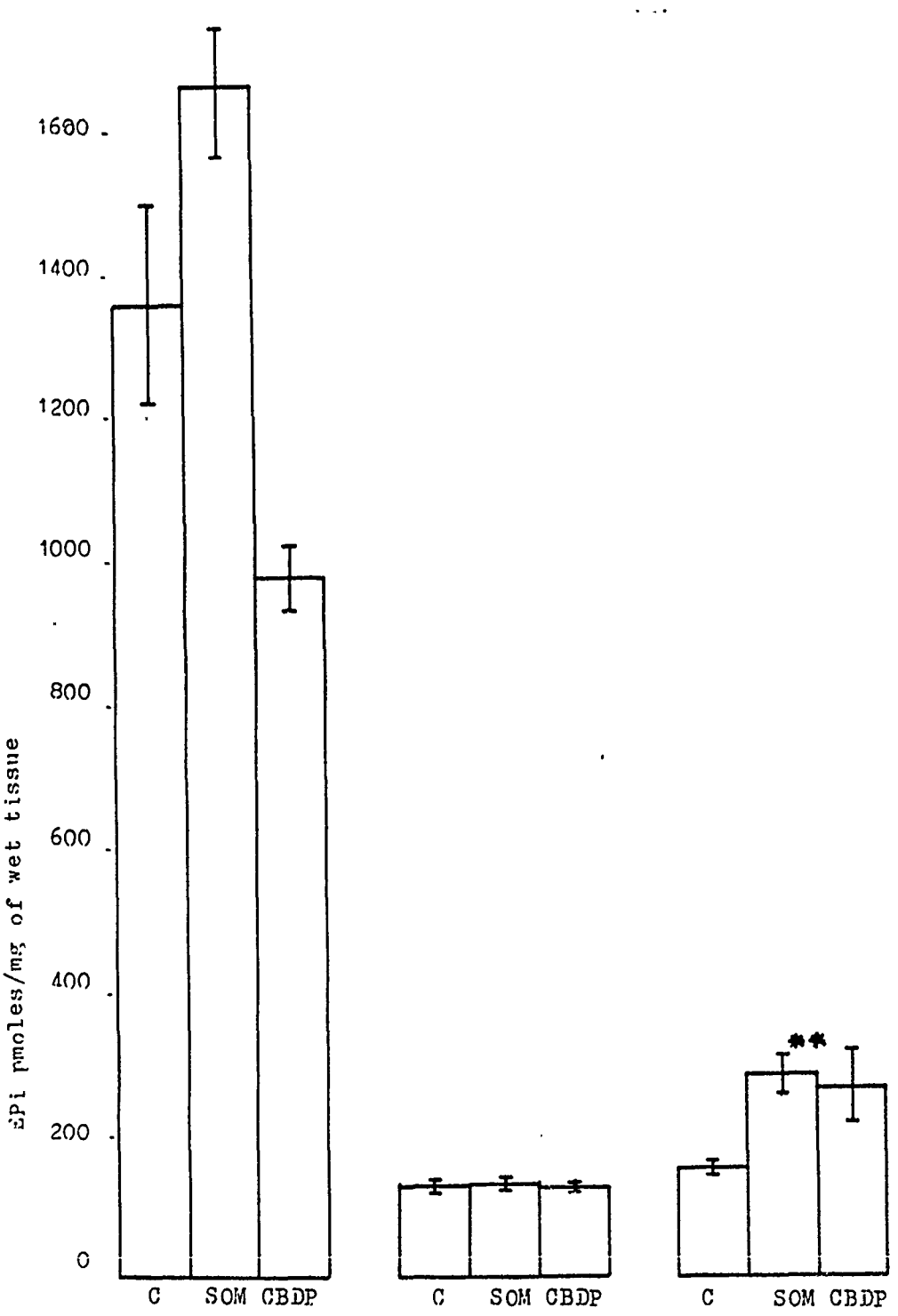


Fig 3. The level of dopamine (DA) in control (C) and following exposure to soman (SOM, 75 ug/Kg) or CBDP (4 mg/Kg) of rat striatum, pyriform cortex, frontal cortex and amygdala. Values represent mean ± S.E from Table 1. *= P<0.05

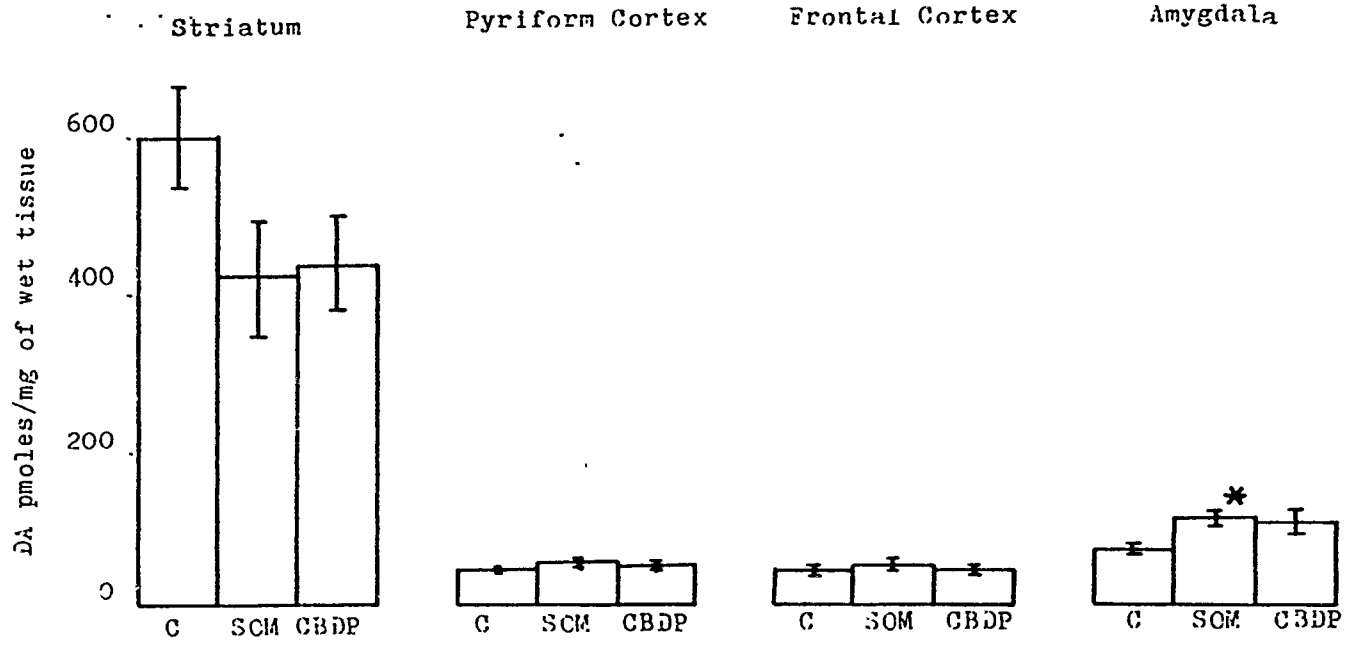


Fig 4. The level of 5-hydroxytryptamine (5-HT) from five brain regions before(C) and after treatment with soman (SOM, 75 ug/Kg) or CBDP (4 mg/Kg). Values are mean ± S.E. **=P<0.01

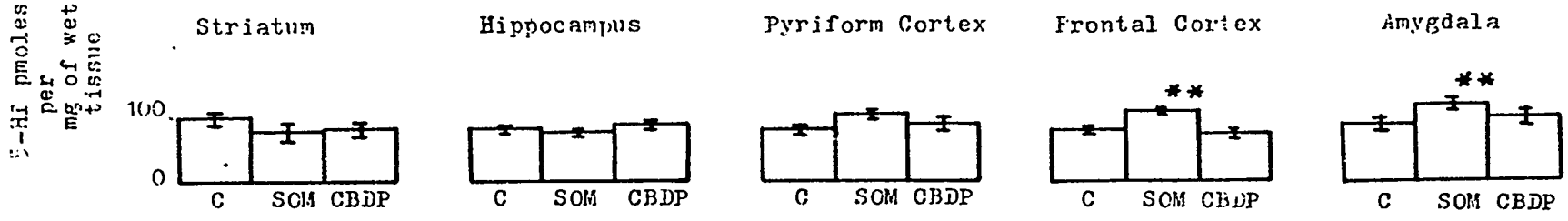
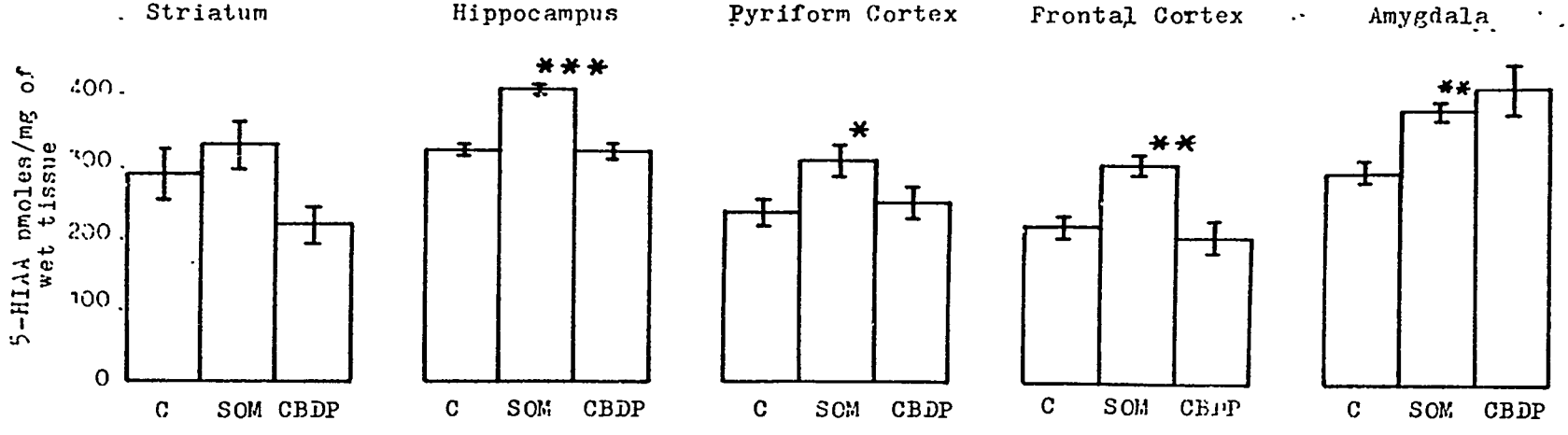


Fig 5. The effect of soman or CBDP on 5-hydroxyindoleacetic acid (5-HIAA) in the rat striatum, hippocampus, pyriform cortex, frontal cortex and amygdala. Control (C), soman (SOM, 75 ug/Kg), and CBDP (4 mg/Kg) are represented. Values are mean ± S.E. from Table 1. * = P<0.05; ** = P<0.01; *** = P<0.001



References:

1. King, F.A., C.J. Yarbrough, D.C., Anderson, T.P., Gordon, and K.G. Gould. Primates. Science, 240: 1475-1482, 1988.

2. Koshland, D.E., Jr. Biological Systems. Science, 240: 1385, 1988.

3. Calibrese, E.J. A reevaluation of the rat as a laboratory animal. In: Principles of Animal Extrapolation. R.L. Metcalf and W. Stumm (eds), pp 283-288. New York. John Wiley and Sons, 1983.

4. Altman, P.L., and D.D. Katz. Inbred and genetically defined strains of laboratory animals, Part 1. Mouse and Rat. Fed Amer. Soc. Exp. Biol. 1979.

5. Oser. B.L. The rat as a model for human toxicological evaluation. J. Toxicol. Environ. Health. 8: 521-542, 1981.

6. Jimmerson, V.R., T.M. Shih, D.M., Maxwell, and J. Hernandez. Preliminary attempts to improve a rat for the study of soman toxicity and its treatment. Proc. 6th Med. Chem. Def. Biosci. Rev., US Army Med. Res. Inst. Chem. Def. pp. 271-273, Aberdeen Proving Ground, MD 1987.

7. Bryson, P.K. and T.M. Brown. Reactivation of carboxylesterase hydrolase following inhibition by 4-nitrophenyl organophosphinates. Biochem. Pharmacol. 34: 1789-1794, 1985.

8. Gupta, R.C. and W.D. Dettbarn. Iso-OMPA-induced potentiation of soman toxicity in rat. Arch Toxicol. 61: 58-62, 1987.

9. Maxwell, D.M., K.M. Brecht, and B.L. O'neil. The effect of carboxyesterase inhibition on interspecies differences in soman toxicity. Toxicol. Let. 39: 35-42, 1987.

10. Maxwell, D.M., D.E. Lenz, W.A. Groff, A., Kaminski, and H.L. Froehlich. The effects of blood flow and detoxification on in vivo cholinesterase inhibition by Soamn in rats. Toxicol. Appl. Pharmacol. 88: 66-76, 1987.

11. Davis, C.S. and Richardson, R.J. Organophosphorus compounds in experimental and clinical neurotoxicology, eds. P.S. Spencer and H.H. Schaunburg, Baltimore, Williams and Wilkins. 527-544, 1980.

12. Holmsteadt, B. Pharmacology or Organophosphorus Cholinesterase Inhibitors. Pharmacol. REV. 11: 567-688, 1986.

13. Russell, R.W. and Overstreet, D.H. Mechanisms underlying sensitivity to organophosphorus anticholinesterase compounds. Progress in Neurobiology. 28: 97-129, 1987.

14. Dasheiff, R.M., Einberg, E., and Grenell, R.G., Sarin. Adrenergic-cholinergic interaction in rat brain. Exp. Neuro. 57: 549-560, 1977.

15. Freed, V.H., Martin, M.A., Fang, S.C., and Kar, P.P. Role of striatal dopamine in delayed neurotoxic effects of organophosphorus compounds. Eur. J. Pharmacol. 35: 229-232, 1976.

16. Johnson, M.K., Organophosphorus esers causing delayed neurotoxic effects. Mechanism of action and structural activity studies. Arch. Toxicol. 34: 259-288, 1975.

17. Karczman, A.G. Acute and long-lasting central actions of organophosphorus agents. Fund. Appl. Pharmac. 4:51-517, 1984.

18. McLeod, C.B., Jr., Singer, A.W., and Harrington, D.G., Acute neuropathology in soman-poisoned rats. Neurotoxicology 5:53-58, 1984.

19. Russell, R.W., Cholinergic system in behavior: The search for mechanisms of action. A. Rev. Pharmac. Toxicol. 1: 242, 1981.

20. Seaman. Brain dopamine receptors. Pharmacol. Rev. 32: 229-313, 1981.

21. Fernando, J.C.R. Hoskins, B. and Ho, L.K. Effect on striatal dopamine metabolism and differential motor behavioral tolerance following chronic cholinesterase inhibition with diisopropylflurophosphate. Pharmac. Biochem. Behav. 20: 951-957, 1984.

22. Bignami, G., Rosic, N., Michalek, H., Milosevec, M. and Gatti, G.L. Behavioral toxicity of anticholineserase agents. Methodological, neurrochemical and neurophysiological aspects. In: Behavioral Toxicology, pp. 155-215, 1975. Eds. B. Weiss and V.G. Laties. Plennum, New York.

23. Blick, D.M., Murphy, M.R., Brown, G.C., Yochmowitz, M.G., and Hartgraves, S.L. Primate equilibrium performance following soman exposure. Effects of repeated dialy exposures to low soman doses. Proc. 6th Med. Chem. Def. Biosci. Rev. Aug. 4-6, 317P, 1987.

24. Jovic, R.C. Correlaction between signs of toxicity and some biochemical changes in rat poisoned by soman. Eur. J. Pharmacol. 25: 159-164, 1974.

25. Petras, J.M. Soman neurotoxicity. Fund. Appl. Toxicol. 1:242, 1981.

26. Fonnum, F. A rapid radiochemical method for the determinattion of choline acetyltransferase. J. Neurochem. 24: 407-409, 1975.

27. Mayer, G.S., Shoup, R.E. Simultaneous Multiple Electrode Liquid Chromatographic-Electrochemical Assay for Catecholamines, Indolamines and Metabolites in Brain Tissue. J. Crom. 255-533-544, 1983.

28. Heath. D.F. Organophosphorous Poisoning. 300-319, Perganon Press, 1961.

29. Taylor, P. Anticholinesterase agents. InGoodman and Gilman's The Pharmacological Basis of Therapeutics, 6th ed., eds. A.G. Gilman, L.S. Goodman, and A. Gilman, pp. 100-119, New York, Macmillan, 1980.

30. Johnson, M.K. The aging reaction of inhibited neuropathy target esterase-fundamental studies and toxicological significance. In. cholinesterases, M. Brzin, E.A. Barnard and D. Sket (eds). 463-482, Walter de Gruyter, 1984.

31. Lotti, M., C.E. Becker, and M.J. Aminoff. Organophosphate polyneuropathy: Pathogenesis and prevension. Neurology. 34:6580662, 1984.

32. Koelle, G.B. Organophosphate Poisoning-An Overview. Fund. Appl. Toxicol. 1:120-134, 1981.

33. O'Brien, R.D. Insectisides, Action and Metabolism. New York, Academic Press, 1967.

34. Hoskins, B., Fernando, J.C.R., Dulaney, M.D., Lim, D.K., Liu, D.D., Watnabe, H.K., and Ho, 1.K., Relationship between the nurotoxicities of soman, sarin and tabun and acetylchoninesterase inhibition. Toxicol. Lett, 30: 121-129, 1986.

Final Report

1987 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCINTIFIC RESEARCH

Conducted by the

UNIVERSAL ENERGY SYSTEMS, Inc.

Cleansing of Bone-marrow by lymphokine Activated Killer Cells (LAK-Cells)

Prepared by	Parsottam J. Patel;DVM,PhD
Academic rank	Associate Professor
Department, School and University	Department of Microbiology School of Medicine Meharry Medical College
Research Location	Department of Microbiology School of Medicine Meharry Medical College
USAF Research Focal Point	Col. Richard N. Boswell
Contract No.	S-760-MG-131

SPECIFIC AIMS

The main objective of the study was to determine whether Interleukin-2 activated killer (LAK cells) are capable of destruction of tumour cells within the bone-marrow, without harming the normal bone-marrow cells. The following specific aims were set forth:

1. To establish optimum conditions for generating LAK-cells from mouse bone-marrow.

2. To determine effect of LAK-cell on normal bone-marrow stem cells, and

3. To determine if LAK-cells could be successfully used to eliminate tumour cells deliberately spiked into normal marrow cells.

METHODS AND RESULTS

Mice

Experiments were performed using C57BL/6 mice. The mice were purchased from the Trudeau Institute, Inc., Saranac Lake, N.Y. The mice were specific pathogen free stock originally derived from a germ-free mouse colony. While under experimentation, the mice were kept in a barrier sustained caging system fitted with a 0.22 micron Hepafilter, and fed ad libitum. The care and maintenance of animals strictly adhered to guides and policy published by the NIH.

Preparation of Interleukin-2

Interleukin-2 was prepared according to the method provided by Dr. John Yannelli, of Biotherapeutics, Inc., Franklin, TN. Twenty female Sprague-Dawley rats were sacrificed by placing them in a CO_2 chamber. The spleens, aseptically removed and teased apart in RPMI 1640, were passed through 22 gauge wire mesh screens; after which they were centrifuged twice at 1000x g for ten minutes at 4° C. The pellet was resuspended in ten mls RPMI 1640. The cells were passed through needles (26g-30g) to obtain a single cell suspension, then counted in white blood cell dilution fluid. Also, a viable count was obtained. The remaining cells were centrifuged at 1000x g for ten minutes at 4° C. The cell volume was adjusted to 5×10^6 viable cells/ml. Concanavalin A (Con A) was added to 50 ml of cells which had been previously placed in a 150 cm^2 culture flask (Falcon 2098, Becton Dickinson and Co., Oxnard, CA) at a concentration of 5 ug/ml. The suspension was mixed thoroughly and incubated at 37° C and five percent CO_2 in air for 24 hours to harvest. The suspension was centrifuged in a GSA rotor at 1000x g for 20 minutes at 4° C.

The supernatant was kept at 4° C until it was passed over a Sephadex G-50 column to remove the Con A. The Sephadex G-50 column (12 ml column prepared from G-50 that had been swollen overnight, then degassed) was equilibrated with five column volumes of complete RPMI. The flow through was collected in sterile bottles. All work was done as sterilely as possible at 4° C. After the flow through had been collected, it was centrifuged in a GSA rotor at 1000x g for 20 minutes to remove Sephadex particles that may clog the filter to be used in the subsequent sterile filtration step. After sterile filtration, small aliquots from each bottle were removed and placed in a 24 well culture plate at 37° C in a five percent CO_2 incubator to check for contamination.

The thymocyte proliferation assay waas used to test for the activity of the interleukin-2 preparation.

Thymocyte Proliferation Assay

The thymocyte proliferation assay is an accepted method of determining the IL-2 activity. Three mice were sacrificed by cervical dislocation. The thymus was removed and gently teased apart in RPMI 1640. The cells were filtered through 200 gauge wire mesh, washed twice in RPMI 1640 and centrifuged at 1000x g for 10 minutes at 4° C.

Following resuspension in ten mls of RPMI 1640, the cells were passed through 26 gauge then 30 gauge needles to obtain a single cell suspension then counted to obtain the total number of cells. Remaining cells were washed in RPMI 1640 and centrifuged at 1000 x g for 10 minutes at 4° C. The final cell pellet was resuspended at a concentration of 1×10^6 cells/ml in RPMI 1640 complete media. The cells were cultured in 96 well, U-bottom tissue culture platees (Nunclon, Denmark) in a 0.1 ml volume of media to which a 0.1 ml volume of appropriately diluted Cr-TCGF (Collaborative Research, Lexington, MA) or a crude preparation of TCGF (prepared as described above) was added. All cultures were incubated at 37° C in a five percent CO_2 in air mixture. After 48 hours of incubation, the cultures were pulsed labeled with one uCi of ^{3}H-Thymidine (specific activity 100 mCi/mg). Cultures were incubated an additional 24 hours then harvested by use of a cell harvester (Skatron, Inc., Sterling, VA).

Figures 1 and 2 present the results of thymocyte proliferation assay. The results are averages of quadruplate cultures in three different experiments. The data are presented as cpm (counts per minute) of 3H-thymidine versus the percentage of IL-2 containing supernatants. It shows that a significant stimulation could be achieved at 25% Il-2 concentration which peaks at 50% level of IL-2 supernatant. The IL-2 supernatant so prepared were freeze stocked and used in all other subsequent experimentation.

Generation of LAK cells

Initially, experiments wre performed to establish optimal conditions for the generation of LAK-cells from the mouse bone-marrow. This involved culturing mouse bone-marrow cells under varying conditions in terms of: (1) number of cells cultured, (2) amount of IL-2 added and (3) length of time in culture. The LAK-activity so generated was then tested in a Cr^{51} release assay as described below. As a result of these studies, a standard protocol was adopted to generate LAK cells. This involved culturing bone-marrow cells ina 50 ml volume of the culture medium aat $2x10^6$ cells/ml density in the presence of 25% IL-2 supernate for a period of 5 to 7 days at 37° C under 5% CO_2 atmosphere.

^{51}Cr Release Assay

Tumour (EL4) cells actively growing in culture were used as target cells. A 5 ml suspension of target cells at a density of 10^6 cells/ml was washed two times and suspended in 0.5 ml medium. To this 200 uCi ^{51}Cr was added and incubated at 37° C for 1 hour. The cells were then washed three times and resuspended at 10^5 cells/ml. The 51Cr labeled target cells were incubated with effector cells at different target:effector cell ratios. Triplicate test combinations were assayed in 200 ul. Volume placed in round bottomed 96 well plates, incubated at 37o C to 4 hours. 75 ul supernatant aliquots were removed from each well and counted to determine the % specific lysis of target cells according to standard published procedures. A 1:75 target:effector ratio provided approximately 85% specific lysis of the target cells, indicating successful generation of LAK cells.

LAK effect on normal bone-marrow cells

It is established tha while LAK cells do lyse fresh or cultured tumour cells, they have no harmful effect on normal human blood cells. No information is available about effect of LAK-cells on normal bone-marrow stem cells. Accordingly, mouse bone-marrow stem cells were co-cultured with syngeneic bone-marrow stem cells for a varying time period and then the viabillity of bone-marrow cells determined in an trypan-blue dye exclusion test. These studies revealed no harmful effect by LAK cells on bone-marrow stem cells, as the viability of stem cells culture with LAK cells was comparable to that cultured in the absence of LAK cells. These studies, however should be extended in terms of culturing stem cells exposed to LAK cells to differentiate into colony forming units (CFU). If CFU studies proves that the LAK does not harm the bone-marrow stem cells, then the experiments could be performed to determine the ability of LAK cells to destroyed deliberately spiked tumour cells within the bone-marrow cell population.

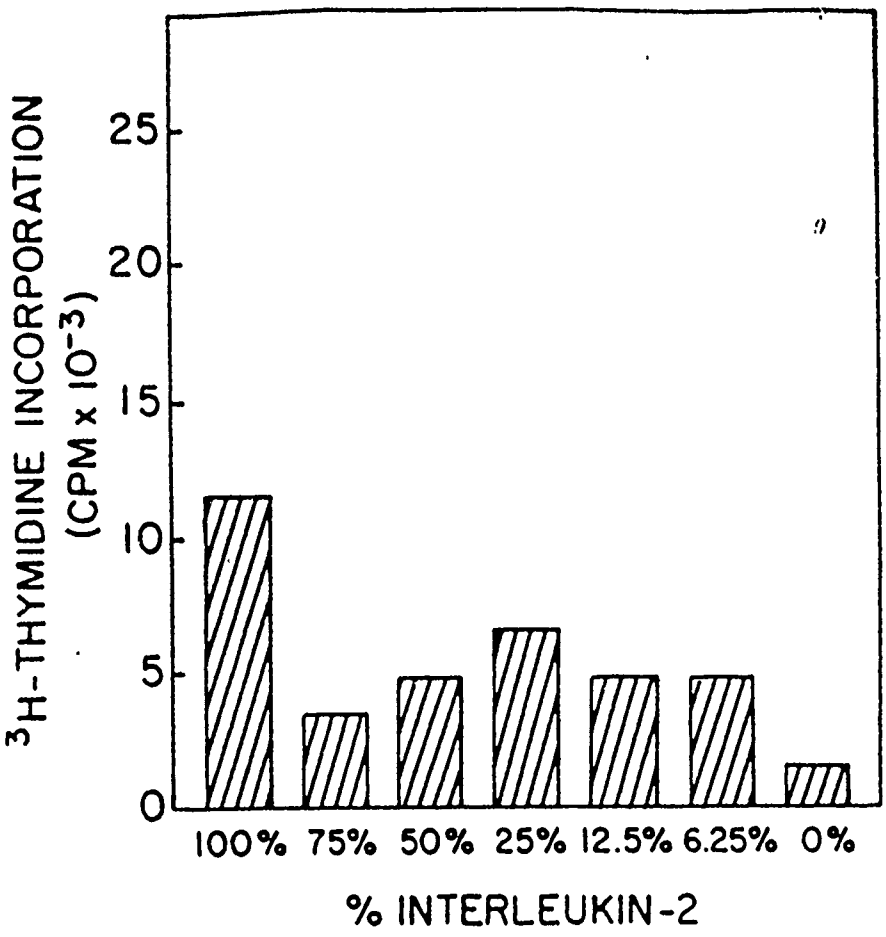


Figure 1. **Measurement of the interleukin-2 activity as determined by the thymocyte proliferation assay.** Thymus cells were cultured in the presence of various concentrations of interleukin-2 containing supernatant. Cellular proliferation was determined by tritiated thymidine incorporation.

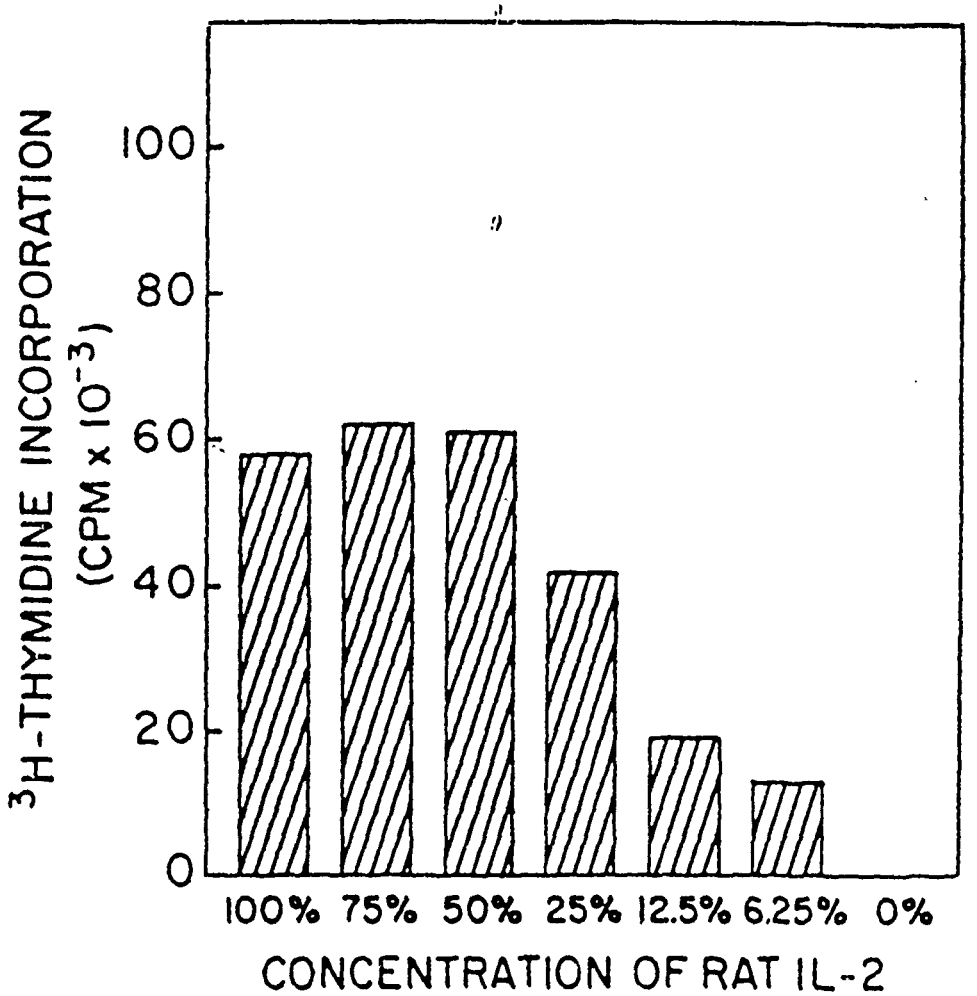


Figure 2. Measurement of the interleukin-2 activity as determined by the thymocyte proliferation assay. Thymus cells were cultured in the presence of various concentrations of rat interleukin-2 containing supernatant. Cellular proliferation was determined by tritiated thymidine incorporation.

AFOSR-RIP FINAL REPORT

TRANSCUTANEOUS OXYGEN DELIVERY

William Z. Plachy, Ph.D.
Professor of Chemistry and Biochemistry
San Francisco State University

March 15, 1990

Objective:

To investigate the rate at which air oxygen can diffuse into and across the stratum corneum.

ESR Spin Labels in the Skin:

The stratum corneum, or SC, does not normally contain a sufficient number of free radical species to give Electron Spin Resonance, or ESR, spectroscopic signals. Thus we can add small amounts of stable nitroxide spin probes to the SC and observe their ESR signals in the absence of significant background signals. Perdeuterated spin probes are used to insure that the ESR line width is as narrow as possible and thus has the greatest possible sensitivity to dissolved dioxygen. A broad ESR signal is observed from a population of the spin probe that is immobilized in a viscous SC microenvironment. Another population of the spin probe is in a nonpolar (lipid) and rather fluid (nonviscous) microenvironment. This latter population has a narrow intrinsic ESR line width, and it is this ESR signal that we use for the dioxygen studies described below.

Skin Respiration:

Interest in the degree to which human skin absorbs dioxygen directly from the air goes back at least 150 years. Measurements dating back to about 1930 established that the total amount of dioxygen absorbed via the skin is probably less than 1% of the dioxygen requirement of a resting adult. The absorption of dioxygen by the skin increases with temperature with an activation energy of about 15 kcal per mole. The rate of air oxygen absorption by the skin is not sufficient to satisfy the dioxygen requirement of the viable epidermis (in spite of the fact that the viable epidermis is less than 0.1 mm removed from an infinite dioxygen reservoir). Therefore most of the dioxygen required by the viable epidermis must be satisfied by "blood oxygen" diffusing from the dermis rather than by "air oxygen" diffusing across the stratum corneum. The stratum corneum thus appears to be an efficient diffusional barrier to air oxygen uptake.

In these experiments we are attempting to measure this barrier function of the stratum corneum to dioxygen uptake from the air. We take advantage of the paramagnetism of dioxygen which causes dissolved dioxygen to increase the ESR line width of our spin probes in a manner which depends on the product of the dioxygen diffusion coefficient times the dioxygen solubility (see the THEORY section). We monitor these line width effects indirectly by observing the line intensity of the spin probe while in the second derivative mode of the ESR spectrometer. Notice that we are observing the dioxygen while it is in the stratum corneum. By monitoring the rate at which dioxygen saturates the stratum corneum we can calculate the apparent dioxygen diffusion coefficient in the stratum corneum.

Sample Preparation:

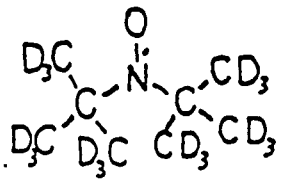
Adult male hairless mice were used to obtain stratum corneum sheets by sequential trypsinization. Full thickness skin was excised from the mice that were 9-12 weeks old. The subcutaneous fat was scraped off with a #10 surgical blade and the skin floated dermis side downward on 0.5% trypsin in phosphate-buffered saline (PBS) at 37°C. After one hour the trypsin solution was replaced with fresh trypsin solution and the dermis was peeled off from beneath. The PBS was then replaced with fresh trypsin solution and the epidermis was incubated for another hour. The resulting epidermal tissue was then rinsed with PBS several times with vortexing to dislodge the remaining adherent unrelated cells from the stratum corneum sheet, and stored at -70°C. Water was removed by vacuum as required.

Experimental Procedures:

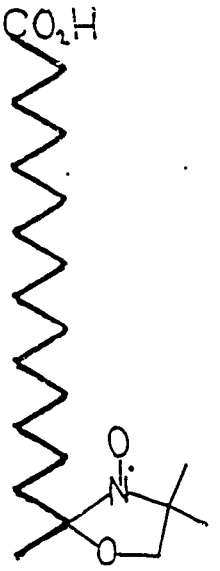
The SC sheet was spin labeled by exposure to the spin probe vapor until a sufficiently strong ESR signal was obtained. The spin labeled sheet was allowed to equilibrate overnight before oxygen diffusion experiments were performed. Measurements were also performed on a lipid film with a composition similar to that of SC lipid. All samples were mounted on flat quartz plates which were placed in a quartz dewar in the ESR sample chamber (Varian, model E-12 ESR Spectrometer). The temperature and the composition of the gas flowing over the sample was controlled. The gas composition was changed (air to nitrogen or nitrogen to air) in a time less than 0.2 seconds. The ESR intensity data was logged by a microcomputer interfaced data system (Keithley, model 195A) at a rate of 20 points per second. Data analysis was done with the ASYSTANT scientific software package (Macmillan, New York).

OUR SPIN PROBES

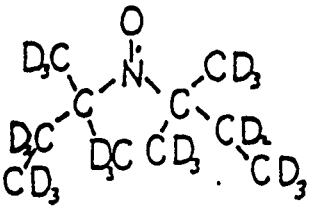
pdDTBN $K_{par} = 15$



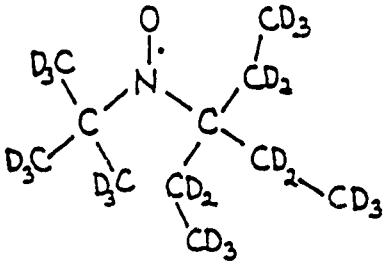
17DSA ^{15}N, d_8



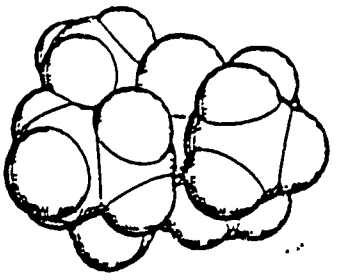
pd5,5NO $K_{par} = 160$



pd4,7NO $K_{par} = 300$



4,7NO CPK Model



THEORY OF THE ESR EFFECT OF DIOXYGEN

The probe ESR line width in the presence of oxygen is:

$$W(x,t) \equiv W(N_2) + k * D(O_2) * [O_2(x,t)]$$

For oxygen diffusing into a slab of thickness L

$$[O_2(x,t)] \equiv [O_2(t=\infty)] \left\{1-\frac{4}{\pi} \cos(\alpha x) * e^{-\beta t}\right\}$$

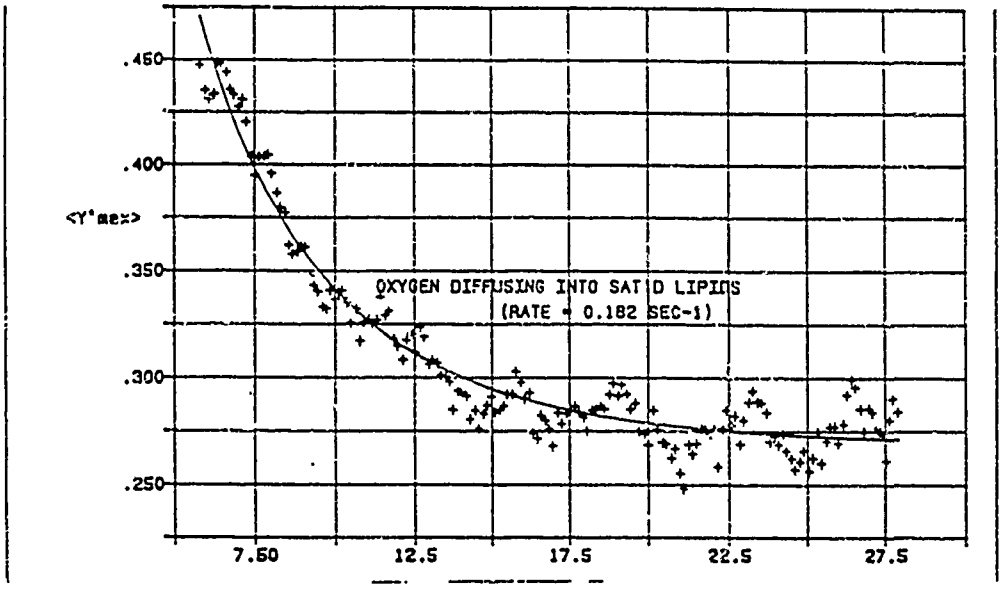
after $t > \frac{0.2}{\beta}$, where $\alpha = \frac{\pi}{2L}$ and $\beta = \frac{\pi^2 D(O_2)}{4L^2}$

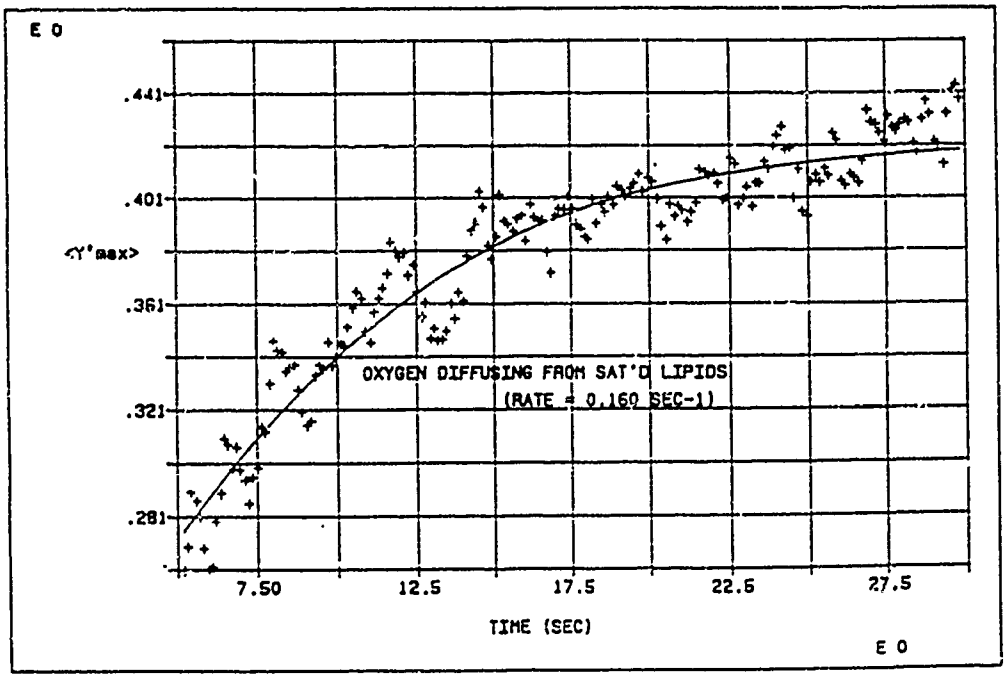
The second derivative signal intensity is averaged over x,

$$\langle Y''(t)\rangle_x = \frac{1}{L}\int_0^L Y''(x,t)\,dx = \frac{1}{L}\int_0^L k' W(x,t)^{-3}\,dx,$$

giving the final expression which is fit to the observed data.

KINETIC CURVES WITH LEAST SQUARE FIT





OXYGEN DIFFUSION AND SOLUBILITY IN VARIOUS SOLVENTS

SOLVENT	T(°C)	10^5 X D (cm^2/s)	α (Bunsen)
ethanol	20	1.64	0.226
ethanol	29.6	2.64	0.222
CCl_4	25	3.82	0.279
benzene	29.6	2.89	0.208
cyclohexane	29.6	5.31	0.259
DMSO	25	3.23	0.0496
water	25	2.31	0.0285
the values below were determined in our laboratory			
DMPC bilayers (T_c=24°C)	30	1.3	0.10
" "	20		0.027
SC lipid mixture	30	0.21	
" "	25	0.15	
SC sat. lipid mixture	22	0.09	
Mouse SC (3.4 μm)	30	0.0025	
Human sunburn scale (3.8 μm)	22	0.0027	

Some Advantages of ESR Spin Labels in Skin Research

1. **Sensitivity**-less than one microgram of spin probe is required.
2. **Ease of labeling**-the sample may be labeled from the spin probe vapor.
3. **Sample**-samples can be as small as 100 micrograms and heterogeneous.
4. **No background signal**-this is both an advantage and a disadvantage.
5. **Multiple probe sites**-the probes may report from two or more microenvironments of a heterogeneous sample.
6. **Viscosity information**-the spin probes are sensitive to the viscosity of their microenvironment via their ESR line width.
7. **Polarity information**-the spin probes report on the polarity of their microenvironment via the ESR hyperfine splitting.
8. **Sample variables**-spectra can be studied as a function of O_2, T, and R.H.
9. **Unique spin probes**-variety of probes available, some are biomimetic.
10. **Relation to other techniques for skin research**-ESR spin probes provide microscopic sample information that is often complementary to information obtained by IR or NMR, or from macroscopic techniques like DSC.

Discussion and Conclusion:

ESR spin probes are well suited for physical studies of the stratum corneum. The spin probes are small molecules at high dilution that should not seriously distort the structure of the stratum corneum. The ESR spectra of the spin probes provides us with molecular information about the detailed microenvironment of similar foreign molecules that we might add to the skin (drugs, cosmetics, etc.).

In the present study we demonstrate that dioxygen diffuses rapidly in a SC lipid mixture. The dioxygen diffusion coefficient for these lipids is typical of that found in other lipids (e.g. a lecithin dispersion). The apparent dioxygen diffusion coefficients for whole SC sheet are smaller than these lipid values by a factor of 50 to 100. We propose that dioxygen diffuses in the SC via a lipid pathway. If we assume that the effective length of the lipid pathway is about eight times the true thickness of the SC, then the tortuosity of the lipid pathway would be sufficient to account for the slow dioxygen diffusion rate we observe in the SC. Using our measured SC dioxygen diffusion coefficient, 0.003×10^{-5} cm^2/s, we can calculate that the maximum flux of air oxygen through the stratum corneum can account for only one third of the dioxygen demand of the viable epidermis.

Chronology:

The chronology of our efforts and accomplishments while funded under the AFOSR-RIP grant is supplied below:

January 1, 1989: Work begins during the Winter Intersession (three weeks between the Fall and Spring semesters). I was joined by a new graduate student, Ms. Mary Hatcher, who continues to work on this project as part of her MS thesis research. Ms. Hatcher was supported under this grant in the summer of 1989.

May 11-13, 1989: I attended the NIH sponsored workshop entitled "The Loop Gap Resonator in ESR Spectroscopy" held at the National Biomedical ESR Center in Milwaukee Wisconsin. This meeting proved to be very helpful in understanding the theory and practice of these new ESR sample resonators. In late May we placed our order for the loop gap accessory. The total price was $8,800 including variable temperature capability. Funds were obtained from this AFOSR RIP research grant ($1,500), a Chemistry and Biochemistry Department research mini-grant ($2,500), a university mini-grant ($2,500), a contribution from Alza Corporation ($800), and from NIH research grant funds of Dr. Ann Walker ($1,500).

June: I submitted an NIH AREA grant on ESR studies of molecular oxygen in skin. This proposal would not have been possible without the preliminary findings we obtained while our work was sponsored by the AFOSR RIP grant. In January '90 I was informed that the NIH AREA project would be funded for a period of two years with a start date in late March.

June and July: I continued work in the lab with Mary Hatcher developing our new ESR techniques for monitoring oxygen in skin and in model systems. We adapted the theory of gaseous diffusion into and out of a plane for the purpose of our ESR experiments. At this time we also took delivery of the new loop-gap resonator ESR accessory. The new resonator was installed and several trial experiments were performed. New techniques were developed for handling the small sample sizes required with a loop gap resonator (a few microliters at most). Our previous experience with ESR samples contained in thin wall gas permeable Teflon tubes was very helpful. Many of the sample handling techniques that I saw demonstrated at the Milwaukee Loop Gap Resonator conference were put to use.

In Spring 1989 and again in Fall 1989 (while I was on a Sabbatical leave) I was able to travel and give several invited talks based on our experiments conducted to date. These talks included:

April 25-30, 1989: I attended the Society for Investigative Dermatology meeting in Washington D.C. We presented two poster papers based in part on work in progress under this grant. Partial travel assistance for the trip was obtained from this ASOSR RIP grant.

August14-18: I was an invited speaker at the Gordon Conference "Barrier Function of Mammalian Skin" (R. Potts and R. Guy, Chairs). New Hampshire.

August 18-23: Travel in New England, Quebec City, and Montreal.

Septembe· ^: I was invited to visit Upjohn Pharmaceuticals, Kalamazoo, I the guest of Dr. Michael Holland. I presented a talk on my research and we discussed possible collaboration.

October 18-20: Attended the Pacific Conference on Chemistry and Spectroscopy (Pasadena, CA, Dick Keyes organizer). Mary Hatcher and I presented a paper on our ongoing dioxygen studies in the stratum corneum.

October 30: Seminar presented to the Dermatology group at UCSF (Dr. H. Maibach), "Spin Label Studies of Dioxygen in the Stratum Corneum."

November 8-10: I attended a symposium on "Present Developments and Future Challenges in Transdermal Drug Delivery" sponsored by Leiden University in The Netherlands (Dr. Harry Bode, Chair). I presented a paper on our dioxygen diffusion work. Some travel funds for this trip were obtained from this AFOSR RIP grant.

December 4-7: I presented an invited talk to the "Bioengineering and the Skin" symposium held in New York City. I made new industrial contacts at this meeting.

Summary:

We have accomplished many of the original objectives of this project. The project is clearly feasible and worth further effort. The results to date have been presented in meetings and in invited lectures as outlined above. Journal publication will follow in the summer of 1990 after further experiments are performed with animal tissue. The funding I have recently obtained from the NIH will allow me to complete the project, including investigation of the effect of perturbation enhancers on the rate of transcutaneous oxygen transport. I wish to thank UES and AFOSR for their timely support in the important early stages of this continuing project.

FINAL REPORT NUMBER 135
REPORT NOT ACCEPTABLE AT THIS TIME
Dr. Ralph Peters
760-7MG-091

FINAL REPORT

A COMPUTER MODEL OF THE HUMAN SYSTEMIC ARTERIAL TREE

THOMAS R. ROGGE
PROFESSOR
DEPARTMENT OF ENGINEERING SCIENCE AND MECHANICS
BIOMEDICAL ENGINEERING PROGRAM
IOWA STATE UNIVERSITY
AMES, IOWA

JANUARY 1990

Abstract

A computer model to simulate blood flow in the human arterial tree has been developed and two numerical techniques, finite difference and the finite-element, have been used to solve the governing equations. This model includes the effect of +G acceleration, the effect of stenosis, the nonlinear effect of material behavior, and a modified Windkessel model for the terminal beds. Examples of results are given demonstrating how the model may be used to simulate blood flow with different parameters which govern the system.

Acknowledgement of Sponsorship

Iowa State University wishes to acknowledge the support of this project by the following agency:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
BOLLING AFB, DC

The principal investigator wishes also to thank the above sponsor, Universal Energy System and the personnel of the School of Aerospace Medicine Brooks AFB, San Antonio, Texas.

Introduction

The problem studied in this research is to develop a computer model to simulate the human arterial system. This model should be sufficiently flexible so that an accurate representation of arterial blood flow could be produced. In order to simulate the blood flow in the arteries this model requires a significant amount of accurate data. Most of this data is available in the literature and at present a certain amount of this data has been collected and used in this study.

Description of the Model

The differential equations governing the pressure and flow in the arterial segments to be used in this study are given below. The continuity equation is

$$\frac{\partial Q}{\partial x} + A_0 C_0 \frac{\partial P}{\partial t} + A_0 C_1 P \frac{\partial P}{\partial t} + SP = 0$$

where P, Q, x, and t are the pressure, flow, length and time; A_0 is the lumen area at $P=P_0$, C_0 and C_1 are a measure of the tube compliance and S is the seepage factor for an arterial segment. The momentum equation is

$$C_u \frac{\partial Q}{\partial t} + \frac{2Q}{A} \frac{\partial Q}{\partial x} + \frac{A}{\rho} \frac{\partial P}{\partial x} + \frac{8C_v \pi \mu}{\rho A} Q - Ag_x = 0$$

where ρ is the fluid density, μ is the fluid viscosity, g_x is the component of G acceleration along the artery and the wall shear stress used in the momentum equation is approximated by a term of the form (Weerappuli)

$$\tau_w = -\frac{\rho}{2\pi R} \left[\frac{8C_v \pi \mu}{\rho A} Q + (C_u - 1) \frac{\partial Q}{\partial t} \right]$$

where C_u and C_v are parameters which can be obtained from the Wormersly solution and which depend upon the frequency of the flow pulse.

The constitutive equation which relates the pressure and lumen area used in the continuity equation is

$$A(x) = A_0(x) \left[1 + C_0^a (P - P_0) + C_1^a (P - P_0)^2\right]$$

where

$$C_0^a = \frac{1}{\rho a_0^2}, \quad C_0^1 = (C_0^a)^2$$

and a_0 is the wave speed in the tube.

For the case of a stenosis in an arterial segment the relation between pressure drop and flow is

$$\Delta P_s = A_s Q_s + B_s Q_s |Q_s| + C_s \frac{dQ_s}{dt}$$

where Q_s is the flow through the stenosis and A_s, B_s, and C_s are independent of pressure for a rigid stenosis and are functions of pressure if the stenosis is compliant in nature (Young, Stergiopulos). The terminal endings of the arterial system are modelled by an equation (a modified Windkessel model) of the form

$$C \frac{dp}{dt} - R_1 C \frac{dB}{dt} + \frac{P}{R_2} - \left(1 + \frac{R_1}{R_2}\right) B = 0$$

where B is the flow, P the pressure, R_1 and R_2 are resistances, and C is the compliance of the distal beds.

The model includes the effect of +G acceleration and a seepage term to account for the number of small arteries emanating from the main arteries. The model does not take into account the effect of viscoelastic behavior of the arterial segments.

A schematic of the arterial segments used to construct the arterial tree is shown in Figure 1 and the appropriate segments are listed in Table 1. The above differential equations govern the flow and pressure in each of the segments with the parameters chosen for that particular segment. At the bifurcations of the segments into two separate branches the pressure is assumed to be equal at the entrance to each branch and the sum of the flows into the branches is set equal to the incoming flow from the parent artery. The input boundary condition is taken as the pressure pulse at the entrance to the aorta, and can be determined by measuring the heart pressure pulse.

Two different methods have been investigated to solve the resulting set of differential equations. One method is the finite-element method, which is discussed in detail in Porenta et al, and the other is a finite-difference method. The reason that both methods were attempted was that the finite-element method displayed some instability in the solution of the governing equations. This instability is at present being examined and studied. At this time the finite-difference method is stable for a linearized set of equations. The effect of each nonlinear term in the governing equations is being investigated with each method to see what the effect of these terms will have on the solution and also the stability of the solution. The use of the finite difference is not as convenient for describing the behavior of the flow and pressure at the bifurcations and also is not as amenable to inclusion of stenosis in the arterial segments. For this reason the finite-element will continue to be studied as a possible method to solve the system of governing equations.

The initial terminal model of the distal beds was taken to be pure resistance instead of the modified Windkessel model. This can easily be accomplished by setting the compliance in the modified Windkessel model equal to zero. The value of this compliance is not an easy parameter to obtain and in order to quickly test the model for other types of errors, such as proper connectivity etc., the value of the compliance was set equal to zero. The data for the resistances and compliances of the distal beds is still being gathered and will be used in the more

complete model of the system. All the results will eventually be compared to other models which have been used as well as actual measured results obtained from the literature.

Results obtained for a linear model using a normal heart pressure pulse are shown in Figs. 2 and 3. These results are from the finite-difference method and can be compared with the results obtained from the finite-element method shown in Fig 4 and 5. As would be expected these results compare very well. It is to be noted though that in the second cycle of both the pressure and flow the finite-element begins to show some numerical instability. This instability can be controlled by increasing the value of the viscosity but at the expense of decreasing the peak pressure. The pressure and flow at a number of locations in the arterial system is shown. In Fig.2 the input pressure pulse is shown (Node 1), with pressure pulses from the Thoracic Aorta artery (Node 74), Abdominal Aorta artery between the splenic and renal arteries (Node 134), and the Abdominal Aorta artery just proximal to the Illiac bifurcation (Node 151). Flows are shown at the same location in the same arterial segments. Figs 4 and 5 shows the same results from the same arterial segments but in this case the node numbers are different.

Figs 6 through 11 show the effect of +G acceleration on the pressure and flow waveforms. The input pressure pulse is a normal pressure pulse and it is evident that to get a more accurate representation of the effect of high G acceleration that a modified heart pulse should be used. In the case considered here it is assumed that the arterial tree is in a standing position with the arms extended perpendicular to the torso so that the arms experience a zero gravity force. The remaining portions of the arterial system thus experiences the same acceleration force which is assumed to be a constant. The equations solved here are the linear set and comparison with previous results show an increase of flow and pressure as the G force is increased from zero to 3 g.

It is expected that there is some control mechanism in the body which will change the terminal resistances and possibly constrict certain arteries so that the flow to the lower body will be restricted while flow to the upper body and head will be enhanced. These additional modules of the model need to yet be studied and included into the model.

Figs 12 and 13 show the pressure and flow at the same points in the arterial tree when an abnormal heart pressure pulse is used as the input pressure. The pressure pulse is for an aortic valvular stenosis (Balar, et al), and one can compare the results with the normal results shown in Figs. 2 and 3.

Figs 14 and 15 show the pressure and flow pulse for normal input pressure pulse and zero acceleration of gravity but includes five cycles to demonstrate that after the first cycle the transients are damped out and the solution is periodic.

The above is merely the beginning of simulating blood flow in the human arterial tree and a significant amount of comparison and collecting of data still need to be completed. Below are listed additional areas which need to be investigated and included into the model.

Recommendations for Continued Development of the Model

The following improvements of the model are planned:

1. As additional data can be obtained the number of segments will be increased to somewhere near 127 rather than the 45 which are presently being considered.

2. The modified Windkessel model for termination of the branches and the main arteries will be added and the appropriate values of the compliance and resistance will be determined. It may also be necessary to include in this lumped model the effect of inertia.

3. Comparison with published model studies will be investigated and if clinical data can be obtained comparison of the model results with this data will be undertaken.

4. The effect of +G acceleration will be studied given that not all the arterial segments are affected the same. In this study some form of feedback control mechanism may need to be included in the model to simulate the changes which the cardiovascular system undergoes when subjected to these types of forces.

5. The inclusion of viscoelastic effects in the constitutive equation for the arterial segments.

6. The study of abnormal heart pressure pulses and of arterial stenoses and attempt to determine if the pressure and flow waveforms can be used to diagnose such problems.

References

Balar, S. D., Rogge, T. R., and Young, D. F., (1989) Computer Simulation of Blood Flow in the Human Arm, J. Of Biomechanics, 22, 691-697.

Porenta, G, Young, D. F., and Rogge, T. R. (1986) A Finite-Element Model of Blood Flow in Arteries Including Taper, Branches, and Obstructions, J. Biomechanics, 7, 77-91.

Stergiopulos, N. (1987) Pulsatile Flow Through a Compliant Stenosis, M.S. Thesis, Iowa State University, Ames, Iowa

Weerappuli, D. P. (1986) Simulation of Pulsatile Flow in Arteries Using the Finite-Element Method. Ph.D. Thesis, Iowa State University, Ames, Iowa.

Young, D. F. (1979) Fluid Mechanics of Arterial Stenoses, Trans. Am. Soc. of Mech Eng., J. Biomechanical Engr., 101, 157-175.

Caption for Figures

Figure 1 Schematic of Human Systemic Arterial Tree

In all the figures which follow the pressure and flow are plotted at the four points:

Node 1. The proximal end of the system (Entrance of the Aorta)
Node 74. The Thoracic Aorta Artery
Node 134. The Abdominal Aortic Artery between the Splenic and Renal Arteriesnal Arteries
Node 151. The Abdominal Aorta just before the Illiac Bifurcation

Figure 2 Pressure waveforms with normal pressure pulse input and the acceleration of gravity G=0. (Finite Difference)

Figure 3 Flow waveforms with normal pressure pulse input and the acceleration of gravity G=0. (Finite Difference)

Figure 4 Pressure waveforms with normal pressure pulse input and the acceleration of gravity G=0. (Finite-Element)

Figure 5 Flow waveforms with normal pressure pulse input and the acceleration of gravity G=0. (Finite-Element)

Figure 6 Pressure waveforms with normal pressure pulse input and the acceleration of gravity G=1g. (Finite Difference)

Figure 7 Flow waveforms with normal pressure pulse input and the acceleration of gravity G=1g. (Finite Difference)

Figure 8 Pressure waveforms with normal pressure pulse input and the acceleration of gravity G=2g. (Finite Difference)

Figure 9 Flow waveforms with normal pressure pulse input and the acceleration of gravity G=2g. (Finite Difference)

Figure 10 Pressure waveforms with normal pressure pulse input and the acceleration of gravity G=3g. (Finite Difference)

Figure 11 Flow waveforms with normal pressure pulse input and the acceleration of gravity G=3g. (Finite Difference)

Figure 12 Pressure waveforms with aortic valvular stenosis pressure pulse input and the acceleration of gravity G=0. (Finite Difference)

Figure 13 Flow waveforms with aortic valvular stenosis pressure pulse input and the acceleration of gravity G=0. (Finite Difference)

Figure 14 Pressure waveforms with normal pressure pulse input and the acceleration of gravity G=3g. (Finite Difference showing five cycles)

Figure 15 Flow waveforms with normal pressure pulse input and the acceleration of gravity G=3g. (Finite Difference showing five cycles)

HUMAN SYSTEMIC ARTERIAL TREE

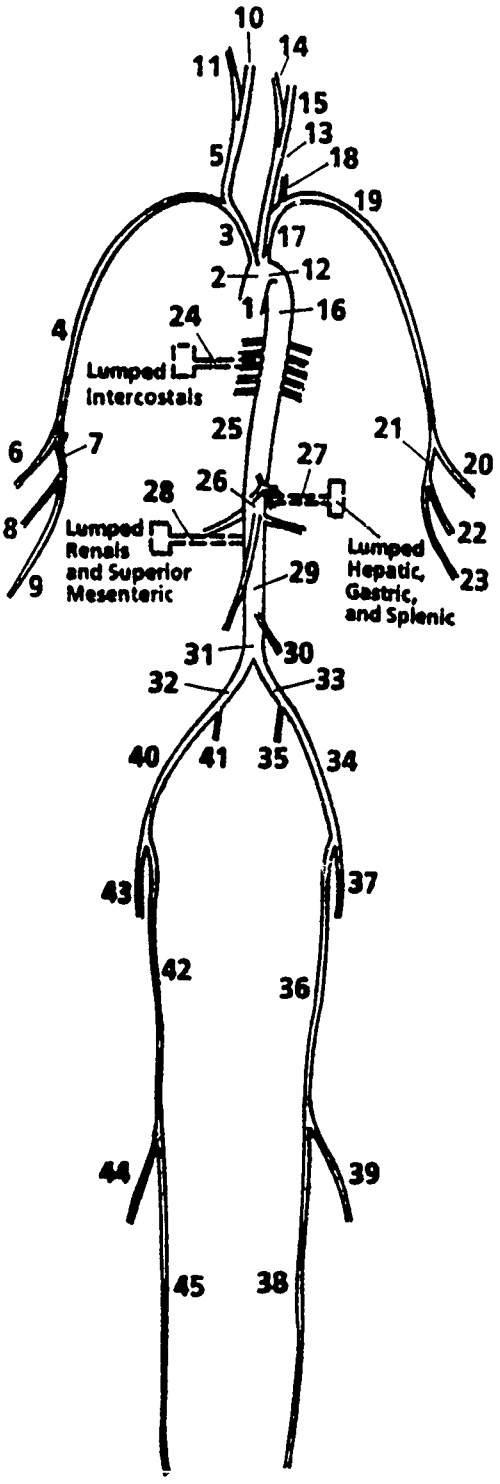


Figure 1

Table 1

Seg. #	Name	Length	Prox. R	Dist. R	COa (E-5)	C1a (E-10)	R (E9)	Prox. A	Distal A	CO (E-5)	C1 (E-10)
1	Ascending Aorta	2.100	1.170	1.140	3.120	9.734	0.000	4.301	4.083	0.531	19.469
2	Aortic Arch A	2.100	1.120	1.030	3.120	9.734	0.000	3.941	3.333	0.531	19.469
3	Inominate	6.400	0.550	0.380	1.470	2.161	0.000	0.950	0.454	0.895	4.322
4	R. Subclavian	45.400	0.380	0.270	1.470	2.161	0.000	0.454	0.229	0.895	4.322
5	R. Carotid	3.700	0.290	0.310	2.730	7.453	0.000	0.264	0.302	0.748	14.906
6	R. Radial	24.500	0.170	0.130	0.744	0.554	5.280	0.091	0.053	0.597	1.107
7	R. Ulnar A	4.900	0.210	0.200	0.744	0.554	0.000	0.139	0.126	0.597	1.107
8	R. Interosseous	8.600	0.085	0.085	0.943	0.889	84.300	0.023	0.023	0.706	1.778
9	R. Ulnar B	20.400	0.200	0.180	0.564	0.318	5.280	0.126	0.102	0.479	0.636
10	R. Internal Carotid	5.400	0.160	0.080	1.160	1.346	13.900	0.080	0.020	0.802	2.691
11	R. External Carotid	5.400	0.160	0.080	1.160	1.346	13.900	0.080	0.020	0.802	2.691
12	Aortic Arch B	2.100	1.010	0.910	3.120	9.734	0.000	3.205	2.602	0.531	19.469
13	L. Carotid	3.700	0.290	0.310	2.730	7.453	0.000	0.264	0.302	0.748	14.906
14	L. Internal Carotid	5.400	0.160	0.080	1.160	1.346	13.900	0.080	0.020	0.802	2.691
15	L. External Carotid	5.400	0.160	0.080	1.160	1.346	13.900	0.080	0.020	0.802	2.691
16	Thoracic Aorta A	10.500	0.910	0.710	3.050	9.302	0.000	2.602	1.584	0.576	18.605
17	L. Subclavian A	3.400	0.380	0.360	1.470	2.161	0.000	0.454	0.407	0.895	4.322
18	Vertebral	17.800	0.175	0.175	0.943	0.889	6.010	0.096	0.096	0.706	1.778
19	L. Subclavian B	42.000	0.360	0.270	1.470	2.161	0.000	0.407	0.229	0.895	4.322
20	L. Radial	24.500	0.170	0.130	0.744	0.554	5.280	0.091	0.053	0.597	1.107
21	L. Ulnar A	4.900	0.210	0.200	0.744	0.554	0.000	0.139	0.126	0.597	1.107
22	L. Interosseous	8.600	0.085	0.085	0.943	0.889	84.300	0.023	0.023	0.706	1.778
23	L. Ulnar B	20.400	0.200	0.180	0.564	0.318	5.280	0.126	0.102	0.479	0.636
24	Intercostals	8.000	0.160	0.160	0.888	0.789	1.390	0.080	0.080	0.678	1.577
25	Thoracic Aorta B	10.500	0.710	0.470	2.950	8.703	0.000	1.584	0.694	0.635	17.405
26	Abdominal Aorta A	5.300	0.470	0.470	2.850	8.123	0.000	0.694	0.694	0.689	16.245
27	Celiac (lumped)	2.500	0.310	0.310	2.620	6.864	0.000	0.302	0.302	0.794	13.729
28	Renals & S. Mesenter	7.400	0.350	0.350	2.620	6.864	0.930	0.385	0.385	0.794	13.729
29	Abdominal Aorta B	0.100	0.470	0.470	2.840	8.066	0.000	0.694	0.694	0.695	16.131
30	Inferior Mesenteric	5.900	0.140	0.140	1.930	3.725	6.880	0.062	0.062	0.939	7.450
31	Abdominal Aorta C	0.100	0.430	0.430	2.760	7.618	0.000	0.581	0.581	0.734	15.235
32	R. Common Illiac	5.800	0.300	0.290	2.010	4.040	0.000	0.283	0.264	0.935	8.080
33	L. Common Illiac	5.800	0.300	0.290	2.010	4.040	0.000	0.283	0.264	0.935	8.080
34	L. External Illiac	8.300	0.280	0.210	1.890	3.572	0.000	0.246	0.139	0.940	7.144
35	L. Internal Illiac	8.000	0.200	0.180	1.830	3.349	1.220	0.126	0.102	0.939	6.698
36	L. Femoral	43.600	0.220	0.200	1.310	1.716	0.000	0.152	0.126	0.854	3.432
37	L. Deep Femoral	15.400	0.250	0.180	0.655	0.429	4.770	0.196	0.102	0.541	0.858
38	L. Posterior Tibial	29.800	0.220	0.120	0.509	0.259	4.770	0.152	0.045	0.440	0.518
39	L. Anterior Tibial	32.600	0.120	0.120	0.431	0.186	5.590	0.045	0.045	0.382	0.372
40	R. External Illiac	8.300	0.280	0.210	1.890	3.572	0.000	0.246	0.139	0.940	7.144
41	R. Internal Illiac	8.000	0.200	0.180	1.830	3.349	1.220	0.126	0.102	0.939	6.698
42	R. Femoral	43.600	0.220	0.200	1.310	1.716	0.000	0.152	0.126	0.854	3.432
43	R. Deep Femoral	15.400	0.250	0.180	0.655	0.429	4.770	0.196	0.102	0.541	0.858
44	R. Anterior Tibial	32.600	0.120	0.120	0.431	0.186	5.590	0.045	0.045	0.382	0.372
45	R. Posterior Tibial	29.800	0.220	0.120	0.509	0.259	4.770	0.152	0.045	0.440	0.518

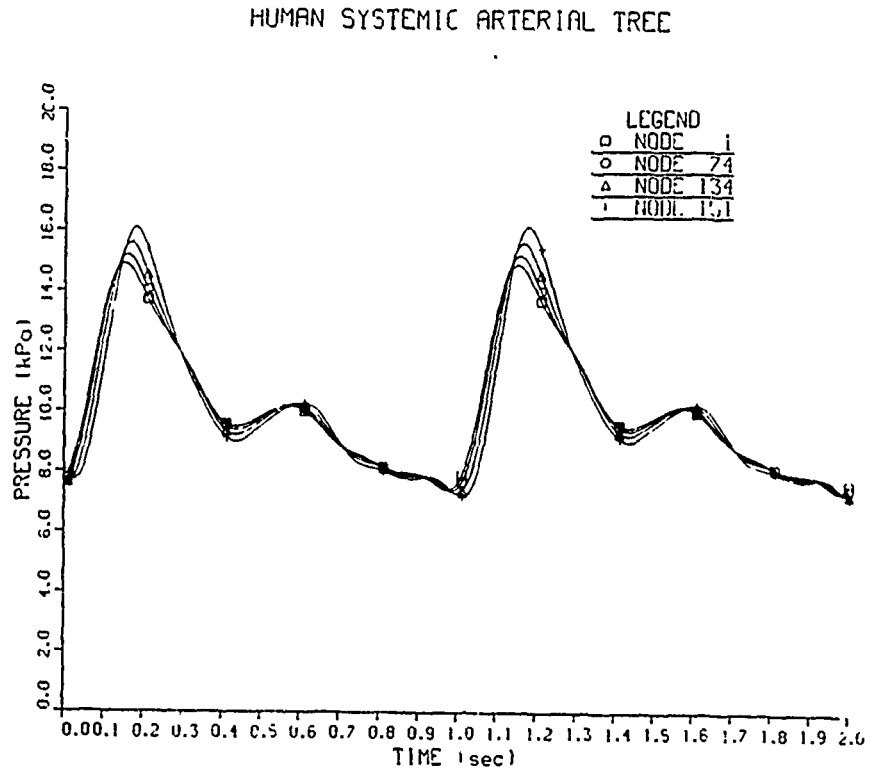


Figure 2

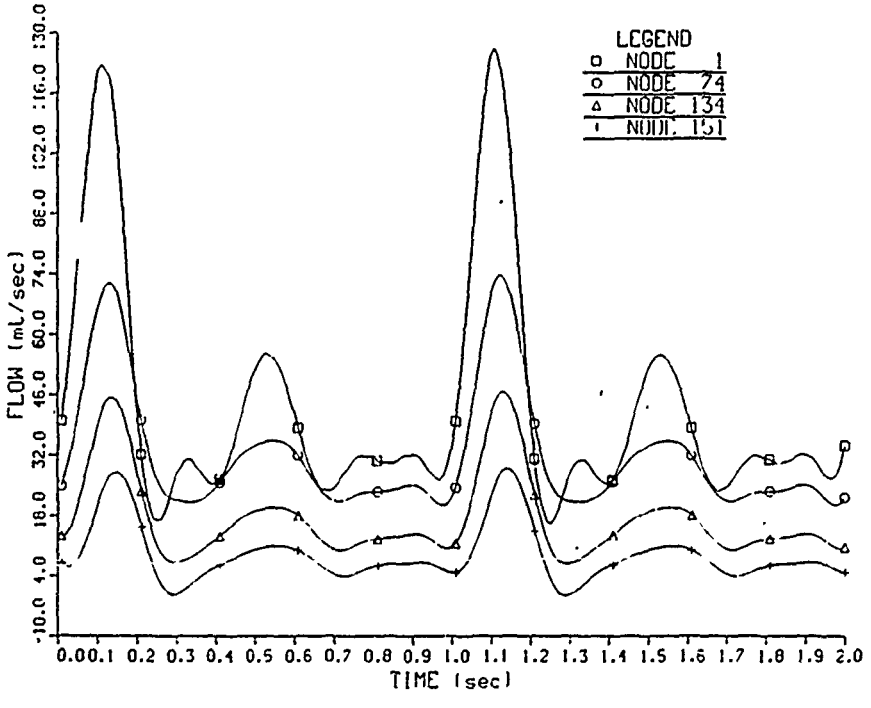


Figure 3

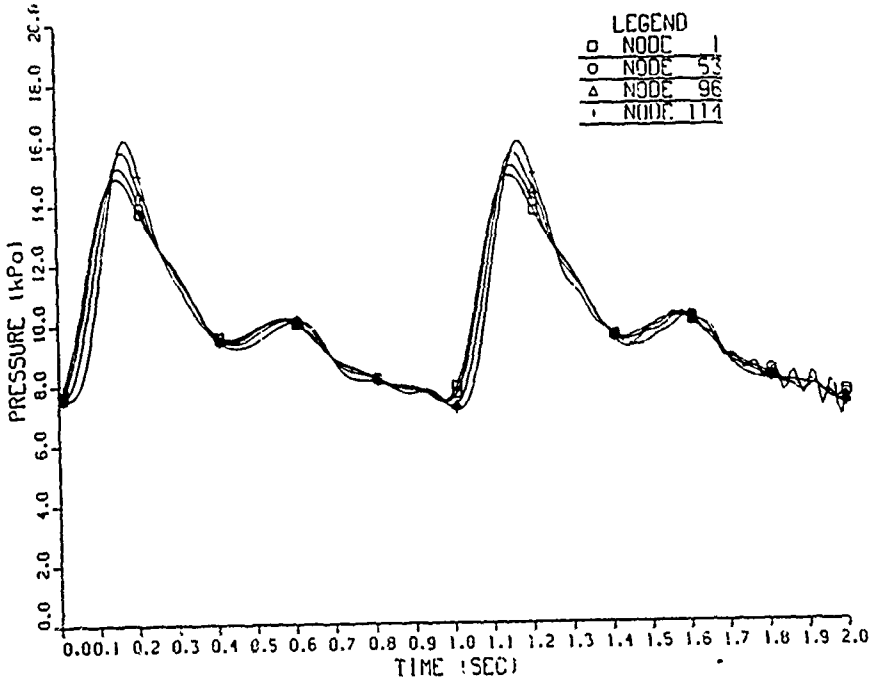


Figure 4

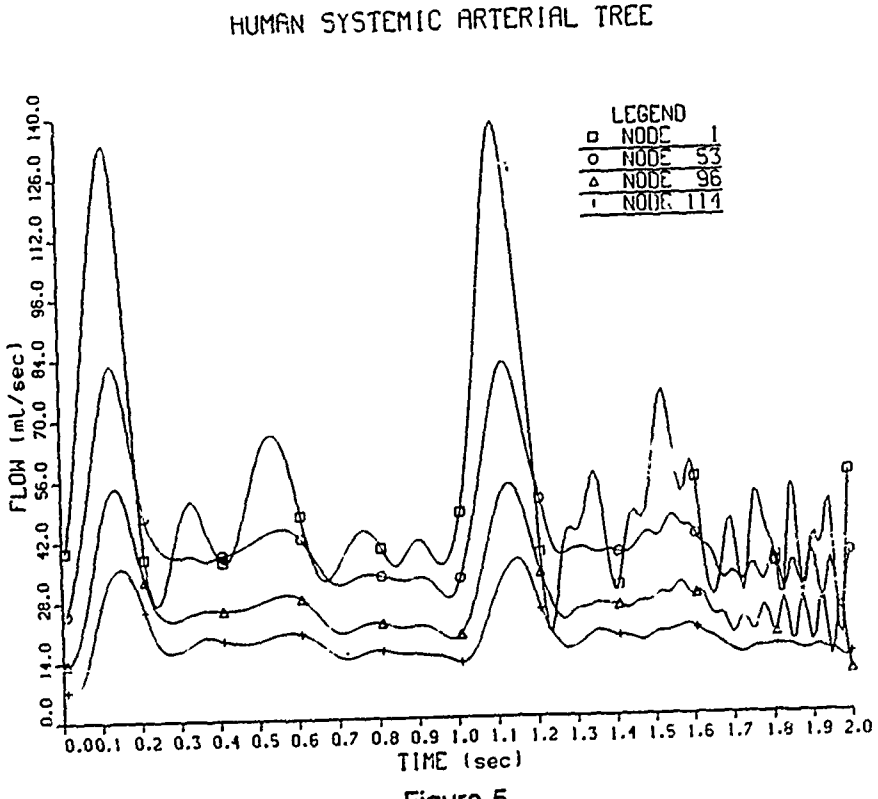


Figure 5

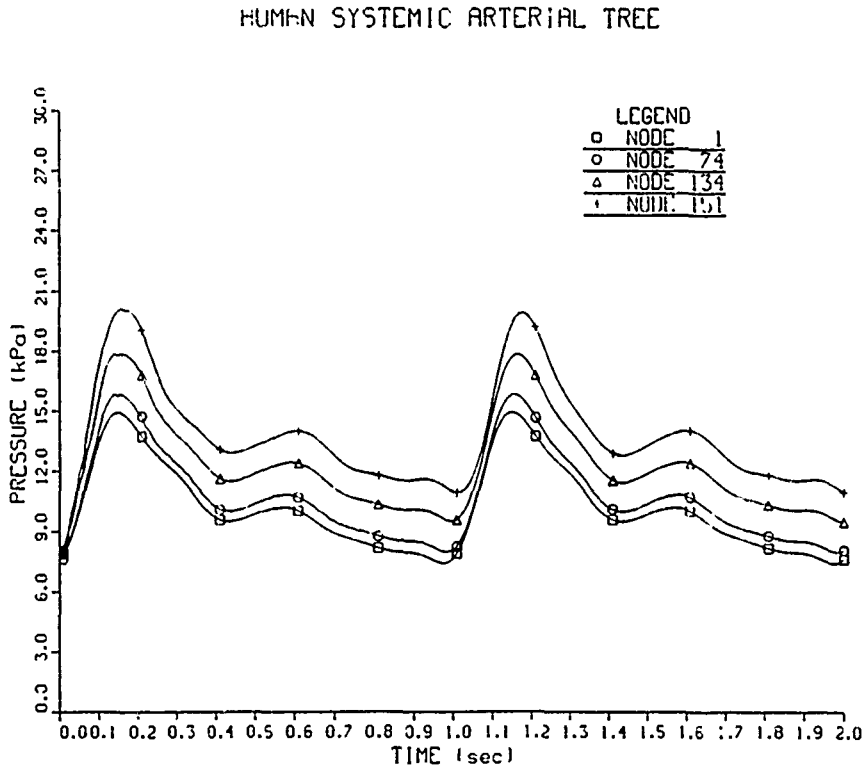


Figure 6

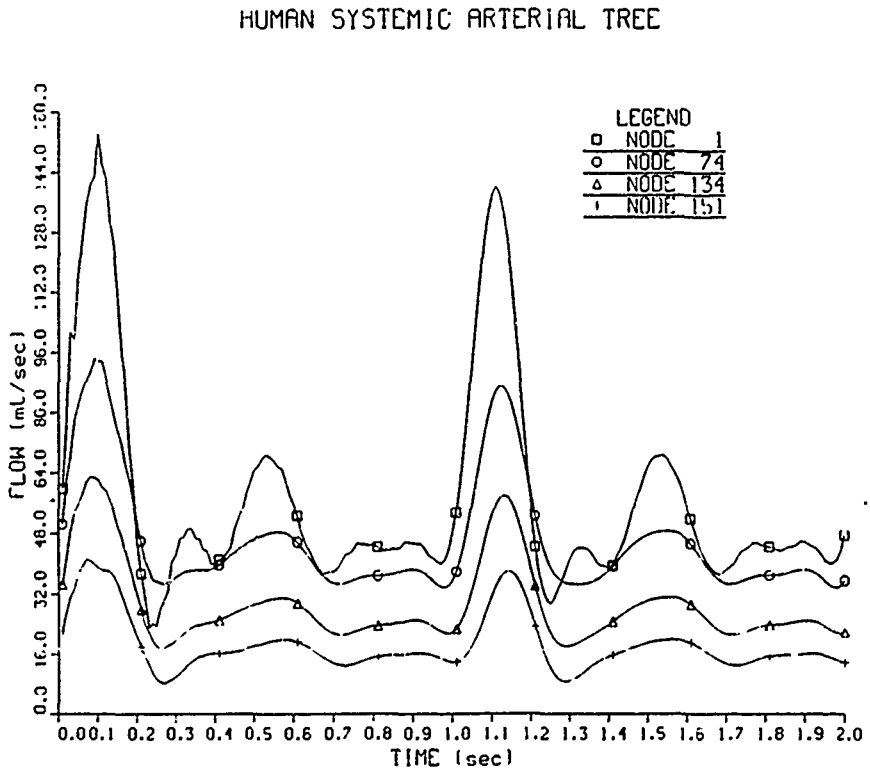


Figure 7

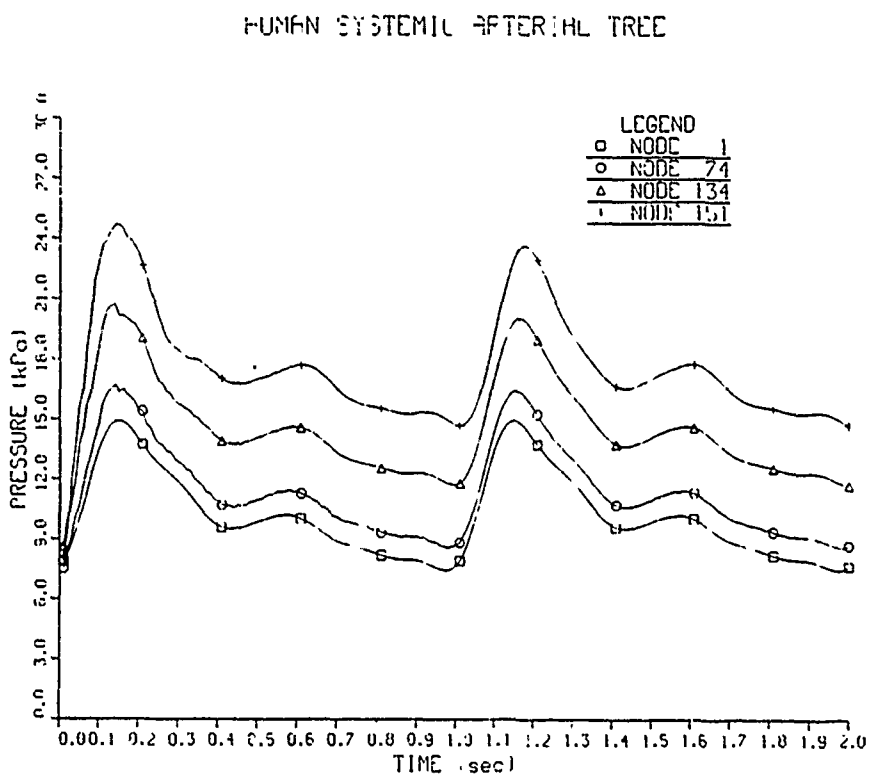


Figure 8

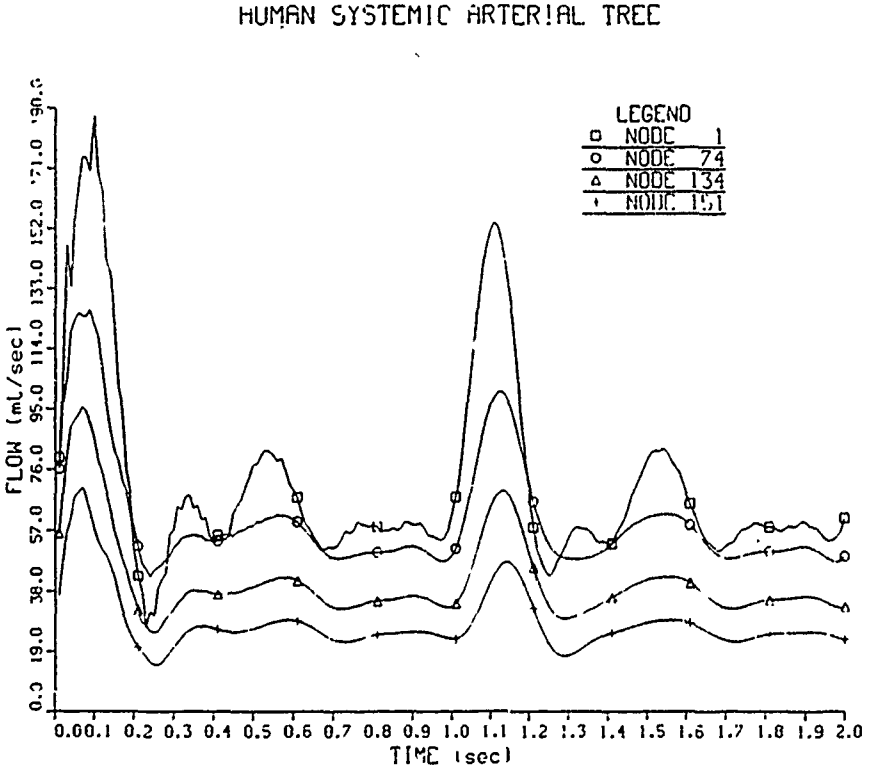


Figure 9

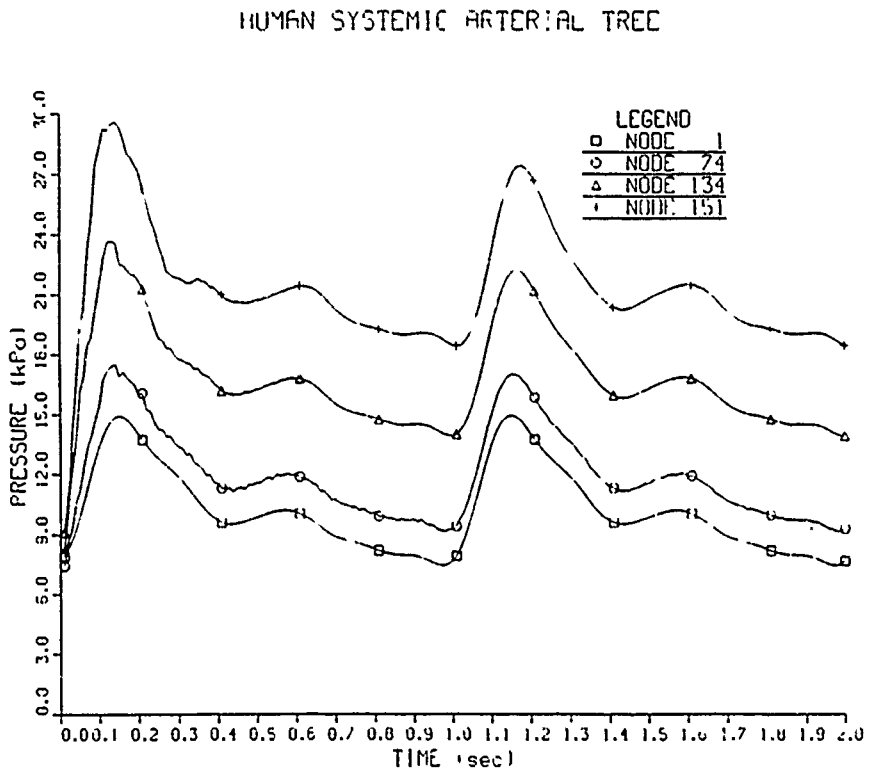


Figure 10

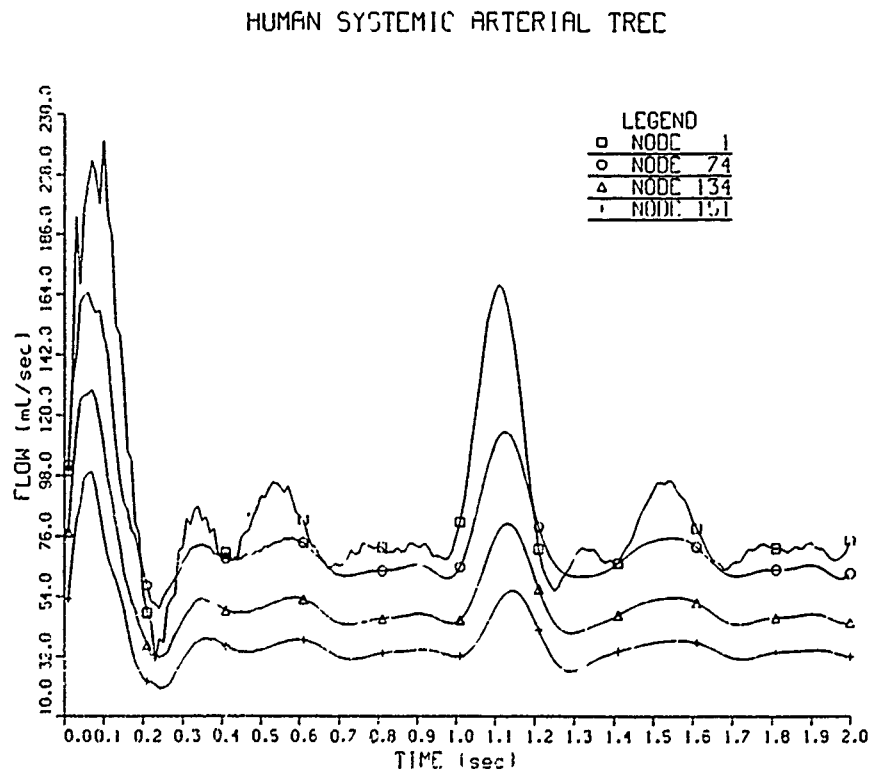


Figure 11

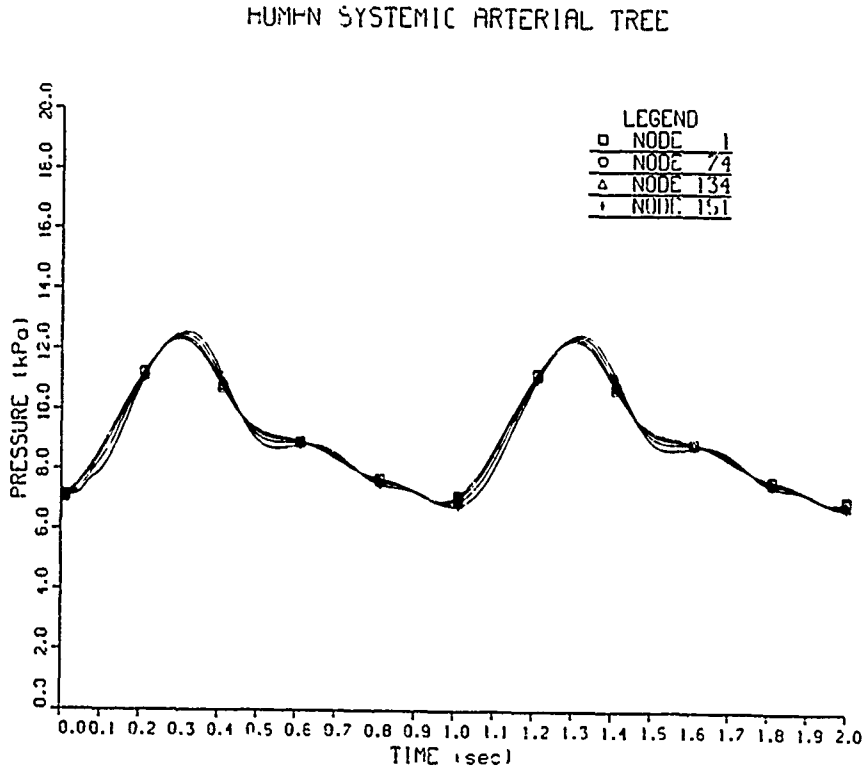


Figure 12

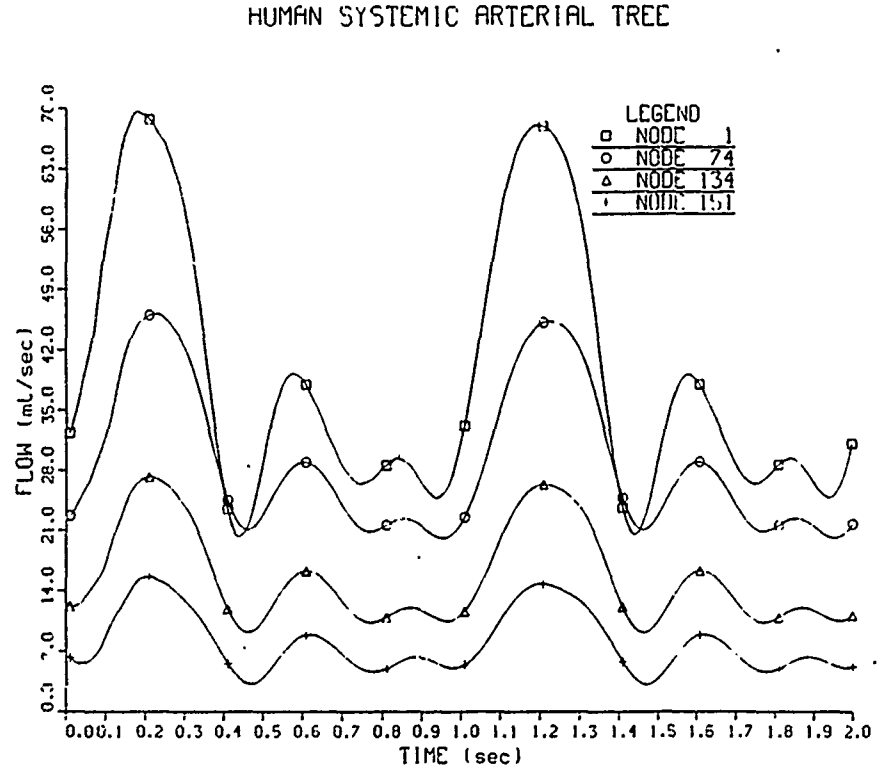


Figure 13

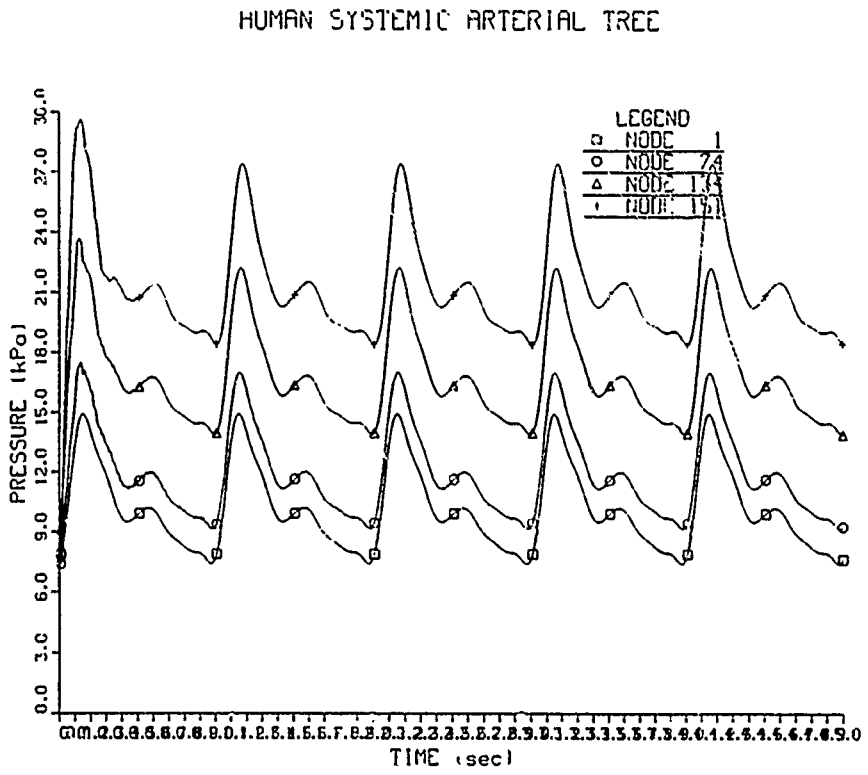


Figure 14

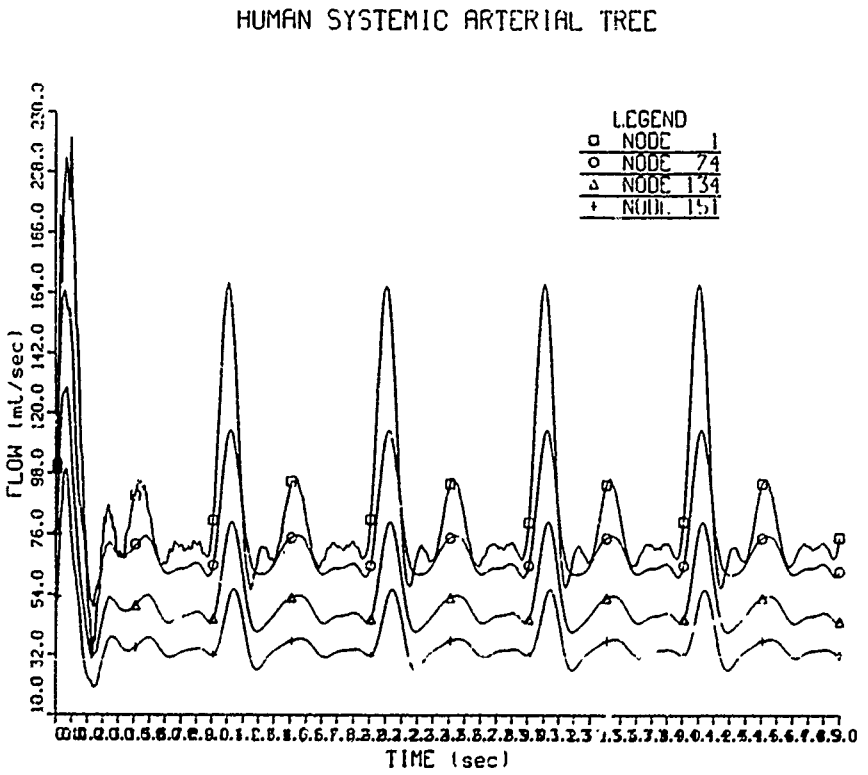


Figure 15

1989 UES - RESEARCH INITIATION PROGRAM

Sponsored By The

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted By The

UNIVERSAL ENERGY SYSTEMS

FINAL REPORT

Prepared:	Sonia D. Hart, M.S.
Academic Rank:	Teaching Associate
Department and	Education Department
University:	The University of Texas at San Antonio
Research Location:	Shaw AFB, Hill AFB, Eglin AFB, Langley AFB
Date:	November 1989
Contract No:	F49620-88-0053/SB588a-0378

The Effect Of Age, Family Status, And Physical Activity On Select Dietary Components Of TAC Pilots

Diet and athletic performance have been investigated a great deal over the past twenty years. The performance demands of elite athletes are similar to those of pilots, yet unlike athletes, there is limited information about specific dietary needs of pilots. This preliminary study gathered dietary data on TAC pilots for descriptive and analytical purposes. METHODS One hundred eighteen TAC pilots who fly F-15's and F-16's from five bases were interviewed concerning their diet, age, family configuration, and exercise level. Dietary information was recorded using PruCal three day dietary report form, while the remaining information was recorded on a personal interview form. RESULTS Three-way ANOVAS (Age X Family Configuration X Exercise) were used to determine if there were significant differences for pilots on their intake of protein, carbohydrates and fat. Although no differences between groups were found, few pilots met the recommended daily allowances for any of these three food groups. CONCLUSIONS Many athletes opitmize their diet in order to meet the strenuous demands of training and competition. These preliminary findings indicate pilots do not have diets well suited to the demands of their jobs. There appears to be a need for dietary education programs for TAC pilots and further study of the diets of TAC pilots in stress situations.

ACKNOWLEDGEMENT

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Universal Energy Systems and The University of Texas at San Antonio must be mentioned for their concern and help in all administrative and directional aspects of this program. I would like to thank TAC Headquarters for their support and assistance, especially Wing Commandor Peel at Langley AFB. Without the hospitality and professional expertise of the local flight surgeons involved (Col. Gillis at Langley AFB, Major Knauck at Shaw AFB, Major Kleinsmith at Hill AFB, and Major Depriest at Eglin AFB), the success of this project would not be possible. I thank the 118 TAC fighter pilots who gave their time and energy to assist me with this study, especially Lt. Col. Bates at MacDill AFB, who was an inspiration and a constant source of Air Force background information. Without the foundation provided by the School of Aerospace Medicine at Brooks AFB, this project would never have gotten off the ground, stayed on track, and been completed. Thanks to Dr. Craig Morrison of Oklahoma State University for his mastery of statistical and scientific information.

INTRODUCTION

Studies conducted since the 1940's have furnished information about the food consumed and the energy expenditures of the general population (1,7,8,15). Research has also provided dietary information about subgroups of the general population. One of these subgroups is athletes. Over the past 20 years, nutrition profiles of different athletes have been developed (4,6,9). The performance demands of flying sophisticated jets mimic the demands of elite athletes. Fighter pilots have been compared to athletes such as: weight lifters, body builders and sprinters (2,5). Comparisons to specific sports may or may not produce exact matches. However, they are useful in understanding some of the physical requirements of pilots. Other parallels between athletes and pilots are: emotional stress, hectic travel schedules, training pressures and competition (combat). In spite of the pilot/athlete analogy, dietary profiles have not been developed for pilots.

Athletes demonstrate little dietary knowledge (4,14). As no studies have been done with pilots, this may also be the case with pilots. Lack of nutritional knowledge and demanding lifestyles may lead to nutritional deficiencies. Studies have shown that irregular and insufficient caloric intake may cause hypoglycemia (3). Hypoglycemia can affect the performance of pilots by reducing G-tolerance and even provoking G-loss of consciousness (2). Stewart (12) showed that in some instances,

reduced blood sugar may lead to unconsciousness during positive acceleration. In 1956 and 1957, known cases of unconsciousness in flight were analyzed and hypoglycemia was found (11,12). It is possible that a poor diet (missed meals or hurried snacks) contributes to hypoglycemia. Dietary deficiencies can be significant factors in pilot performance and aircraft accidents.

Lyons and Wilson (9) indicate a correlation between dietary disturbances and peptic ulcers in the pilot population. As with hypoglycemia, incipient peptic ulcers can affect performance. Since there is a connection between diet and health, and good health is the foundation for performance, it is important to develop a nutrition profile of TAC fighter pilots. The purpose of this study was to examine the effect of age, family status and exercise on the diet of TAC pilots.

METHODS

Subjects

118 TAC pilots who fly F-15's and F-16's at TAC bases (Langley AFB, Shaw AFB, Eglin AFB, MacDill AFB and Hill AFB) were used for this study. For analysis, these pilots were divided into groups by age (20-30, 30-40, 40+), family configuration (single, couples no children, couples with children), and exercise level done on a weekly basis (no exercise, 1 or 2 days, 3 or more days).

Instrument

PruCal Analysis Computer Program was used to analyze a typical three day diet to determine the percent protein, carbohydrates and fats for each pilot. Information was also collected on dietary fiber, cholesterol and sodium. A personal interview form was used to gather descriptive information about the lifestyle of the pilots.

Procedure

A personal interview with the primary investigator followed the three day dietary recording period to clarify responses on the PruCal analysis. The investigator also asked the pilots questions concerning family configuration, physical activity done on a weekly basis and age.

RESULTS AND DISCUSSION

Three-way ANOVAS (Age X Family Configuration X Exercise) were used to determine any differences ($P<.05$) for the pilots on their intake of protein, carbohydrate and fat. Although no differences were found among the pilots, few met the recommended daily allowances for any of these food groups.

Table 1

Means and standard deviations for percent of dietary protein, carbohydrate and fat

	X	S.D.	Minimum	Maximum	RDA
Protein	15.05	2.66	8	21	12-15
Carbohydrate	44.56	8.27	25	77	48-57
Fat	34.25	6.31	15	47	30

The mean protein percentage for the 118 pilots was 15%. The pilots were within the recommended range of 12% to 15% set by the U.S. Guidelines. When compared to athletes, football players had 15% protein: body builders 19%: gymnasts 15%: and sprinters 14% (14). Twenty elite triathletes studied using a 7 day dietary intake recording reported a 13% protein intake (4). Although the average RDA for the pilots fell within the normal range, many of the pilots were outside the norms. Certainly the extremes of 8% and 21%, indicate dietary counseling is advisable.

The mean carbohydrate percentage for the pilots was 44.5%. The RDA for carbohydrates of 48%-57% is well above that recorded for the pilots in this study. Current TAC guidelines for pilots of 50% to 60% carbohydrate also exceed recorded carbohydrates. Comparisons with athletes yield the following percentages: football players 30%: body builders 36%: sprinters 49%: and gymnasts 44% (14). Elite triathletes had a carbodydrate consumption of 59.9% (47). The U.S. population has a current diet of 22% carbohydrate (7). The mean carbohydrate recorded for pilots did not fall within the range suggested by either the Air Force or U.S. Guidelines. As

with protein, the extremes of 25% and 77%, indicate reason for concern. The TAC Pilots as a whole fell short in their carbohydrate analysis.

The recommended guideline for fats is 30%. The pilot population had 34% of their diet in fat. When compared to other athletes, body builders had 39% fat, sprinters 36% fat, and gymnasts 39% (14). Triathletes had 27% fat in their diet (4). The current diet of the U.S. population is composed of 42% fat (7). As can be seen, the fat component of TAC pilots' diet, is higher than the recommended daily allowance. Although it is lower than the normal population's, it is still too high, especially for personnel in performance oriented positions.

Although the focus of this study was in the three major food groups, other components of the pilots diets were measured: cholesterol, sodium and dietary fiber.

Table 2

RDA's, TAC Pilot's consumption and percentage of TAC Pilots' meeting guidelines

	RDA	TAC Pilots (Average)	% Meeting Guidelines
Sodium	2000 mg/day	4000mg/day	6%
Cholesterol	300 mg/day	500/mg/day	45%
Fiber	15kg/1000kg/day	7kg/1000kg/day	4%

As revealed by Table 2, most TAC pilots are not within the RDA's for these dietary components. This is surprising since some of these dietary components along with the various nutrient groups affect cardiac function. High blood pressure or arterial plaque may result from excess sodium or too much fat and cholesterol in the diet. Cardiac function is a major concern of any pilot let alone TAC pilots.

CONCLUSION

Diet is crucial to the maintenance and improvement of performance. Although no dietary differences among the pilots were found, few pilots met the RDA's for the major food groups. Not enough pilots met the RDA's for cholesterol, sodium and fiber. These two findings present reasons for concern. It is important to note that most (96%) diets were recorded under normal working conditions. The interesting question arises, how do diets change under "game" conditions? Do the stresses of training and combat simulation lead to better or worse diets? Athletes are accustomed to modifying diet before games. In some cases (endurance athletes), diet is modified for several days when carbohydrate loading is employed to enhance performance before a major event. Are there any provisions for, or do pilots prepare "pregame" meals designed to meet the demands of competition?

A large percentage of the pilots (above 75%) are aware of the importance of their diet and were interested in the results of their 3 day dietary

in take analysis. They seemed knowledgeable about cholesterol, simple sugars and the consumption of too much alcohol. As a group, the pilots would like to know more about their own diet and what type of diet would be optimal for the stresses of their particular assignments. The results of the analysis show not enough is known by pilots to design optimal diets.

Since pilots must perform in a very demanding environment. They cannot afford to neglect something as important as diet. Further study is needed to examine dietary intake in the very stressful environments of: multiple sorties, night flying, remote tours, overseas tours with family and cross country flights. Required physical conditioning programs are an important part of pilot training. Nutrition education should be incorporated in the pilot training program either as a part of the conditioning program or as a separate component. It would be of interest to see if nutrition education programs lead to improved diets both in the mission and regular duty environment.

A great deal of time and money is spent to elevate flying performance in the Air Force. Athletes modify their diet to enhance performance. Similarly, pilots should recognize the role diet plays in their performance. The results of this study show there is need to improve the diets of TAC pilots. The expressed interest of the pilots in this survey for improving their diet could lead to dietary changes which would provide an improved basis for flight performance.

1. American Dieticians Association, "Food Consumption Survey", Journal of American Dietetic Association, Vol. 70, 1977.

2. Baldin, U.I., "Physical Training and +Gz Tolerance." Aviation, Space, and Environmental Medicine, August 1983, pp. 1115-1135.

3. Browne, M.K., "The Effect of Insulin Hypoglycemia on Tolerance to Acceleration," FPRC 1044, May 1958.

4. Burke, M.S., "Triathlete", The Physician and Sportsmedicine, Vol. 15 February 1987.

5. DeHart, R.M. "The Aerospace Athlete", TAC Attack, March 1987.

6. Hecker, Arthur, M.D., "Nutritional Conditioning for Athletic Competition", Clinics in Sports Medicine, July 1984.

7. Hegsted, M. "Dietary Goals for the United States", U.S. Senate Select Committee on Nutrition and Human Needs, Washington D.C., Government Printing Office, December 1977.

8. Lamb, L.L., M.D., "How Many Calories Do Men Need", American Journal of Clinical Nutrition, Vol. 46, 1987.

9. Lyons, T.J., M.D., K. Wilson, M.D., "Retrospective Cohort Study of Peptic Ulcer Disensi in U.S. Air Force Pilots". Aviation Space and Environment, 1989.

10. Manjarrez, Clarrissa, R.D., and Richard Birrer, M.D., "Nutrition and Athletic Performance", American Journal of Sports Medicine, Vol. 28, November 1983.

11. Powell, T.J., Carey, M.M., Brent, H.P., & Taylor, "Episodes of Unconsciousness in Pilots During Flight in 1956", Journal of Aviation Medicine, Vol. 28, 1957.

12. Powell, T.J., "Episodic Unconsciousness of Pilots During Flight in 1956", Journal of Aviation Medicine, Vol. 27, 1956.

13. Rayman, R.B., "Inflight Loss of Consciousness", Aerospace Medicine, June 1973.

14. Short, S.H., "Four Year Study of University Athletes Dietary Intake", Journal of the American Dietetic Association, Vol. 82, June 1983.

15. Ten State Nutrition Survey 1968 to 1970, Publish #72-81334, U.S. Department of Health, Education and Welfare, Washington, D.C., Government Printing Office, 1977.

AFOSR RESEARCH INITIATION PROGRAM
FINAL REPORT

TITLE: COMPREHENSIVE LIPOPROTEIN ANALYSIS BY HIGH-PERFORMANCE MOLECULAR EXCLUSION CHROMATOGRAPHY

SUBMITTED BY: Dr. Wesley Tanaka (Signature) Wesley Tanaka
Department of Chemistry (Date) 5/10/89
University of Wisconsin - Eau Claire
Eau Claire, WI 54701

Present Address: Merck & Co., Inc.
P.O. Box 2000
Rahway, NJ 07065

University of Wisconsin - Eau Claire, WI 54701

Administration - Dr. Ron Satz
Dean, School of Graduate Studies
Schofield 201
(715) 836-2721

- Mr. Don Zeutschel
Coordinator Research Services
(715) 836-5733

ACKNOWLEDGMENT

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research and Universal Energy Systems for selecting me to participate in the program.

The research project was rewarding and productive because of the help of a number of people. I am grateful to Ms. Lynn Stanco for her excellent technical help. Her enthusiasm and industry in the face of major technical and logistical problems were much appreciated. To the faculty and staff of the Chemistry Department I owe much. Perhaps the legacy of expertise gained, equipment purchased and the continuation of a commitment to high-quality undergraduate research and education that this project helped to foster was partial payment on that debt. I am particularly grateful to Drs. Joel Klink and Jack Pladziewicz for their advice, support and friendship.

AFOSR Research Initiation Program
Final Report

SUMMARY

Candidate reagents for and preliminary operating characteristics of a high-performance molecular exclusion chromatography system with a cholesterol-specific detector were studied. Preliminary results with a new higher efficiency molecular exclusion column, PWXL series, were not found to increase separation in the molecular weight range of interest compared to previous columns used (PW and SW series). The kinetics of the appearance of chromophore in the cholesterol assays from Sigma Diagnostics and Boehringer Mannheim Diagnostics were studied for use in a cholesterol-specific detector system. Cholesterol reagent from Boehringer Mannheim Diagnostics were found suitable as a post column reagent. Hardware for the post column system is described using parts available from U.S. suppliers. Operating characteristics of this system are described in the discussion.

INTRODUCTION

The eventual goal of the present project is the development of accurate reliable predictors of early coronary artery disease to be used to aid in the screening of individuals requiring intervention or monitoring. There is still considerable controversy regarding the best predictor of coronary artery disease. The present work focused on the use of high performance molecular exclusion chromatography together with post column detection of cholesterol to carry out simultaneous quantitation of all lipoproteins. This would permit analysis of various risk indices as a means for evaluating coronary artery disease risk. Preliminary work done in the summer of 1987 showed the potential feasibility of carrying out lipoprotein separation using high performance molecular exclusion chromatography. The major draw back of this method was the need for a specific detection system for cholesterol.

MATERIALS AND METHODS

1. Cholesterol Reagent Systems

Two commercially available reagent assays for use in the cholesterol-specific detector were characterized. Sigma Diagnostics Cholesterol (Catalog No. 352) and Boehringer Mannheim Diagnostics (BMD) Cholesterol High Performance Kinetics (Catalog No. 725252). Both reagent systems utilized a cholesterol esterase and a cholesterol oxidase catalyzed coupled enzyme arrangement as shown for the Sigma Diagnostics Cholesterol assay system (below).

$$\text{Cholesterol Esters} + H_2O \xrightarrow[\text{Esterase}]{\text{Cholesterol}} \text{Cholesterol} + \text{Fatty Acids}$$

$$\text{Cholesterol} + O_2 \xrightarrow[\text{Oxidase}]{\text{Cholesterol}} \text{Cholest-4-en-3-one} + H_2O_2$$

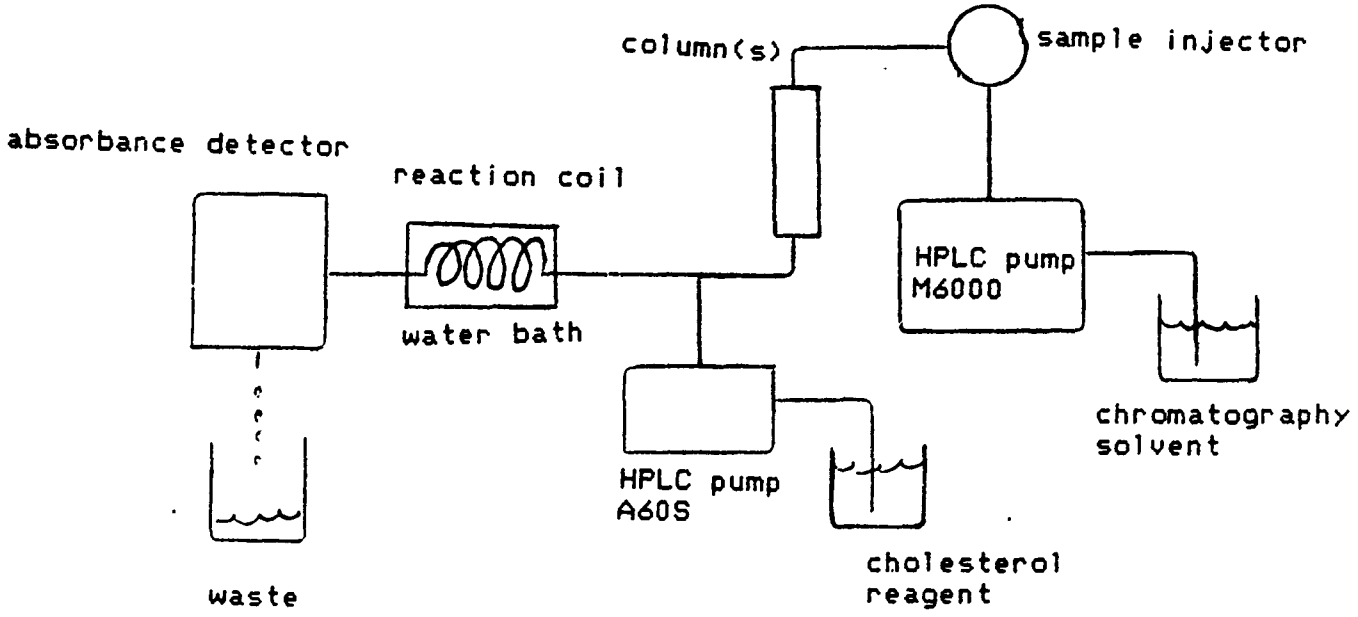
$$2H_2O_2 + \text{4-Aminoantipyrine} + \text{p-Hydroxybenzenesulfonate} \xrightarrow{\text{Peroxidase}} \text{Quinoneimine Dye} + 4H_2O$$

The hydrogen peroxide, H_2O_2, generated in the reaction is coupled to the production of a chromophore which is measured by absorbance at 520 (BMD cholesterol) nm or 500 nm (Sigma Diagnostics) after a 10 minute incubation at 37°C. Both reagent systems use a multipoint standard curve for calculating sample cholesterol concentration. The two assays utilize different enzyme concentrations, different dyes for development of color, and different buffer and detergents. When used according to manufacturer's specifications for the Instrumentation Laboratory (IL) Multistat, an automated analyzer on which the cholesterol

assay was characterized, the BMD cholesterol assay is almost 2.5 times more expensive on a cost per test basis (18.3 ¢/test) compared to the Sigma Diagnostics cholesterol assay (7.7 ¢/test). Linearity for the BMD cholesterol assay is reported as 800 mg/dL while the Sigma Diagnostic cholesterol assay has a reported linear range of 600 mg/dL.

2. HPLC System

The HPLC System consisted of a Waters Model M6000 Dual Piston Pump and an Eldex Model A-60-S Single Piston Pump (Biomedical Engineering Company, St. Paul, MN) hooked up as shown in the diagram below.



Schematic Diagram of HPLC with Post Column Detector

Connecting tubing (except for the reaction coil) is 0.020" ID stainless steel tubing. Fittings were from Chromtech, Apple Valley, MN. Sample injector is a Rheodyne Model 7125 equipped with a 2.0 mL sample loop. Reaction coil consisted of a 75 ft. link of 0.01" ID stainless steel tubing wrapped tightly around a 1" diameter galvanized pipe (total volume in reaction coil is 1.0 mL). Coil is immersed in a constant temperature water both maintained at the temperature indicated in the result section.

3. Detector and Data Collection System

Absorbance detector is an ISCO Model 1840 Variable Wavelength Detector. Analog signal from the detector was digitized via an ADA-LAB PC Interface Board and CHROMAMP-PC Interface Modual (Interactive Microware Inc., State College, PA). Data was collected and processed on a Zenith Model 159 Microcomputer (Zenith Data Systems) using CHROMATOCHART-PC software (Interactive Microware Inc.).

4. Reagents

Specialty reagents included the following: Sigma Diagnostics Cholesterol Assay (No. 352); BMD Cholesterol High Performance Kinetics (No. 725252); Sigma Diagnostics ACCUTROL Abnormal (No. A2034) and Normal (No. A3034) Control; NBS Standard Reference Material 1951-Cholesterol in Human Serum (Frozen); National Bureau of Standards, Office of Standard Materials, Very Low Density Lipoprotein (Sigma No. L-2264), Low Density Lipoprotein (Sigma No. L-2139) and High Density Lipoprotein (Sigma No. L-2014); Peroxidase (Sigma No. P-8415), Cholesterol Oxidase (Sigma No. C-8649); Cholesterol Calibrators (Sigma No. 0534); Serum-based Calibrators (BMD No. 781-827); Apoferritin (Sigma

No. 3660, MW 443,000); Thyroglobulin (Sigma No. T-9145, MW 669,000); High Performance Molecular Exclusion Columns (30 cm X 7.8 mm ID), G3000 PWXL, G4000 PWXL, 2500 PWXL, PWXL Guard Column (Supelco Inc., Bellefonte, PA). All other reagents were of the highest quality commercially available.

RESULTS

Day-to-day precision with the Sigma Diagnostics and the BMD cholesterol assays were similar, approximately 1% - 5% CVs over a range of cholesterol concentrations from 100 to 400 mg/dL. Using a variety of commercially available serum-based samples with known cholesterol concentrations the accuracy of the two assays were similar as shown in Table 1. Reagents were made up and used according to manufacturer's specifications for the IL Multistat. Calibrators used for determining sample cholesterol concentrations were commercially available aqueous-based cholesterol standards at concentrations of 100, 200, 300 and 400 mg/dL. The Sigma Diagnostics Cholesterol assay required 2 µl sample and 150 µl reagent and used absorbance at 500 nm after a 10 minute incubation at 37°C to quantitate cholesterol concentrations. The BMD Cholesterol assay required 2 µl sample and 200 µl reagent and used absorbance at 520 nm after a 10 minute incubation at 37°C to quantitate cholesterol concentration. Both assays were linear to the highest cholesterol concentration tested, 400 mg/dL.

Table 1. Comparison of Accuracy for Sigma Diagnostics and BMD Cholesterol Assays*

Sample	Actual Conc. mg/dL	Observed BMD Assay mg/dL	Observed Sigma** mg/dL
ACCUTROL-N	163	165	160
Nat'l Bureau Stds			
Low	210.4	214	214
Medium	242.0	237	233
High	282.0	282	274

* Aqueous-based standards used for both assays.

** Concentrations for Sigma assay were calculated manually from absorbaace data collected from Multistat cholesterol program.

The time course of color development for the BMD Cholesterol assay was much faster than the Sigma Diagnostics assay as shown in Figures 1 and 2. In the experiments summarized by these figures, cholesterol reagent was reconstituted as described by the manufacturers for the IL Multistat. Reagent was mixed with serum-based cholesterol samples in a 3:5 (reagent:sample) ratio. Absorbance at the appropriate wavelength was followed in the Multistat with time. The data show that the rate of absorbance increase at 520 nm with the BMD Cholesterol assay is much faster than the rate of absorbance increase at 500 nm with the Sigma Diagnostics assay. At one and two minutes absorbancies reached 76% to 95% and 95% to 100%, respectively, of their maximum plateau value in the BMD Cholesterol assay. With the Sigma Diagnostics Cholesterol assay the percentages were much lower and much more dependent on cholesterol concentration. Supplementing the Sigma Diagnostics Cholesterol reagent with additional cholesterol esterase and peroxidase increased the rate of absorbance increase but not to the rate achieved with the BMD Cholesterol reagent. Temperature effect on the rates of absorbance increase was not examined.

A series of studies was undertaken to assess the separation capabilities of the PWXL series columns using commercially available purified proteins (catalase, MW 232,000; apoferritin, MW 440,000; thyroglobulin, MW 669,000) and purified lipoproteins (very low density lipoprotein, VLDL; low density lipoprotein, LDL; high density lipoprotein, HDL). Each protein and lipoprotein was run alone and in various combinations on the 2500 PWXL, 3000 PWXL and 4000 PWXL columns. The 2500 PWXL was unable to separate any of the lipoproteins or any of the purified proteins at a flow rate of 1 ml/minute. No additional work was done with this column. The separations achieved on the 3000 PWXL and 4000 PWXL are shown in Tables 2 and 3. Samples were introduced in volumes from 20 to 80 µl by partially filling the 2 ml sample loop. Mobile phase, consisting of 0.2M Tris-acetate 0.2M NaCl, pH 7.0, was pumped at a flow rate of 0.5 ml/min. Absorbance of the eluant was monitored at 280 nm using an ISCO variable wavelength detector and data was captured and stored using an A to D converter (ADA-LAB-PC) and software (CHROMAMP-PC) resident on a microprocessor (Zenith Model 159).

Table 2. Elution Times for Solute from 3000 PWXL Column

Analyte	Peak Elution Time(s)	Width at Half Peak Ht(s)
LDL	1104	75
HDL	1266	75
Catalase	1305	90
Apoferritin	1185	90
Thyroglobulin	1105	90

Solvent flow rate = 0.5 ml/min

Buffer - 0.2 M Tris-acetate, 0.2 M NaCl, pH 7.0

Table 3. Elution times for solute from 4000 PWXL Column

Analyte	Peak Elution Time(s)	Width at Half Peak Ht(s)
Catalase	1620	90
Apoferritin	1550	90
Thyroglobulin	1410	90

Solvent flow rate = 0.5 ml/min

Buffer - 0.2 M Tris-acetate, 0.2 M NaCl, pH 7.0

As expected the 3000 PWXL column separated better in the "lower" molecular weight range (230,000 - 440,000) and the 4000 PWXL column separated better in the "higher" molecular weight range (440,000 - 670,000).

The 3000 PWXL column was placed in series with the 4000 PWXL (3000 PWXL first) and the resolution power of the combined column arrangement was studied. A 20 µl sample containing thyroglobulin, apoferritin and catalase was chromatographed on the 3000 PWXL/4000 PWXL column system. Figures 3a and 3b show the absorbance at 280 nm of the eluant at two different flow rates, 1.0 ml/min (Figure 3a) and 0.5 ml/min (Figure 3b). Identity of each peak was established by chromatographing each protein separately. Elution time for the catalase peak, the last protein eluted, was 14.5 minutes at a flow rate of 1 ml/min and 35.1 minutes at a flow rate of 0.5 ml/min. By comparison, baseline separation of these same three proteins can be achieved with a column system consisting of a 3000 PW and 5000 SW column in series. Total separation time was 68 minutes (final report AFOSR summer work, W.K.T.). Figures 4a and 4b show the absorbance at 280 nm of the eluant from a 60 µl sample containing three lipoproteins VLDL, LDL and HDL at flow rates of 1.0 ml/min

(Figure 4a) and 0.5 ml/min (Figure 4b). Identity of each peak was established by chromatographing each lipoprotein separately. Elution time for the HDL peak, the last lipoprotein eluted, was 14.4 minutes at a flow rate of 1 ml/min and 35.2 minutes at a flow rate of 0.5 ml/min. By comparison, complete separation of the three lipoproteins can be achieved with a column system consisting of a 3000 PW and 5000 SW. Baseline separation can be achieved with a 68 minute run time with acceptable separation achieved in a 30 minute run (final report, AFOSR summer work, W.K.T.).

To study the operating characteristics of the post column reaction, a series of solutions containing different concentrations of an aqueous-based cholesterol calibrator was injected into a mobile phase consisting of 0.2 M tris acetate, pumped at a flow rate of 0.5 ml/min. Downstream from the injector a "tee" connector was used to mix mobile phase with cholesterol reagent. The cholesterol reagent used was the Sigma Diagnostics Cholesterol assay reagent reconstituted as described for use with the TL Multistat (7.5 ml distilled water/cholesterol 50 bottle) and was pumped with a second pump at a flow rate of 0.5 ml/min. The mixed stream containing mobile phase, sample and cholesterol reagent was then sent through a 75 ft. reaction coil consisting of 0.01 inch ID stainless steel tubing immersed in a constant temperature water bath. Total inside volume of the reaction coil is 1 ml. Figure 5 shows the absorbance profile of a series of cholesterol solutions injected at various times. The Figure 5 insert shows the relationship between the absorbance peak and the cholesterol concentration.

Due to technical difficulties with the M6000 HPLC pump analysis of samples using the 3000 PWXL/4000 PWXL columns with post column cholesterol detection was not carried out.

SUMMARY AND CONCLUSION

The present column configuration consisting of a 3000 PWXL/4000 PWXL offers less power in separating the three types of lipoproteins, VLDL, LDL, HDL compared to a 5000 PW/3000 SW configuration. A 5000 PW/3000 SW column configuration can do an adequate separation of the three types of lipoproteins in approximately 30 minutes. The 3000 PWXL/4000 PWXL column configuration yields only partial resolution of the three lipoproteins in 35 minutes. The 2500 PWXL is of little use in separating any of the lipoprotein types. There is no indication that it would be useful in sbufractionation of HDC species.

The post column detection of cholesterol using a cholesterol-specific reaction is still an attractive option. There are differences between cholesterol assay reagents not only in cost but also performance characteristics relevant to a post column reaction system. The time course of appearance of chromophore for the Boehringer Mannheim Diagnostics Chloesterol High Performance Kinetics Assay is much faster than the time course for the Sigma Diagnostics Cholesterol Assay. In spite of this difference a suitable linear response of signal (absorbance) to cholesterol concentration can be generated using the Sigma Diagnostic Cholesterol assay reagent in a continuous flow cholesterol reaction at 45° C. This continuous flow system is capable of detecting cholesterol in a 0.1 ml sample containing 10 mg/dl injected directly into the mobile phase system. Whether this is adequate sensitivity for a 1.0 ml sample of plasma containing 200 mg/dL cholesterol, a typical sample size and cholesterol concentration, remains to be demonstrated. Clearly, additional modifications to the reaction conditions are possible to increase sensitivity and extend linear range for cholesterol.

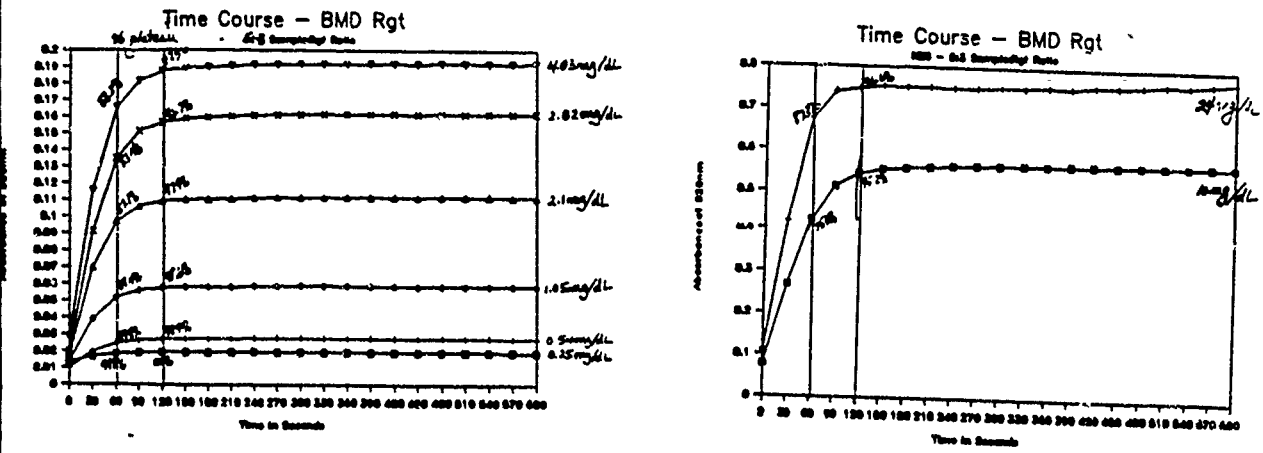


Figure 1. Time Course for the Appearance of Absorbance at 520nm for the BMD Cholesterol Assay.
Cholesterol reagent was reconstituted as described for use in the IL Multistat. NBS cholesterol standards were diluted to the concentrations shown on the graph. 125 ul of sample and 75 ul cholesterol reagent were loaded into the rotor and the absorbance of the solutions measured at 520 nm every 30 seconds for 10 minutes. All data were collected simultaneously by the Multistat. Pathlength of absorbance cell is 0.5 cm.

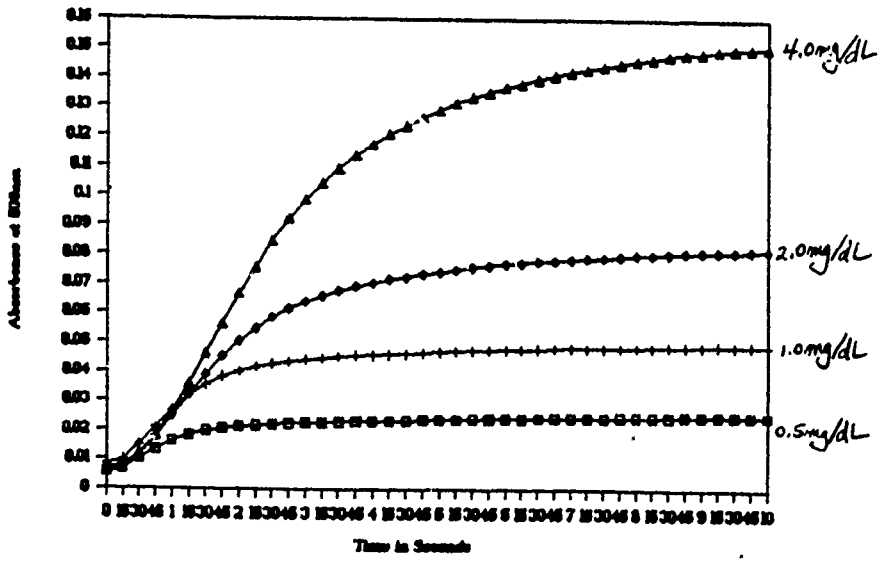


Figure 2. Time Course for the Appearance of Absorbance at 500nm for the Sigma Diagnostics Cholesterol Assay.
Cholesterol reagent was reconstituted as described for use in the IL Multistat. NBS cholesterol standards were diluted to the concentrations shown on the graph. 100 ul of sample and 60 ul cholesterol reagent were loaded into the rotor and the absorbance of the solutions measured at 500 nm every 15 seconds for 10 minutes. All data were collected simultaneously by the Multistat. Pathlength of absorbance cell is 0.5 cm.

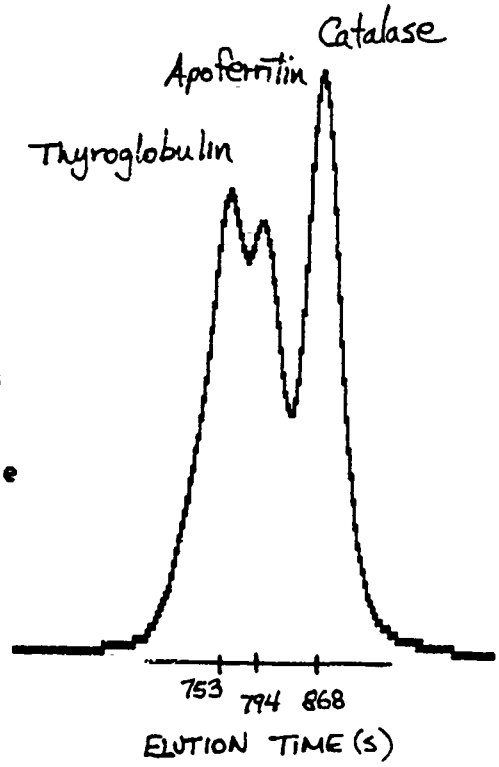


Figure 3a. Elution Profile of Protein Mixture Containing Thyroglobulin, Apoferritin, and Catalase.
A 20 ul sample of a protein mixture was injected onto a chromatographic column system containing a 3000 PWXL and a 4000 PWXL column connected in series. Mobile phase was 0.2M Tris-acetate, 0.2M NaCl, pH 7.0 at a flow rate of 1.0 ml/min. Absorbance of eluant was monitored at 280nm.

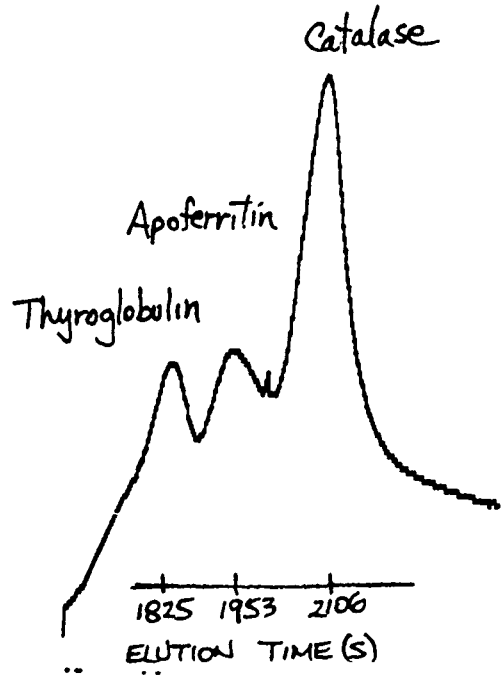


Figure 3b. Elution Profile of Protein Mixture Containing Thyroglobulin, Apoferritin, and Catalase.
A 20 ul sample of a protein mixture was injected onto a chromatographic column system containing a 3000 PWXL and a 4000 PWXL column connected in series. Mobile phase was 0.2M Tris-acetate, 0.2M NaCl, pH 7.0 at a flow rate of 0.5 ml/min. Absorbance of eluant was monitored at 280nm.

Figure 4a. Elution Profile of Solution Containing Purified Lipoproteins, VLDL, LDL, and HDL.
A 60 ul sample containing purified lipoproteins was injected onto a chromatographic column system containing a 3000 PWXL and a 4000 PWXL column connected in series. Mobile phase was 0.2M Tris-acetate, 0.2M NaCl, pH 7.0 at a flow rate of 1.0 ml/min. Absorbance of eluant was monitored at 280nm.

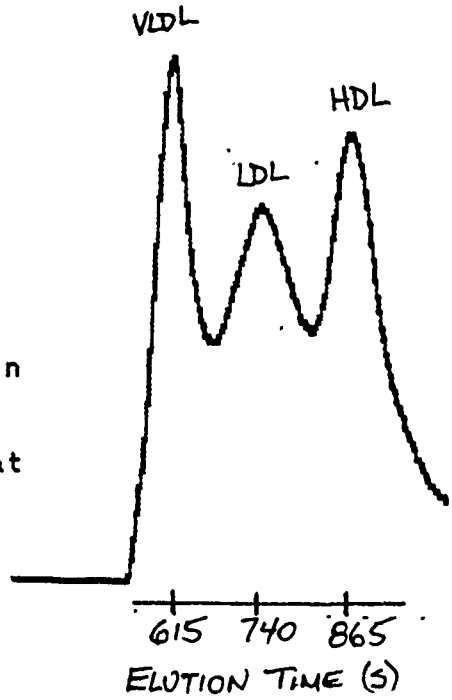
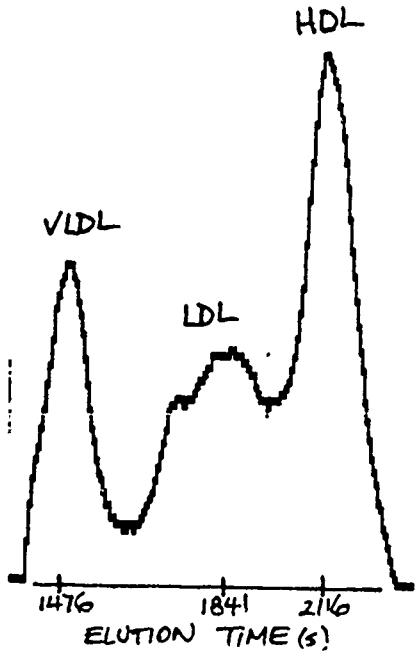


Figure 4b. Elution Profile of Solution Containing Purified Lipoproteins, VLDL, LDL, and HDL.
A 20 ul sample containing purified lipoproteins was injected onto a chromatographic column system containing a 3000 PWXL and a 4000 PWXL column connected in series. Mobile phase was 0.2M Tris-acetate, 0.2M NaCl, pH 7.0 at a flow rate of 0.5 ml/min. Absorbance of eluant was monitored at 280nm.



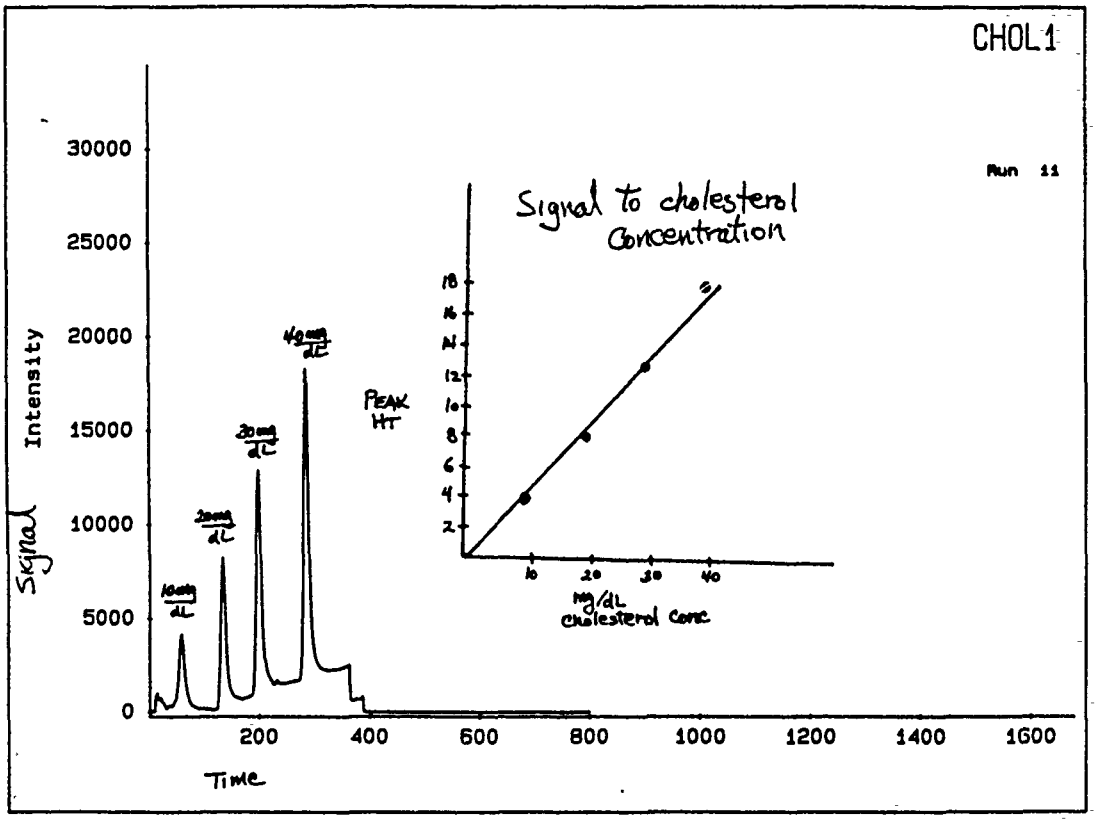


Figure 5. Response of Post Column Cholesterol Detection System.
0.1ml aliquots of cholesterol-containing solutions (10mg/dl, 20mg/dl, 30mg/dl, and 40mg/dl) were injected into a stream of mobile phase at approximately 2 minute intervals. Mobile phase carrying sample was mixed with cholesterol reagent as described in text in a equal volume ratio. The mixed stream was sent through a reaction coil immersed in a water bath at 47°C. Eluant from the reaction coil was monitored at 500 nm. Insert shows peak height (in units calculated by chromatography software) versus cholesterol concentration of injected sample

USAF-UES MINIGRANT

"Nmr and Temperature-Dependence Studies of the

Metal-Ion Catalyzed Chemiluminescence of Luminol"

sponsored by

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

and

UNIVERSAL ENERGY SYSTEMS

FINAL REPORT

Prepared by: John R. Wright, Ph.D.

Academic Rank: Professor of Chemistry

Department and Physical Sciences

University: Southeastern Oklahoma State University

Research Location: Southeastern Oklahoma State University

USAF Researcher: Dr. J. L. Kiel, USAFSAM/RZP

Date: December 29, 1989

Contract No: Proj.#S-210-9MG-037

Award#F49620-88-C-0053/

SB5881-0378

ABSTRACT

Chemiluminescent dosimeters based on luminol/hydrogen peroxide mixtures containing metal oxide catalysts have been shown to be sensitive to heating using 2.4 GHz microwave energy, and these might be used as a means for imaging the deposition of microwave energy in target phantoms. Several metal oxide catalysts were considered, and those of copper, cobalt and iron were found to be potentially the most useful for this type of application. A simple mathematical model was developed which simulates the general properties of the dosimeters. These dosimeters were also examined in strong acoustic fields at 20 kHz, and a dosimeter based on Scandium(III)oxide appeared to be more sensitive to an acoustic field than to heating.

ACKNOWLEDGEMENTS

We wish to express our gratitude to the United States Air Force and to Universal Energy Systems for the financial and administrative support provided for this research project. The P.I. also wishes to express his appreciation of all of the students participating in the work: Gregory Gandy, a physics/chemistry major; Tim Smith, a physics/chemistry major who will soon begin his Ph.D. program; Ann Mackie, a chemistry major; Steven VanWaggoner, a chemistry major; and Melanie Washburn, a high school student with much potential (Melanie did not receive salary support from the minigrant but was supported by a NIH minority program which allows academically strong high school students to participate in on-going research projects). Professor W. J. Polson (physics) offered many helpful suggestions. Finally, the P.I. also expresses his appreciation to Pat James, who helped in administrative matters and in the preparation of the final report.

1. INTRODUCTION

During a summer faculty research project conducted at Brooks Air Force Base in 1988 some attention was given to chemiluminescent dosimeters which are temperature sensitive and which might be used to generate three dimensional microwave absorption maps in target phantoms (1). It was shown that alkaline solutions of luminol in water containing hydrogen peroxide and a trace of copper(II) (as the hydrated oxide colloid) produced a chemiluminescent reaction over an extended period, and the reaction was temperature dependent, emitting more light at higher temperatures. These dosimeters are fairly dilute in all species (1) and thus have microwave absorption properties not markedly unlike those of water containing dilute electrolytes (e.g., biological fluids).

In most analytical applications of chemiluminescence, the reaction produces a brief burst of light, and one uses a photometric luminometer to measure the integrated envelope of this burst (2); however, these procedures stand in contrast to the needs of an imaging dosimeter, which must react slowly enough that one has time to fill and use a phantom. The latter objective is achieved by restricting the rate at which luminol is activated for its chemiluminescent reaction.

All available evidence indicates that the luminol reaction is initiated by free radicals (3). The dosimeter contains a small amount of the catalyst, and the action of the catalyst on hydrogen peroxide presumably produces the necessary activating free radical species (such as hydrogen atoms, hydroxyl radicals, etc.) at an acceptably slow rate. There is a lower limit to this since dilute aqueous

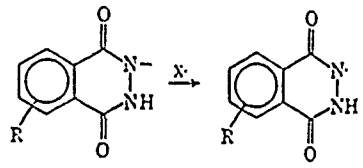
solutions of alkaline hydrogen peroxide containing luminol and no metal ions are weakly chemiluminescent and probably could generate images with sensitive charge-coupled device arrays. Also, this reaction is sensitive to dissolved carbon dioxide, which enhances the level of chemiluminescence, apparently through the formation of formyl or bicarbonate radical species (1,4) that act to increase the probability of the initiation reaction. Gas effects at liquid interfaces would be expected to complicate the interpretation of an imaging dosimeter; thus, dissolved gas effects are probably undesirable.

The work reported herein has been largely oriented to a search for metal ion catalyzed dosimeters which are, like the above-described copper(II)-catalyzed system, temperature sensitive. Varying degrees of temperature sensitivity were indeed observed, and appropriate dosimeters have been identified. Based on known solubilities, all of the catalysts described in the following are hydrated metal oxide colloids.

A possible interesting variant application of temperature sensitive dosimeters is in the detection of transient acoustic waves produced in targets by high power microwave (HPM) pulses (5). Sonochemical reactions are explained as the result of fluctuations of temperature produced in the compression and cavitation zones of an intense acoustic wave (6); thus one would expect that dosimeters which are sensitive to temperature would also show enhanced chemiluminescence during accoustic irradiations. Measurements conducted during this project did not bear this up, but time ran out before some of the observed phenomena could be thoroughly tested.

Most investigations of the luminol chemiluminescent reaction have

led to a conclusion that the mechanism in aqueous solutions begins by free radical action (3), i.e., a single electron is removed from the luminol anion at one of its azine nitrogens:



The best evidence for this comes from the fact that systems which are known to produce radicals also activate luminol, but the evidence for subsequent reaction species on the path to light production is much more nebulous. During this research effort an attempt was made to verify the involvement of radicals by means of nmr CIDNP spectra. CIDNP effects were not seen over a range of conditions, which may place some restraints on the time scale and nature of radical processes involved in the luminol reaction.

2. OBJECTIVES

The objectives of this project were outlined in the RIP, and they have been followed as closely as possible during the past year. The main objective of this research has been to identify metal ion catalyzed luminol dosimeters with the steepest possible temperature dependence, especially near 25 C, since these would be best suited for microwave/thermal dosimetry in target phantoms which are subjected to sudden RF fluences and with sufficient energy input to create a temperature step.

A second objective has been to examine some of these dosimeters for their sensitivity to acoustic fields. It was expected that sonochemical luminescence would correlate with thermal effects since acoustic waves act to cause instantaneous fluctuations in temperature; however, it was recognized that if a dosimeter behaved anomalously in this regard, it would have useful properties since pulsed microwave fluences create acoustic waves.

A third objective has been to detect and characterize, if possible, NMR CIDNP effects in the luminol reaction. It was hoped that this would gain insight into the chemiluminescent reaction mechanism, which is presumed to be initiated by free radicals. One needs to question the nature of metal ion catalyzed chemiluminescence: do all of these act to generate free radicals, or or do some of the metals act directly upon luminol?

3.EXPERIMENTAL METHODS

3.1. Measurements of Chemiluminescence as a Function of Temperature

A Turner model TD-20e luminometer interfaced to an IBM PS/2 (model 30/286) was used to quantify relative chemiluminescence intensities as a function of time or temperature. The rudimentary interface software provided by Turner, Inc. was developed into a data logging program (in BASIC) appropriate for this work, and a listing is included as Appendix II.

The luminometer was connected to a ten gallon circulating bath containing an antifreeze mixture. The bath was equipped with stirring

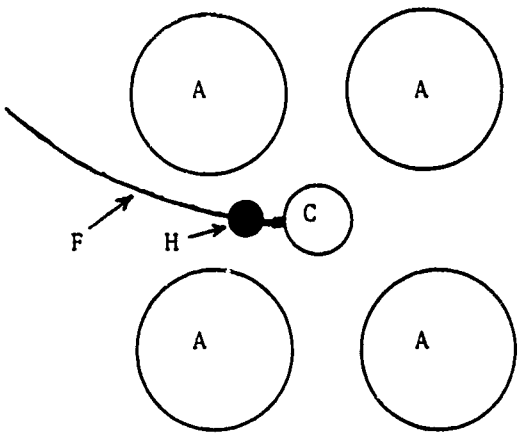
and circulating pumps, and a heating element/YSI Model 71 Thermistemp relay control loop kept temperatures in the bath and the luminometer head to within one degree. Further, the bath could be chilled using a Polyscience model KR-80A refrigeration unit, which provided an extension of temperature control to sub-ambient values.

Luminometric measurements were conducted in 8 mm translucent plastic cuvettes (from Turner) with the parabolic reflector in place. In all measurements a fresh cuvette was left in the luminometric compartment long enough to reach thermal equilibrium (about 10 minutes) while pipet tips and components of the luminometer solution were similarly pre-equilibrated in tubes kept in the circulator bath (it was fitted with test tube racks). The mixing process followed a rigorous timetable in an attempt to minimize systematic errors. The actual variability of data experienced in these measurements is attributed to geometrical imperfections in the luminometer cuvettes and other factors such as procedures for preparing catalysts.

3.2. Microwave Heating Experiments

For quantitative measurements of relative chemiluminescence levels in a microwave (heating) field, our machine shop fabricated a special cover for the Turner luminometer. This was fitted with a flexible fiber optic guide, five feet long, which was routed to the heating compartment of a 2.4 GHz, 500 watt microwave oven. The oven end of the nonmetallic fiber optic guide was anchored to a clear glass measuring cell so that the geometry could be kept constant. The floor of the oven was marked with a positioning pattern (Figure 1) so that additional containers of water could be added in a four-fold pattern

FIGURE 1. Lay Out of the Absorbers and Cuvette on the Microwave Oven Floor.



A - Energy absorbers; each beaker contained 200 mL of water.

C - The dosimeter cuvette.

F - Fiber optic guide.

H - Fiber optic guide holder.

around the measurement cell, allowing control of the heating rate. After some exploration of the conditions, we settled on having a total of 800 mL of water in the reservoirs, divided into four 200 mL portions, which began at 23 C, and 100 ml of the dosimeter in the cuvette, starting at a temperature of 5 C. The dosimeter was loaded into the measurement cuvette two minutes before beginning an irradiation. The dosimeters were prepared from stock solutions which had been pre-equilibrated in a 5 C constant temperature bath, and thermometry detected an approximate 2 C temperature rise during the arbitrary 2 minute delay time before starting the microwave irradiation. The microwave heating rate, B, was measured at 10.0 degrees K/minute (standard deviation 0.2) using an Omega digital thermometer (temperature changes after specified heating times were measured with the probe inserted in the solution volume near the fiber optic guide).

Measurements of doses with an irradiator of this type are beset with many difficulties (7), especially if one is attempting to measure accurate specific absorption rates (SARs), i.e., the dose distribution is not uniform, which may introduce convection, and temperature measurements are at best averages. However, in this case we are maintaining a fixed geometry, and all dosimeters were at the same ionic strength. The time versus luminosity data are equivalent to averaged temperature versus luminosity. This method is especially appropriate for observing changes in relative chemiluminescent intensity expressed as a ratio, I/Io, where Io is the intensity at the beginning of an irradiation and I the intensity at some later time. Data obtained by the method proved to be reproducible, and it was possible to compare the temperature sensitivity of different

dosimeters, which is the real intent of the work.

3.3. Sonochemical Luminescence

The sonochemical irradiator made use of the fiber optic guide described above for microwave irradiations. A blackened light baffle was fabricated for inclosing the transducer element of a Branson model W-185 sonifier, and a 50 mL beaker containing 20 mL of the test dosimeter solution was positioned so that the tip of the transducer dipped half-way to the bottom of the dosimeter solution. The fiber optic terminus was positioned to sample light coming from the region of the dosimeter solution within 1 mm of the sonifier's tip. The Branson sonifier can deliver up to 150 watts of acoustic energy at a frequency of 20 kHz; however, in these irradiations the power was set at 50 Watts to avoid violent swirling action. The dark-adapted human eye can detect sonochemical luminescence in these dosimeters. It appears as a small, flame-like zone immediately below the transducer, and intensity changes appear to be instantaneous when one switches the sonifier on or off.

3.4. Nmr and Other Spectroscopic Measurements

Proton and carbon-13 nmr spectra were obtained using a Chemagnetics A-200 spectrometer. In the attempts to detect CIDNP effects in the luminol chemiluminescence reaction, controls were obtained to verify that these reacting samples were actually producing light during the times of measurement.

Infrared spectra were obtained by means of a Bomem MB-100 FTIR

instrument.

3.5. Dosimeter Preparation

Dosimeters were prepared from two stock solutions: alkaline luminol was prepared fresh daily and consisted of 20.0 mg of luminol (Sigma) in 6.00 mL of 1 N NaOH; 2% solutions of the various transition metal salts were prepared by dissolving 1 g in 50 mL deionized water. Titanium(IV)chloride is a liquid (toxic!), and some of the stock metal solutions were acidified to maintain solubility. The very active dosimeters of cobalt and ruthenium were modified in some experiments; the metallic salts were normally at a 1:200 dilution in solution B, and a specified 1:2000 dilution means that the dosimeter contained the metal at a concentration one-tenth that of a normal dosimeter.

In measurements of dosimeter intensity versus temperature, a 0.600 mL aliquot of the alkaline luminol solution was made up to 10.0 mL in deionized water to obtain solution A, and a 50 microliter aliquot of the transition metal stock solution was made up to 10.0 mL with 3% aqueous hydrogen peroxide to obtain solution B. The final dosimeter solution was prepared by mixing A and B in the volume ratio 1:1, and to stay on-scale, only 25 microliters of the dosimeter was transferred to the luminometer cuvette.

In sonochemical irradiations the total dosimeter volume was 20 mL, i.e., solutions A and B were mixed, and the entire amount was transferred to the sonifier cuvette.

In microwave irradiations, the dosimeter component concentrations were as described above, but the dosimeters were scaled up to obtain a

final volume of 100 mL, all of which was transferred to the measuring cell.

Luminol and the metal oxide colloids are the limiting components of these reaction mixtures. The final dosimeter solutions contain hydrogen peroxide at 0.44 M, NaOH at 0.030 N and luminol at 0.00057 M (present as the anionic form in the alkaline solution). The various metal ions were at an equivalent concentration on the order of 0.0001 M; however, considering the colloid nature of the latter, it is not possible to specify catalyst concentrations.

There was not enough time to explore all of the possible metal oxide catalysts, and colloidal palladium metal (one of our candidates) should receive consideration along with more ions from the second and third transitions.

4. RESULTS AND DISCUSSION

4.1. Dosimeter Chemiluminescence as a Function of Temperature

The results of these measurements were not at all satisfying. An intensity versus temperature plot for the Co(III) dosimeter (Figure 2) shows the typically poor degree of reproducibility associated with the method. Figure 2 is based on three replications of the experiment, and the error bars mark one standard deviation. Excessive data scatter was a characteristic of all of these measurements, although it appears to be worse for some of the catalyst systems.

Averaged dosimeter intensity data as a function of temperature are collected in Table I, and Table II presents a least squares analysis of the contents of Table I to obtain activation energies (Ea) and pre-exponential (Ao) factors (from log I versus 1/T; this is an Arrhenius treatment which presumes that I is proportional to a reaction rate constant, k). The data of Table II reflect sampling at 1, 10 and 20 minutes into the isothermal chemiluminescence versus time curve; also, since the temperature versus intensity data appear to be convex in some cases and concave in others, the analysis for Ea and log Ao at the specified time intervals is shown for the lower range, 5 - 25 C, as well as the full range, 5 - 40 C. Data showing large deviations for both Ea and Ao are clearly suspect, and this was confirmed during subsequent microwave heating experiments (see 4.2.). The dosimeters with more consistent values of Ea and Ao were in fact temperature dependent in the latter experiments; however, the Ea values in Table II do not correlate well with the values of I/Io versus temperature seen in microwave heating experiments.

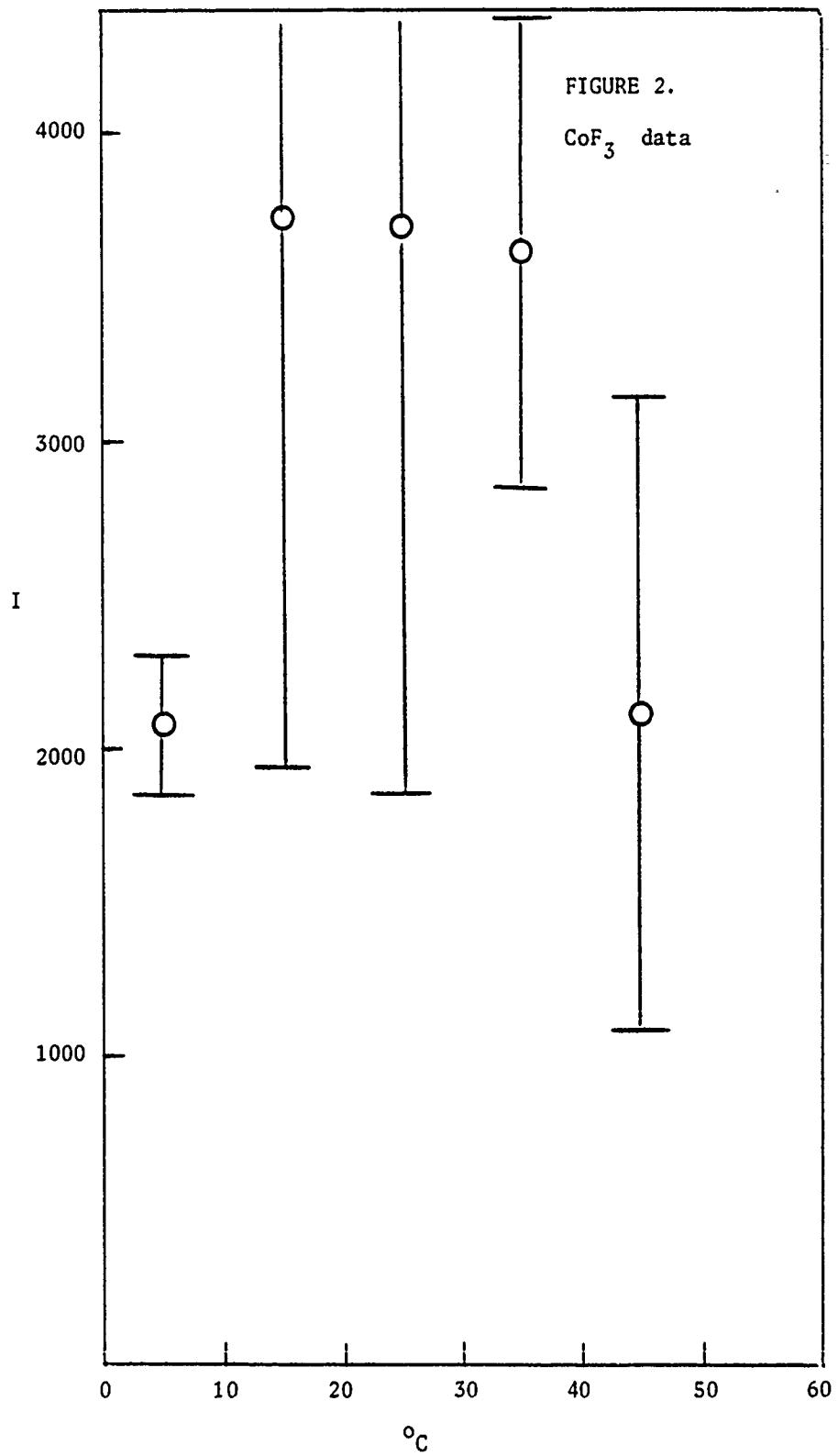


FIGURE 2.

CoF_3 data

TABLE I
Dosimeter Luminosity Versus Temperature

Source of Metal	Time, Min	t °C 5	15	25	35	40
$CuCl_2$	1	2108	3887	4564	5773	5126
	10	2585	3708	4376	4588	3117
	20	1977	2237	1695	2483	1324
$CrCl_2$	1	17.03	9.710	27.14	135.1	60.19
	10	13.10	10.42	15.48	64.22	18.08
	20	7.063	7.216	4.388	25.83	8.910
$Fe_2(SO_4)_3$	1	181.5	896.0	552.3	714.5	659.7
	10	117.5	384.9	564.0	616.1	559.6
	20	120.5	173.2	641.9	712.1	654.0
$CoSO_4$	1	5207	4931	4948	-	5149
(1:200)	10	6491	7278	6395	-	7101
$MnCl_2$	1	99.40	13.87	40.58	61.59	130.9
	10	40.64	19.76	25.03	36.77	47.43
	20	38.86	44.44	36.83	51.83	55.74
$FeSO_4$	1	271.3	178.2	335.6	337.6	366.6
	10	168.0	173.6	346.8	313.8	435.1
	20	211.2	193.9	328.6	374.6	289.1
$NiCl_2$	1	7.760	-	148.7	406.9	144.7
	10	4.361	106.8	122.9	219.1	94.10
	20	3.284	141.6	101.3	153.1	70.08
$ZnBr_2$	1	28.42	33.19	53.76	16.62	22.12
	10	26.77	7.810	37.90	7.491	8.340
	20	23.06	3.308	28.46	3.915	4.469
$RuCl_3$	1	335.0	1900	2924	274.2	467.8
(1:2000)	10	169.7	1371	2083	212.4	380.5
	20	95.10	1429	1520	201.7	372.6
$PbCl_2$	1	4265	6103	7404	4117	6821
	10	1729	3140	6086	3381	6200
	20	1090	3236	4312	2647	3135
$ScCl_3$	1	9.120	-	41.52	15.20	72.10
	10	2.712	-	31.77	11.50	39.11
	20	1.347	-	19.96	8.02	29.30

$NbCl_5$	1	276.7	505.8	413.2	511.6	156.5
	10	136.9	532.7	371.7	169.1	78.55
	20	63.35	387.3	293.9	124.4	58.29
AgF	1	154.6	420.9	37.34	237.4	330.4
	10	96.60	191.6	31.85	132.1	162.7
	20	45.14	108.5	35.60	90.82	95.02
$TiCl_4$	1	74.90	129.7	173.6	50.50	207.2
	10	11.01	191.5	224.2	11.30	71.63
	20	1.974	142.4	165.3	5.04	34.52
CoF_3	1	2078	3726	3697	3617	2115
	10	1925	2228	2281	3512	2962
	20	1709	1647	2229	3320	3525
$AgClO_4$	1	31.60	80.00	232.1	342.0	181.8
	10	9.070	31.60	136.7	168.8	61.50
	20	3.050	14.20	84.20	76.30	27.30
$ZrCl_4$	1	12.90	34.70	83.89	57.72	5.683
	10	3.800	20.80	112.0	56.30	1.457
	20	1.000	9.000	82.00	26.40	0.750
VCl_3	1	95.10	174.8	407.8	296.7	286.3
	10	68.80	68.90	281.4	196.3	126.8
	20	41.10	25.00	173.1	119.3	65.50

TABLE II
Dosimeter Activation Energies and Pre-Exponential Factors

Source of Metal	Upper Temp	Sample Time	Ea	S.D.	$\log_{10}A_o$	S.D.
$CuCl_2$	25	1	6412	1170	8.39	0.89
	25	10	4359	497	6.85	0.38
	25	20	-1232	1120	2.36	0.85
	40	1	4405	829	6.85	0.61
	40	10	1466	1115	4.64	0.82
	40	20	-1061	4	2.50	0.92
$CrCl_2$	25	1	3691	4412	4.02	3.35
	25	10	1317	1733	2.11	1.32
	25	20	-3876	1471	-2.16	1.12
	40	1	10152	3323	8.99	2.45
	40	10	5236	3061	5.14	2.26
	40	20	3593	3179	3.59	2.34
$Fe_2(SO_4)_3$	25	1	9375	5611	9.77	4.26
	25	10	13012	2062	12.35	1.57
	25	20	13690	2767	12.77	2.10
	40	1	4890	2583	6.33	1.91
	40	10	7341	1804	8.00	1.33
	40	20	9576	1675	9.62	1.24
$CoSO_4$	25	1	- 427	184	3.38	0.14
(1:200)	25	10	- 100	679	3.75	0.52
	40	1	- 28	160	3.68	0.12
	40	10	236	355	4.01	0.26
$MnCl_2$	25	1	-7678	8275	-4.24	6.28
	25	10	-4088	2585	-1.67	1.96
	25	20	- 412	891	1.29	0.68
	40	1	3085	4565	4.00	3.37
	40	10	1430	1850	2.56	1.37
	40	20	1595	653	2.83	0.48
$FeSO_4$	25	1	1653	2915	3.66	2.21
	25	10	5914	1881	6.83	1.43
	25	20	3587	1728	5.10	1.31
	40	1	2333	1187	4.18	0.88
	40	10	4760	916	5.94	0.68
	40	20	2667	962	4.40	0.71
$NiCl_2$	25	1	-	-	-	-
	25	10	27822	8083	22.7	6.14
	25	20	28667	10935	23.3	8.30
	40	1	17284	3980	14.60	2.91
	40	10	14486	5304	12.49	3.91
	40	20	13274	6354	11.52	4.69

$ZnBr_2$	25	1	5225	961	5.54	0.73
	25	10	2599	7758	3.27	5.89
	25	20	1345	11268	2.13	8.56
	40	1	- 2150	2191	-0.13	1.62
	40	10	- 4446	3703	-2.14	2.73
	40	20	- 5923	4991	-3.45	3.68
$RuCl_3$	25	1	17990	3378	16.75	2.57
(1:2000)	25	10	20837	4353	18.71	3.31
	25	20	23107	7012	20.31	5.32
	40	1	- 1827	5749	1.53	4.24
	40	10	187	6064	2.86	4.47
	40	20	2205	6505	4.26	4.80
$PbCl_2$	25	1	4564	415	7.23	0.31
	25	10	10371	315	11.38	0.24
	25	20	11417	2073	12.06	1.57
	40	1	987	1397	4.47	1.03
	40	10	5082	1761	7.31	1.30
	40	20	4005	2195	6.37	1.62
$ScCl_3$	25	1	-	-	-	-
	25	10	-	-	-	-
	25	20	-	-	-	-
	40	1	7459	3660	6.86	2.68
	40	10	11215	3826	9.35	2.80
	40	20	13211	3838	10.61	2.81
$NbCl_5$	25	1	3383	2176	5.16	1.65
	25	10	8400	4627	8.86	3.51
	25	20	12853	5586	12.04	4.24
	40	1	- 1409	2643	1.50	1.95
	40	10	- 3484	3881	-0.26	2.86
	40	20	- 1499	4637	1.04	3.42
AgF	25	1	-11390	9543	-6.51	7.25
	25	10	- 8913	6918	-4.84	5.25
	25	20	- 1767	5494	0.41	4.17
	40	1	1211	5154	3.15	3.80
	40	10	836	3845	2.64	2.84
	40	20	2402	2486	3.61	1.83
$TiCl_4$	25	1	6956	631	7.36	0.48
	25	10	25109	7127	20.95	5.41
	25	20	36906	10924	29.57	8.29
	40	1	1698	3098	3.30	2.28
	40	10	1993	7925	3.18	5.84
	40	20	5235	10469	5.24	7.72

CoF_3	25	1	4808	1574	7.14	1.19
	25	10	1411	335	4.40	0.25
	25	20	2158	963	4.91	0.73
	40	1	474	1656	3.82	1.22
	40	10	2607	596	5.32	0.44
	40	20	4081	662	6.38	0.49
$AgClO_4$	25	1	16430	567	14.40	0.43
	25	10	22350	855	18.50	0.65
	25	20	27337	981	21.94	0.75
	40	1	10259	2342	9.67	1.73
	40	10	11714	3732	10.35	2.75
	40	20	12985	4578	10.94	3.38
$ZrCl_4$	25	1	15449	123	13.25	0.09
	25	10	27901	280	22.49	0.21
	25	20	36337	455	28.54	0.35
	40	1	- 395	5987	1.13	4.42
	40	10	1136	9830	2.01	7.25
	40	20	4115	10893	3.86	8.03
VCl_3	25	1	11982	796	11.37	0.60
	25	10	11482	3996	10.76	3.03
	25	20	11672	6836	10.61	5.19
	40	1	5639	1766	6.51	1.30
	40	10	4908	2590	5.72	1.91
	40	20	7013	3516	6.78	2.59

The reason(s) for the observed data scatter are not completely clear. Pipetting technique was monitored, and this could not account for the observed large fluctuations since measured volumes were within 3%. Also, there was a protocol of fixed time intervals between each mixing step, and use of a vortex mixer insured good mixing. Temperature control was within 2 degrees in the worst cases (bath and luminometer compartment) and was more typically within 1 degree C. The Turner luminometer is demonstrably capable of good accuracy.

A further examination of the methods identified several probable sources of variability, and these are not readily controlled. These include the following:

- The 8 mm plastic cuvettes (available only from Turner) are translucent and quite irregular due to moulding imperfections. These can cause variations of I due to wobble or rotational position in the luminometer.

- The placement of the small pipetted aliquot (25 microliters) is difficult to control as one must work fast to minimize temperature changes. The test droplet may end up on the bottom of the cuvette or perhaps slightly to the side. This is a significant source of variability.

- The necessary movements between the temperature bath and the luminometer compartment may introduce temperature variations.

- Since all of these dosimeters produce metal oxide colloids, the particle size is a significant variable. Some of the dosimeters (e.g., Ti(IV)) produced a more turbid mixture. It is likely that fluctuations in the rate of mixing will alter the size of colloid particles and in turn the amount of catalytic surface area; this will

certainly cause variations in the absolute chemiluminescent intensity values and may explain why some of the dosimeters appear to be more reproducible than others.

These results led to a consideration of other methods which would allow better control of the luminometric geometry. Also, it became clear that measurements of relative increases of intensity (I/Io) were preferable to measurements of absolute intensity. A much better approach to this problem is discussed in the following section.

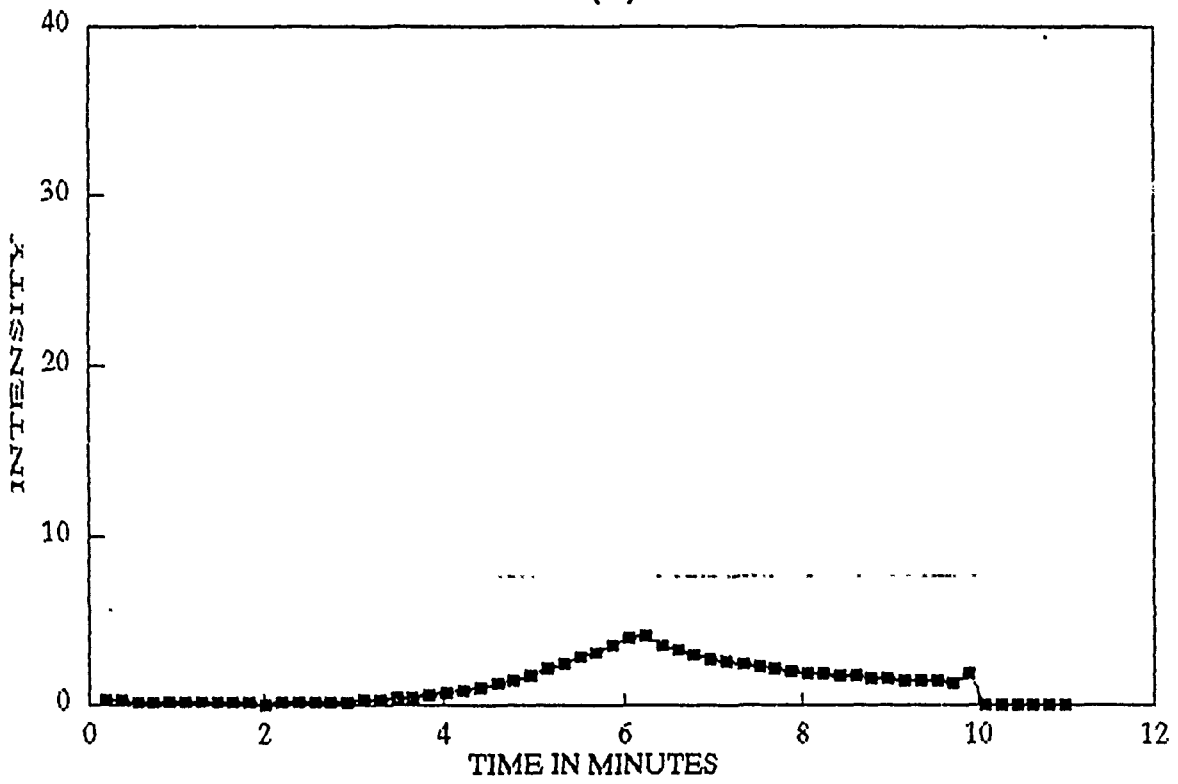
4.2. Microwave Heating Experiments

In the microwave heating experiments, the dosimeter solutions at 5 C were placed in the oven compartment (along with the power dissipating volumes of water) and allowed to stand for exactly 2 minutes before starting the irradiations. After an irradiation period of 4 minutes, the power to the magnetron was shut off. Data collection with the Turner/IBM PS/2 system was started immediately after placing the cuvette in its position (a 150 mL Berzelius beaker; always the same beaker) and closing the light baffle, a delay of about 10 s. Data were thus collected before, during and after the irradiation. Figures 3 through 14 show replicate plots of intensity versus time for the various dosimeters, and it is seen that this method yields reproducible results in spite of the limitations involved in large volume irradiations (7). Appendix I contains the raw data on which these plots are based. It should be noted that the appendix contains three replications of each dosimeter; only two are shown in the figures.

In some of the timecurves the isothermal dosimeter decay rate can

FIGURE 3

COPPER (II) SULFATE



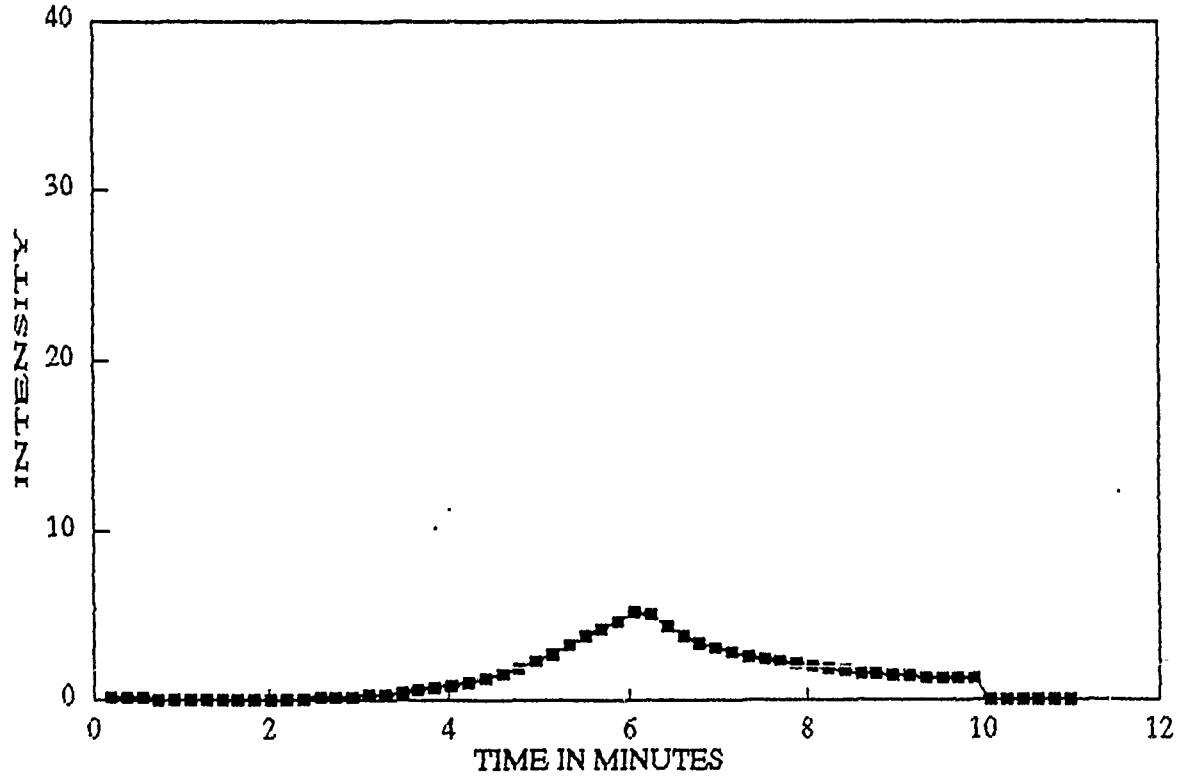


FIGURE 4

COBALT (III) FLUORIDE

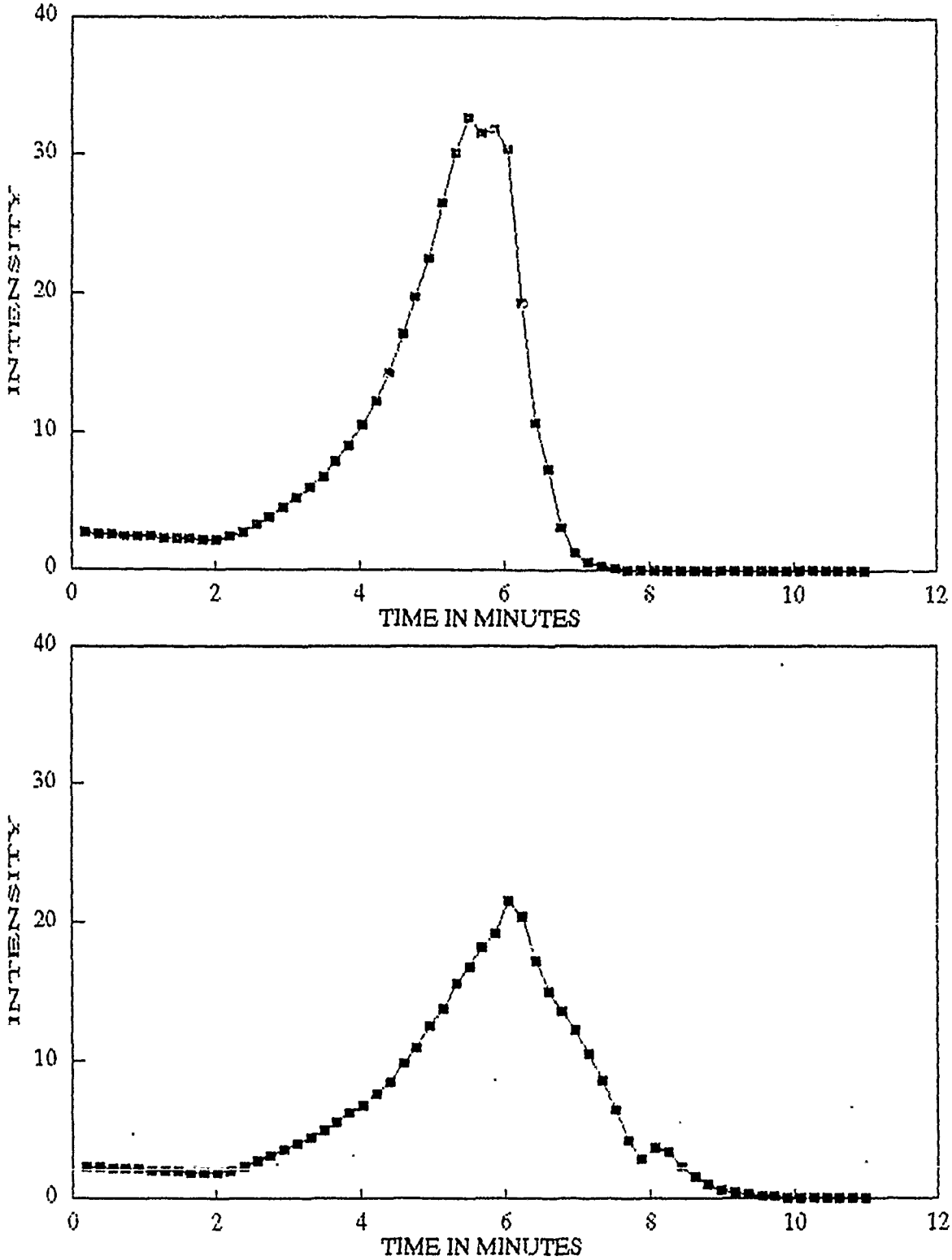
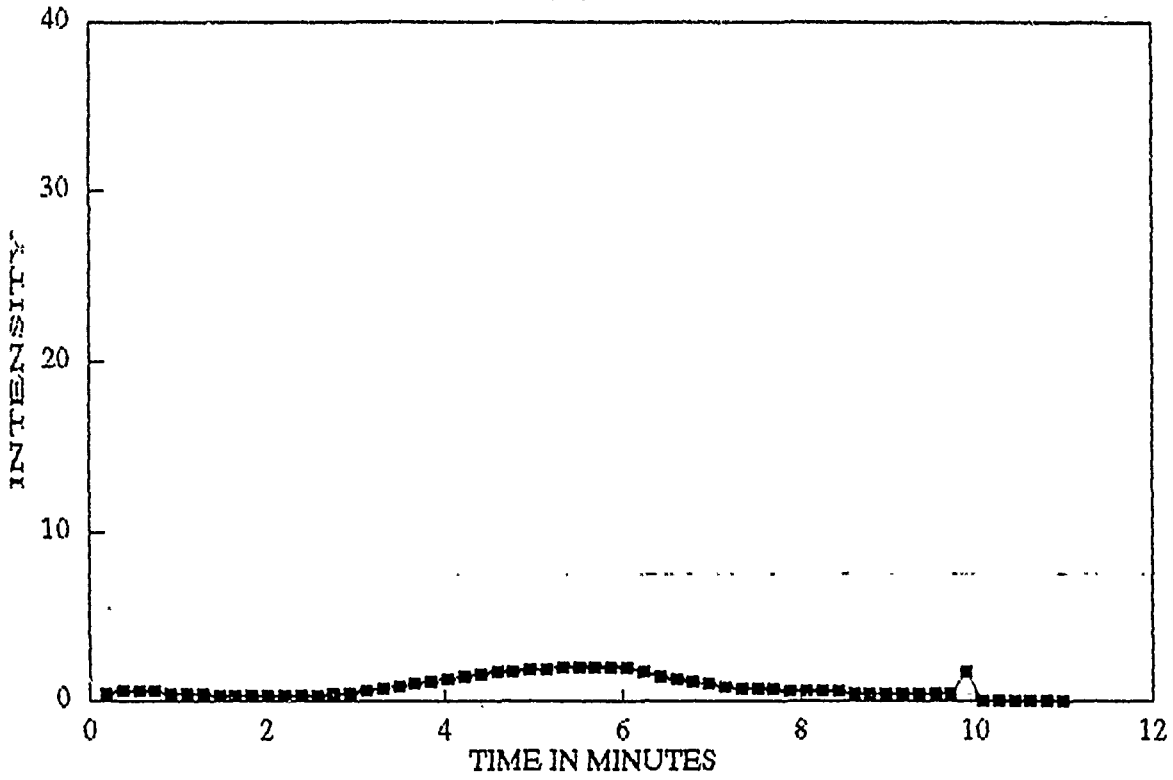


FIGURE 5

IRON (II) SULFATE



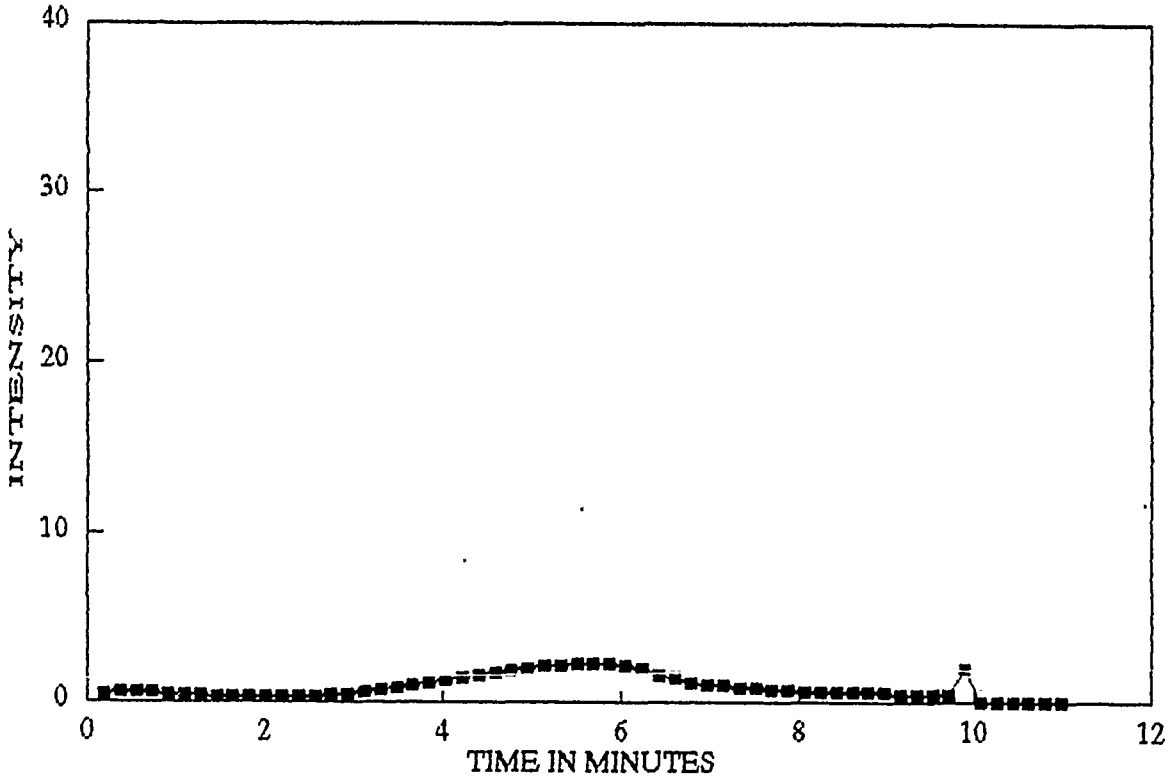
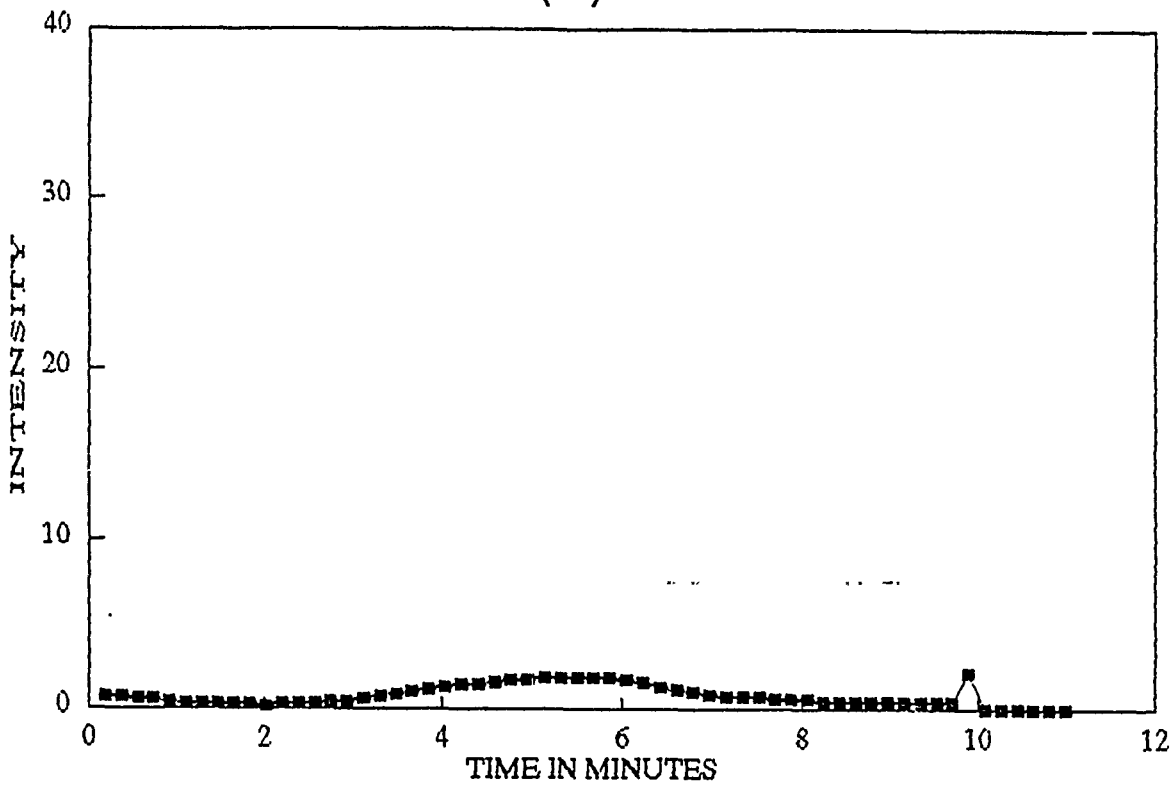


FIGURE 6

IRON (III) SULFATE



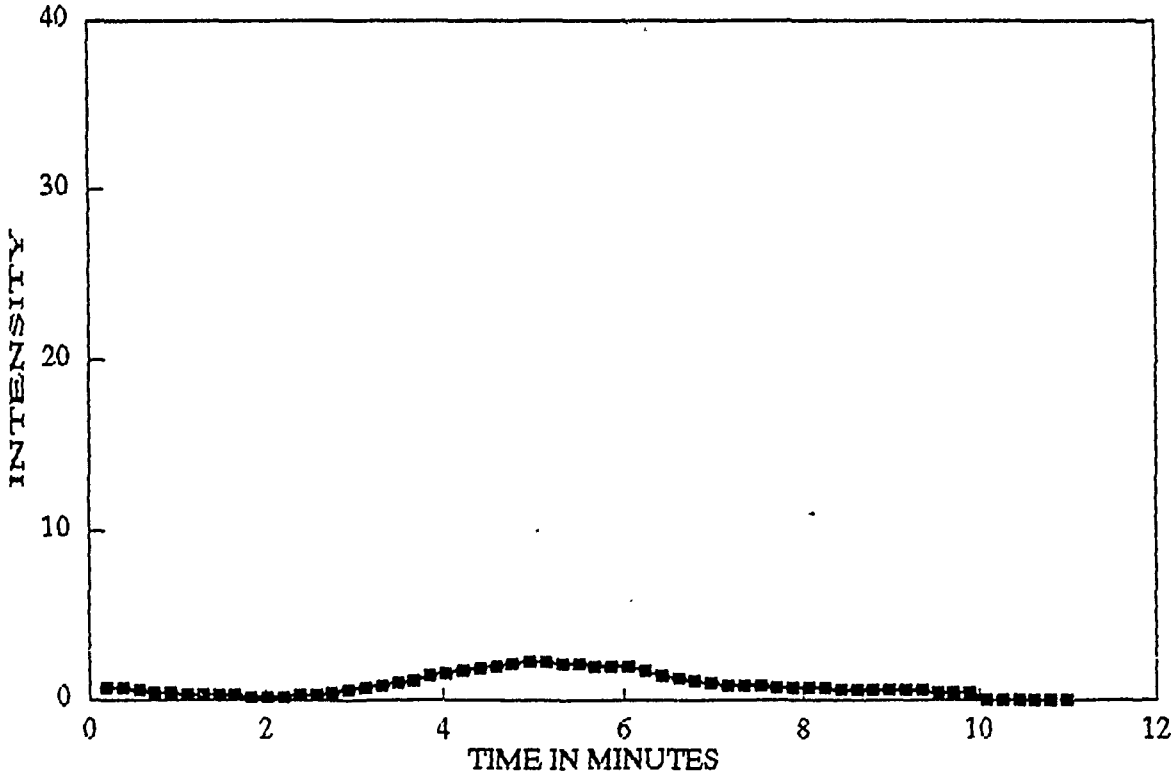


FIGURE 7

COBALT (II) SULFATE

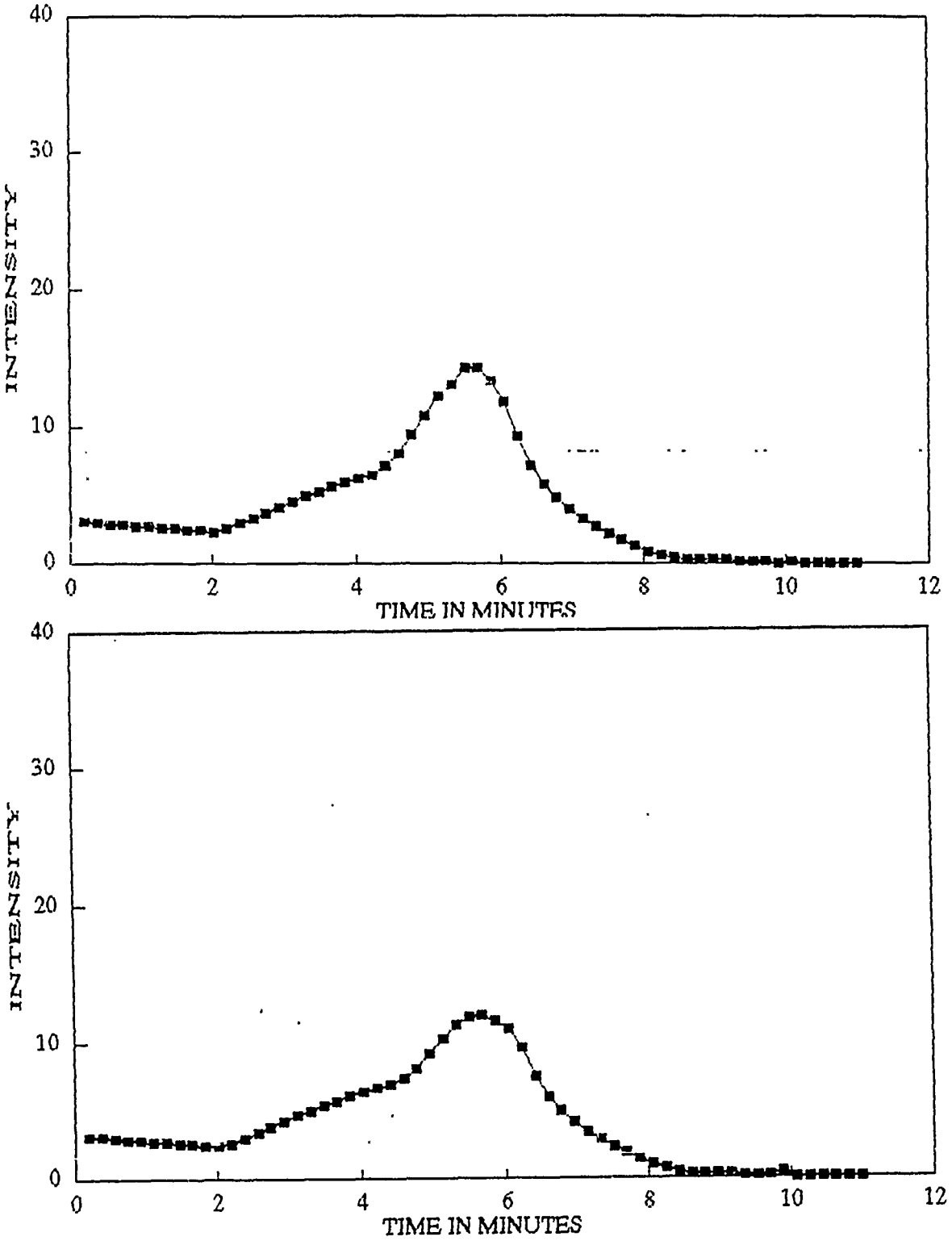


FIGURE 8

NICKEL (II) CHLORIDE

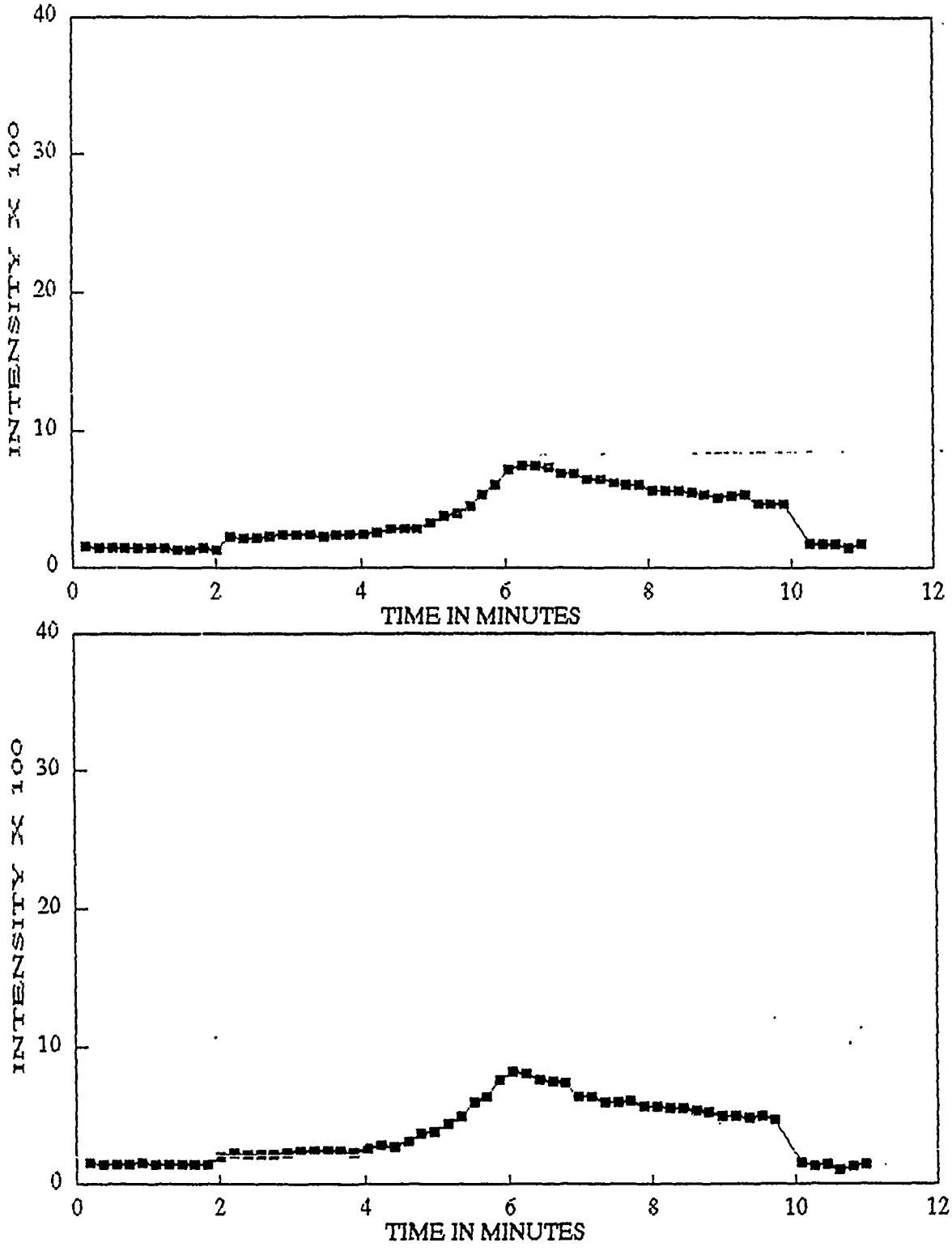


FIGURE 9

ZIRCONIUM (IV) CHLORIDE

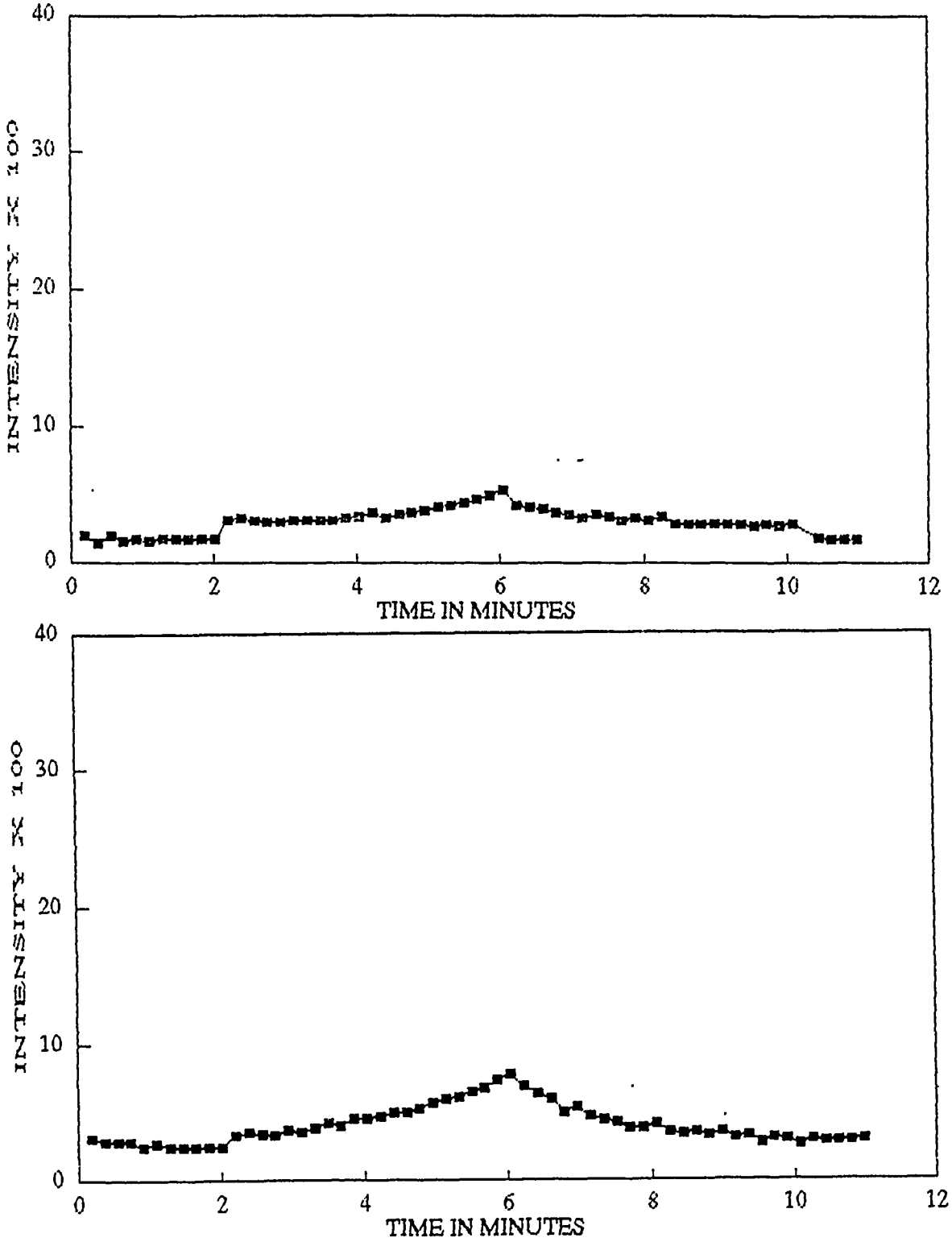


FIGURE 10 SCANDIUM (III) CHLORIDE

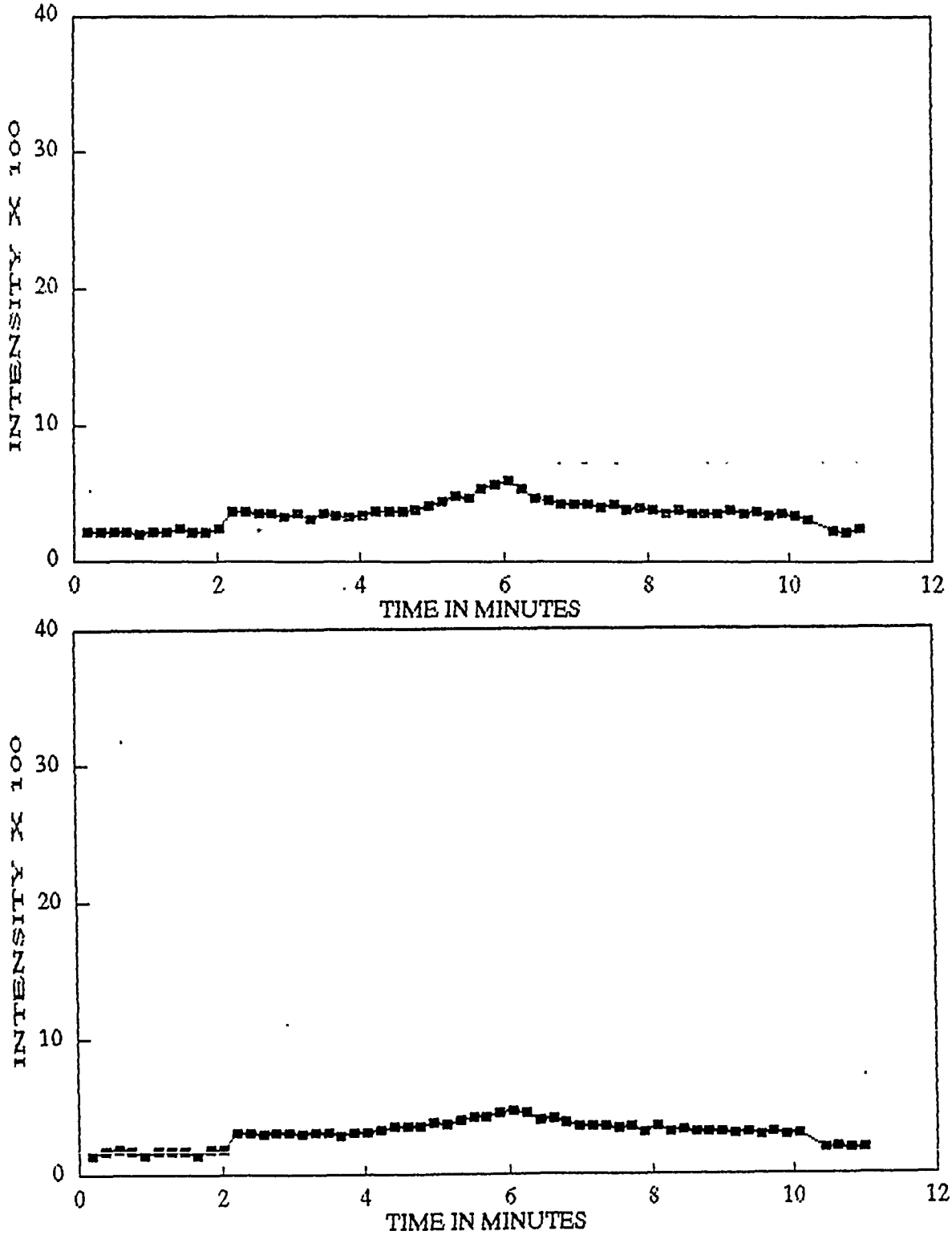
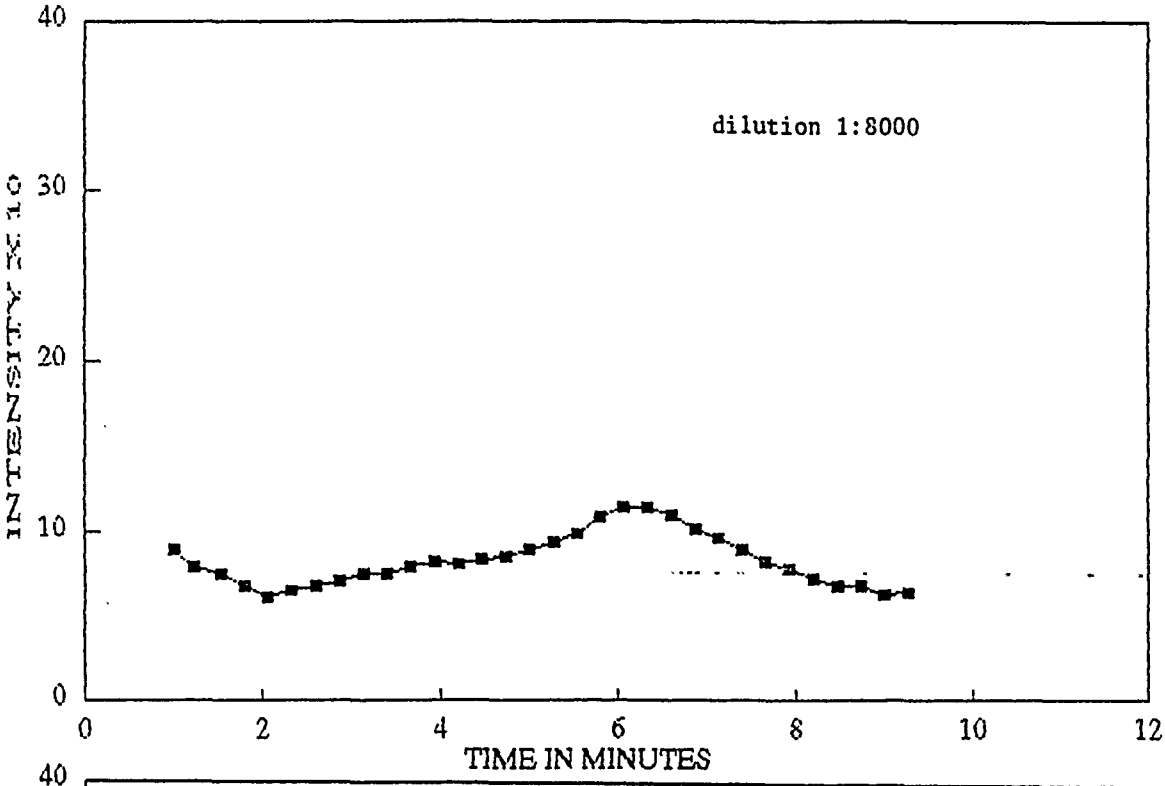


FIGURE 11

RUTHENIUM(III)CHLORIDE



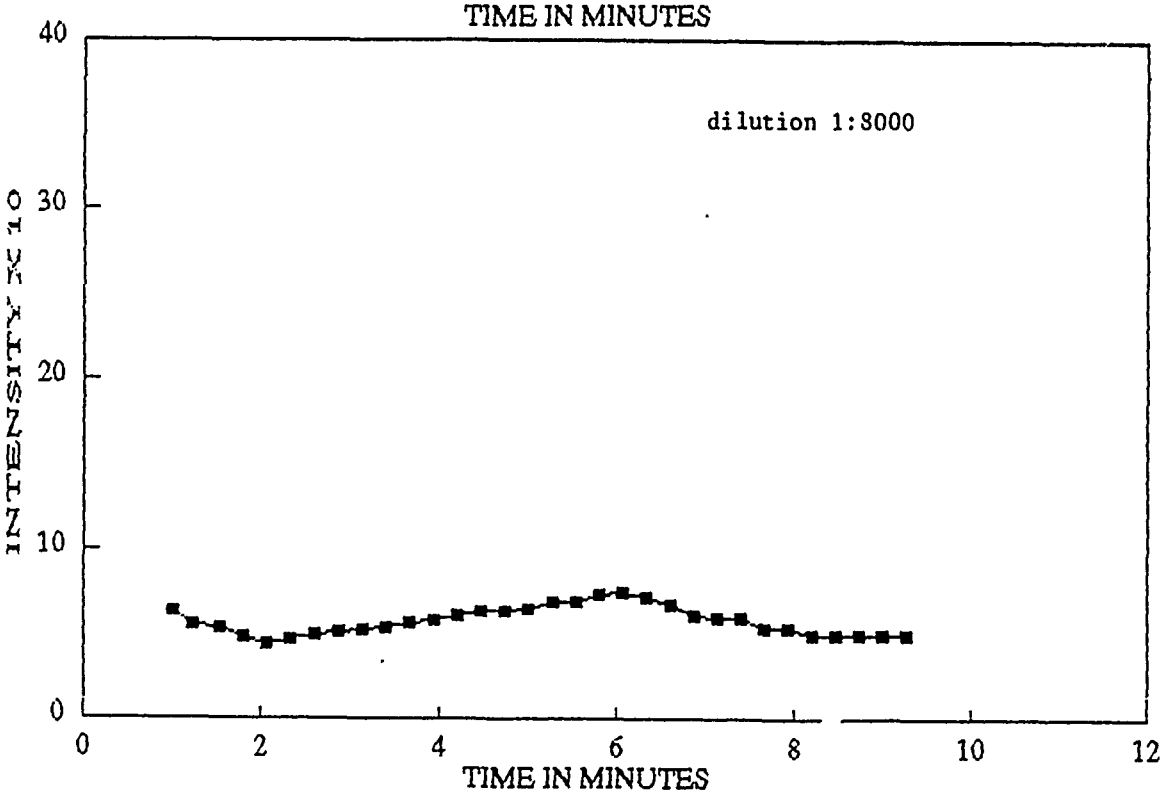


FIGURE 12

NIOBIUM (V) CHLORIDE

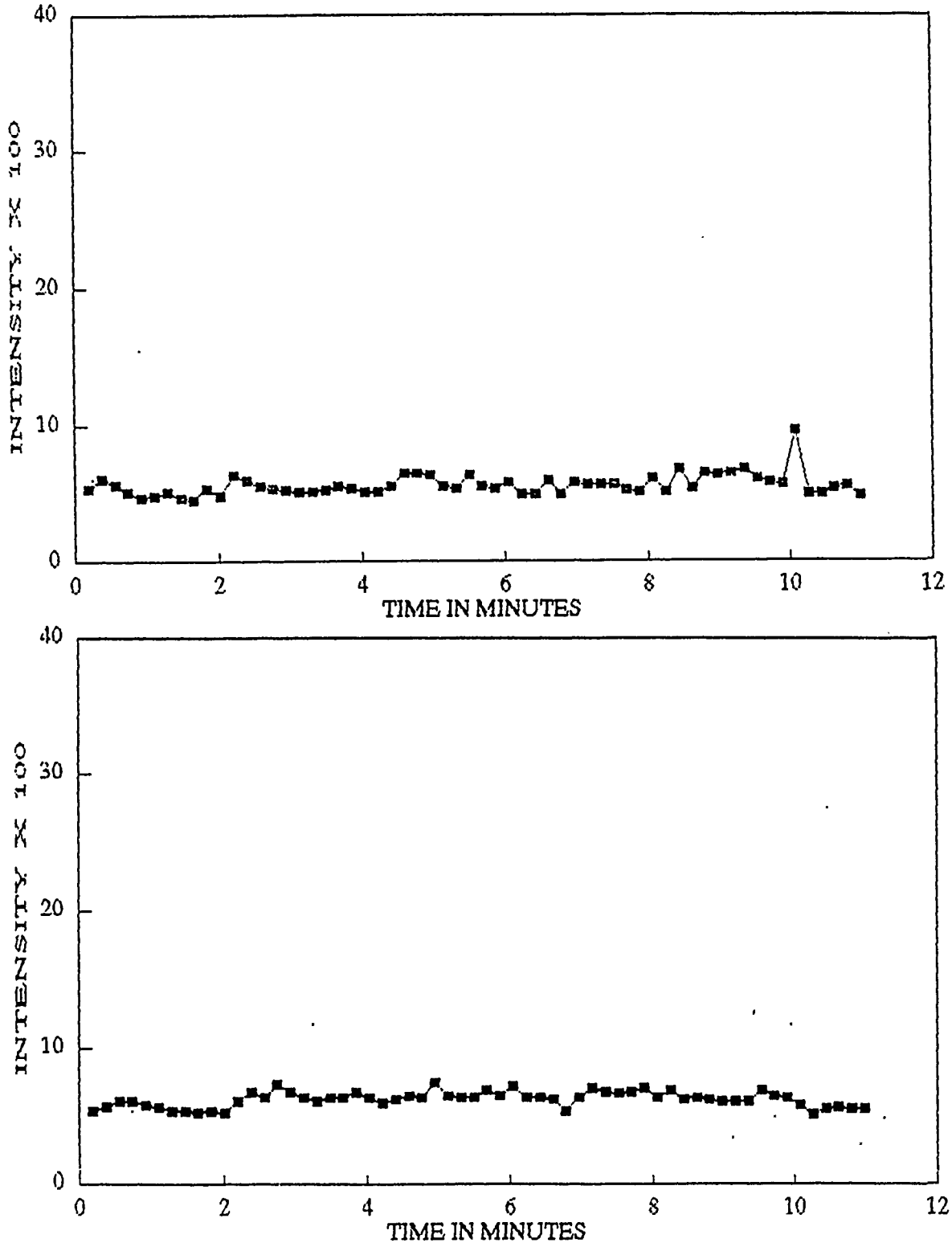


FIGURE 13

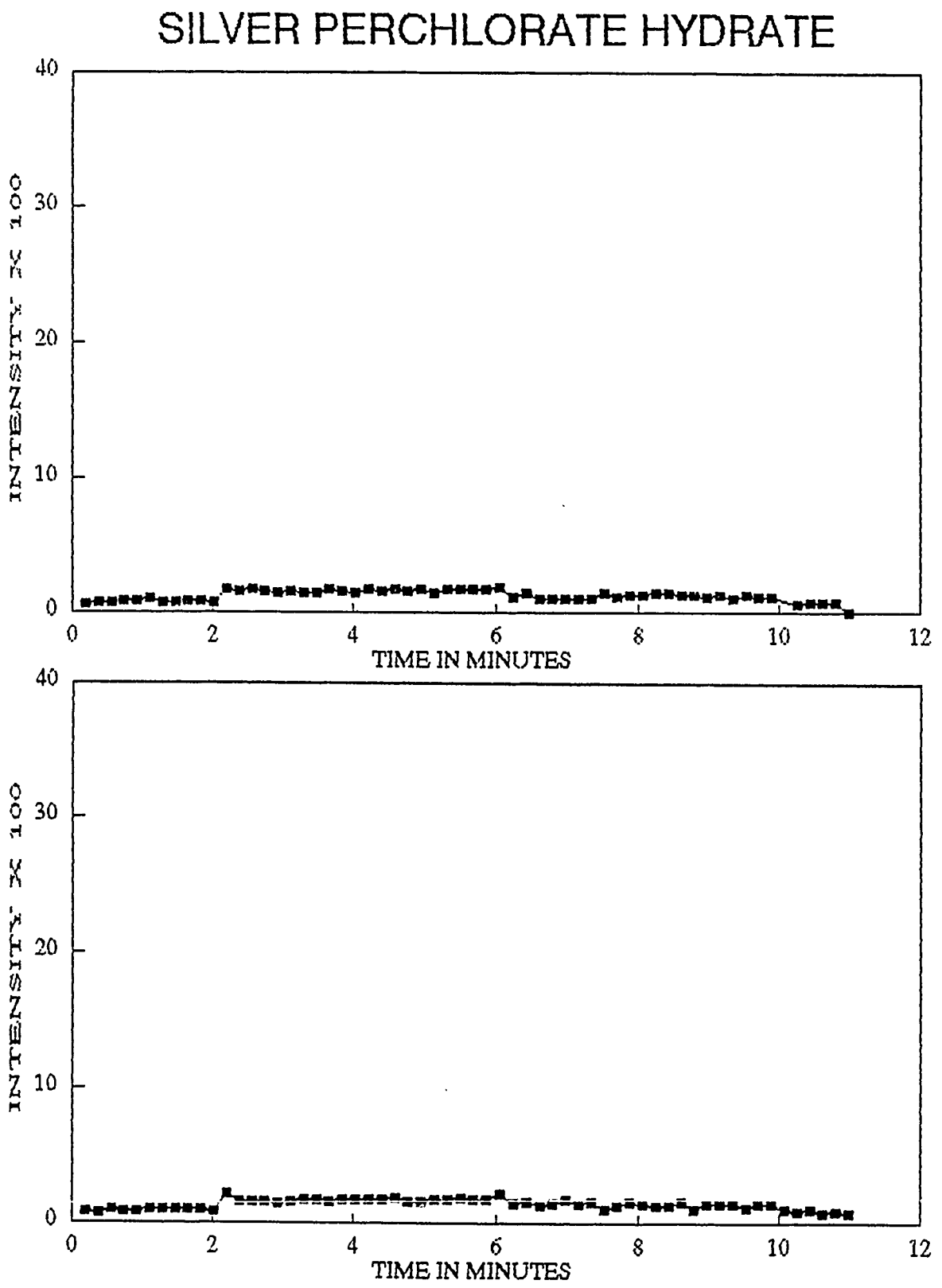
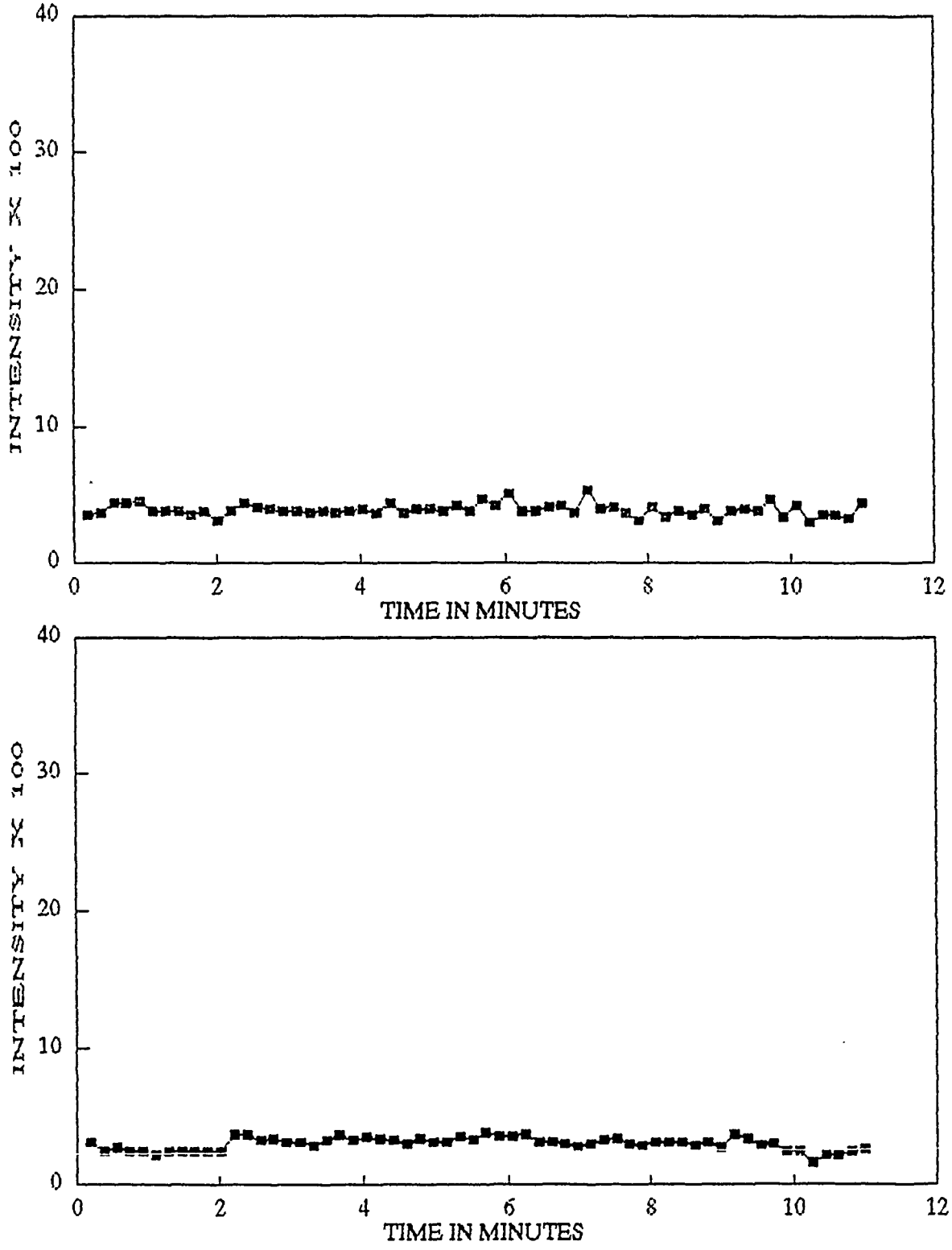

SILVER PERCHLORATE HYDRATE
INTENSITY X 100
40
30
20
10
0
0
2
4
6
8
10
12
TIME IN MINUTES
INTENSITY X 100
40
30
20
10
0
0
2
4
6
8
10
12
TIME IN MINUTES

FIGURE 14

VANADIUM (III) CHLORIDE



be detected during the initial 2 minute delay period. Different dosimeters were quite variable in their response to microwave heating. Table III presents the maximum intensity ratios, I/Io obtained during the heating period; I is the maximum observed intensity (during irradiation) and Io is the intensity at the start of the irradiation. The values in Table III are based on triplicate measurements from Appendix I.

Three of the dosimeters (niobium(IV), silver(I) and vanadium(III)) showed virtually no heating effects, and these appear to be of no use in imaging applications. All three of these have very low Io values (the plotted values are 100 times the measured values), and the small step at 2 minutes may be due to corona glow from the magnetron RF source.

An additional three of the dosimeters showed modest heating effects. In the order of best to worst, these were nickel(II), zirconium(IV) and scandium(III).

The dosimeters of copper(II), iron(II) and iron(III) began at fairly low Io values and attained respectable maxima of I/Io; these should be usable for imaging dosimetry. It is noted that the iron dosimeters peaked before 6 minutes, and the I/Io values refer to that peak. In view of the probable oxidation taking place here, these appear to be the same dosimeter, i.e., iron(III)!

The dosimeters of cobalt(II) and cobalt(III) are initially very luminous, but these also yielded usable heating responses. Both of these were sensitive to dissolved carbon dioxide (see 4.5.), which raises a suspicion that they might be identical, as in the case of the iron dosimeters; however, if that were true, then the differences seen in Table III would have to be attributed to a counterion effect.

TABLE III
Thermal Sensitivities of Various Dosimeters

Source of Metal	I/I_o(max)	S.D.
$CuCl_2$	35.6	16.7
$FeSO_4$	7.4	1.1
$Fe_2(SO_4)_3$	7.6	1.0
$CoSO_4$	6.2	1.3
CoF_3	12.2	2.9
$NiCl_2$	4.4	0.9
$ZrCl_4$	2.6	0.7
$ScCl_3$	2.3	0.8
$NbCl_4$	0	-
$AgClO_4$	0	-
VCl_3	0	-

TABLE IV
Sonochemiluminescent Effects 25 C

Source of Metal	I/I_o*	S.D.
$ScCl_3$	2.6	0.1
$CoSO_4$	2.1	0.1
$ZrCl_4$	1.6	0.1
$Fe_2(SO_4)_3$	1.3	0.01
$CuCl_2$	1.2	0.03

*From I values immediately before (I_o) and after applying the acoustic field

The most luminous dosimeter of all was based on ruthenium(III), but it is too efficient, causing the luminol to be expended in a burst. More importantly, it was not very temperature-dependent. Time ran out before we were able to examine ruthenium(III) under a wider range of conditions, especially for the case of aged preparations of solution B (the catalyst becomes more active on standing!), and other potential candidates such as lead(II) also need to be examined in future work.

Any attempt at a mathematical modeling of these time/heating curves is liable to ignore the great potential for complexity in the reacting systems. Nevertheless, a fairly simple approach developed here does mimic the observed behaviour of the various dosimeters. The model begins with a reasonable assumption that the catalyst acts to decompose hydrogen peroxide to oxygen and water, producing free radicals in the process. The radicals in turn activate the chemiluminescence of luminol (3); free radicals are the currently accepted initiators of this mechanism. All of the luminometers effervesce fine bubbles of oxygen, and the bubbling rate was most pronounced in the more luminous dosimeters of cobalt and ruthenium (the latter behaves anomalously in that the rate of decomposition of hydrogen peroxide <u>increases</u> with time in solution B).

Owing to the very low catalyst concentration, the radical forming reaction is expected to be zero-order in hydrogen peroxide and first order in the effective catalyst concentration; also, this step is presumed to be the most temperature dependent step in the whole sequence of reactions. Thus:

$$H_2O_2 \xrightarrow{\text{catalyst}} R\cdot \qquad 1.$$

From which we write:

$$\frac{d[R\cdot]}{dt} = k_1[\text{catalyst}] \qquad 2.$$

And, sequentially, the radicals react with the luminol anion:

$$R\cdot + \text{luminol}^- \xrightarrow{k_2} R^- + \text{luminol}\cdot \qquad 3.$$

This is reasonable since the chemiluminescence is initially first-order with respect to luminol. Ignoring recombination of radicals, the rate based on the above is:

$$\frac{d[R\cdot]}{dt} = -k_2[R\cdot][\text{luminol}^-] \qquad 4.$$

If the latter is a fast radical exchange mechanism, Ea should be relatively low and k2 relatively insensitive to temperature (compared with k1). Thus (with some license!) the steady state approximation is applied:

$$k_1[\text{catalyst}] - k_2[R\cdot][\text{luminol}^-] \simeq 0 \qquad 5.$$

Or:

$$[R\cdot] \simeq \frac{k_1[\text{catalyst}]}{k_2[\text{luminol}^-]} \qquad 6.$$

In this expression, the catalyst concentration is constant and the increase of k2 with temperature tends to compensate the decreasing concentration of luminol: thus, as a crude approximation the radical concentration is proportional to k1, and the apparent rate constant for a first-order decomposition of luminol, i.e., $k \propto k_1k_2$, should reflect the temperature dependence of the radical-forming process. The rate of decomposition of luminol is thus:

$$\frac{d[\text{luminol}]}{dt} = \frac{dC}{dt} = kC \qquad 7.$$

And we set the chemiluminescent intensity proportional to this rate. (This is not necessarily a good assumption since some of the metals may act to quench excitation energy). There are other approaches which can lead to a similar rate expression. For example, a metal might remove an electron during a direct interaction with luminol, then be reoxidized by hydrogen peroxide.

Integration of the rate equation to obtain C as a function of time must take into account the temperature dependence of k if heating by a microwave source is taking place:

$$\frac{dC}{dt} = A_o \left[e^{\frac{-Ea}{RT}} \right] C \qquad 8.$$

The value of Ea is probably invariant over the 35 C temperature range considered in these experiments. In the last equation, T is time dependent and may be approximated as T = Bt + To, where B is the heating rate in degrees C/min, t is the elapsed time and To the temperature at the start of the irradiation. An evaluation of the reaction rate (thus I/Io) using equation 8 is not straightforward, and it leads to a complex exponential-integral and logarithmic-integral treatment since both the Boltzman factor and C are time dependent. The numerical algorithm of Appendix III achieves the same purpose, and Figures 15 through 17 show the effects of Ea and Ao on time/heating curves; these simulations used the actual conditions of the experiments (the measured heating rate and luminol concentration). It should be noted that in using the computer algorithm, Ao values should be near those predicted by the following equation to avoid over- or underflow conditions. This equation applies only to the range of 15 <

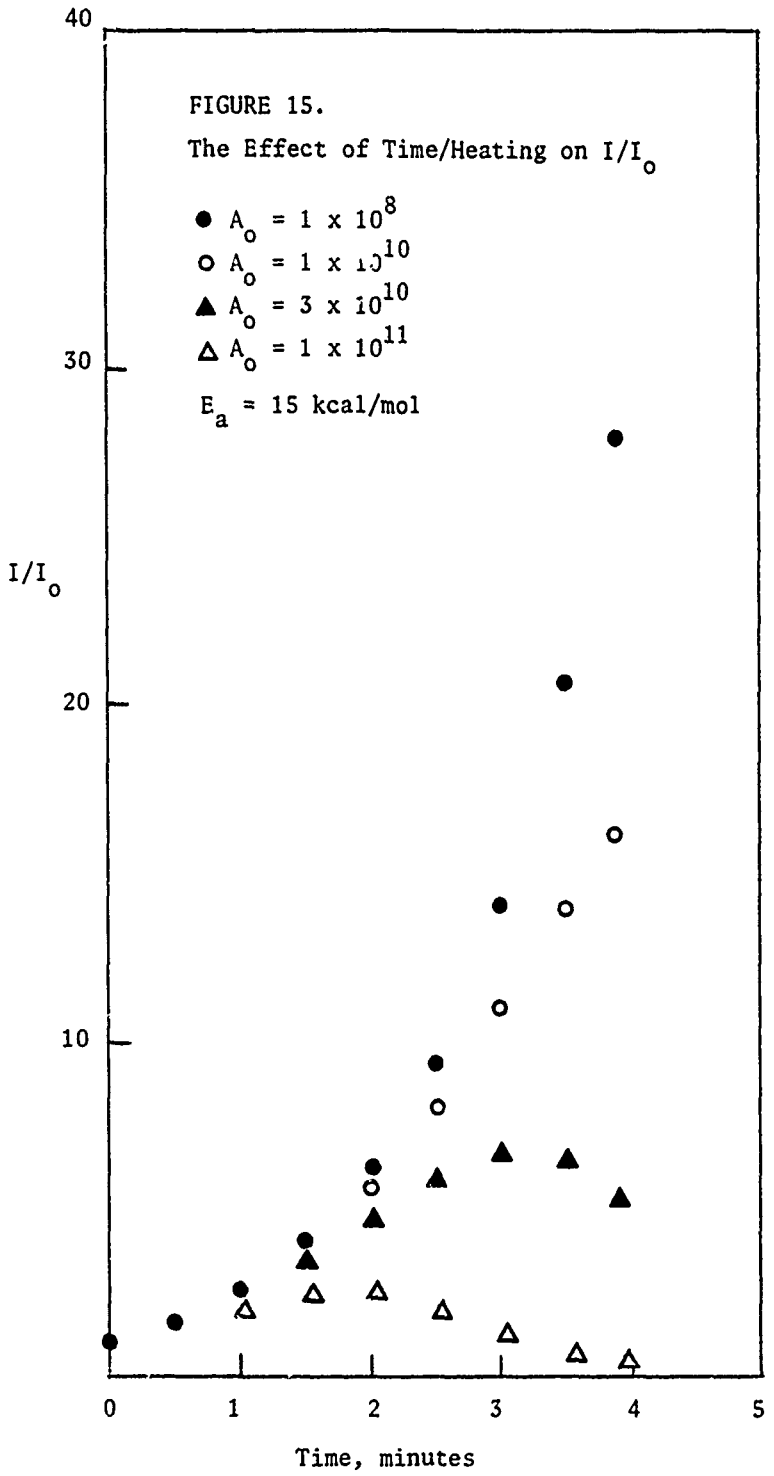
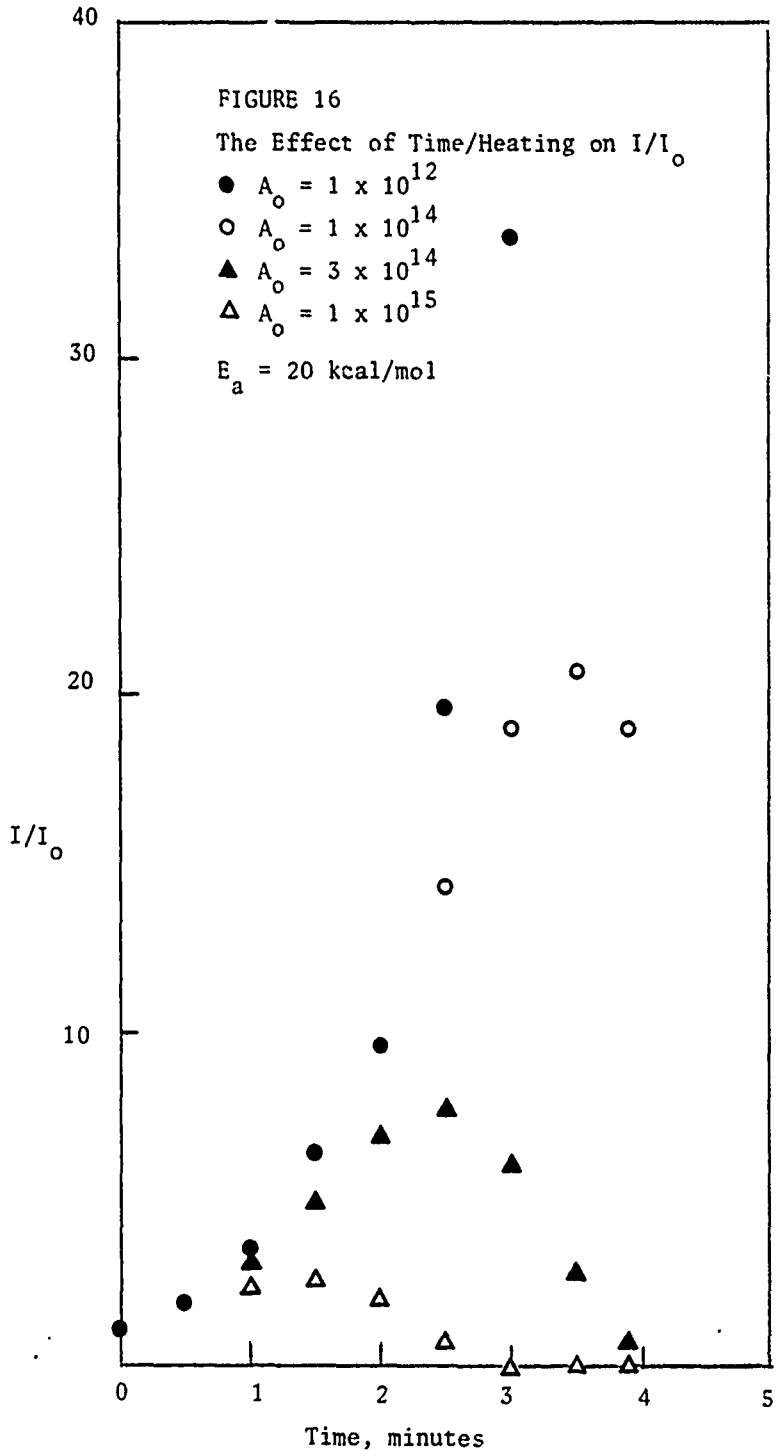


FIGURE 15.
The Effect of Time/Heating on I/I_o


FIGURE 16
The Effect of Time/Heating on I/I$_o$
● A_o = 1 x 10^{12}
○ A_o = 1 x 10^{14}
▲ A_o = 3 x 10^{14}
△ A_o = 1 x 10^{15}
E_a = 20 kcal/mol
40
30
20
10
I/I$_o$
0
1
2
3
4
5
Time, minutes

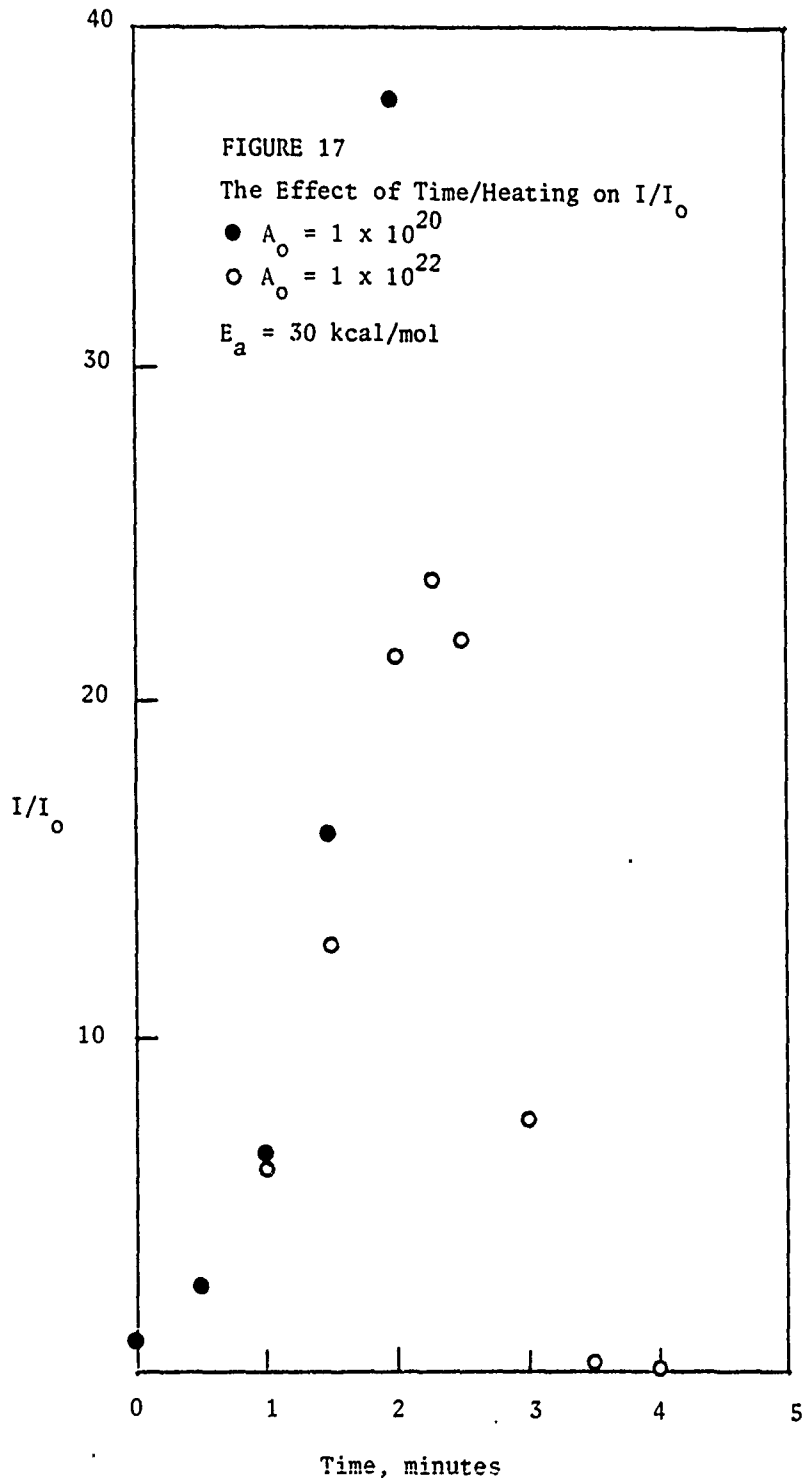


FIGURE 17

The Effect of Time/Heating on I/I_o

● $A_o = 1 \times 10^{20}$

○ $A_o = 1 \times 10^{22}$

E_a = 30 kcal/mol

Ea <40 kcal/mol, and realistic values for metal oxide catalysts are probably in the 15 - 20 kcal/mol. range:

$$A_o = 10^{0.69\ Ea} \qquad 9.$$

One must be careful in interpreting these families of curves. For example, the upper curve of a set (the lowest pre-exponential factor) actually involves the lowest values of chemiluminescent intensity. At higher values of Ao (the lower curves) one begins to see the kind of light pulse associated with the more conventional forms of luminometry. It should also be noted that the initial rate of change of I/Io with time/temperature is independent of Ao and entirely dependent on Ea (the change in T simply alters the value of the Boltzman factor before C is much decreased). A "best" dosimeter is thus one with Ao large enough to produce detectable light in the imaging system and with the highest possible Ea. These conditions obtain in some of the systems under investigation if irradiation times are not allowed to exceed 0.5 minutes. Light intensity increments of 50% or more associated with temperature increments of a few degrees are in fact demonstrated in the data. This should be more than enough to produce good image contrast in a CCD detector array. Convection effects might be minimized if the dosimeter were made somewhat more viscous. Very energetic pulse fluences (as produced in a HPM generator) would achieve the same effect, and a data sampling immediately after the applied T-step would circumvent convective effects.

The theoretical curves of Figure 15 actually resemble the dosimeters under consideration. For example, the copper(II) dosimeter produced a variation of I/Io resembling that of the curve next to the

top in Figure 15, suggesting an Ea of 15 kcal/mol. (a value of about 19 kcal/mol. was measured during the summer program at Brooks AFB in 1988), and the dosimeters of cobalt and iron resemble the curve below that. Concentrated ruthenium(III) dosimeters were similar to the lower curve of Figure 16 since they were completely extinguished before much progress into the irradiation (the most concentrated versions extinguished before the irradiation started!). Table II did not predict the actual temperature dependence of the ruthenium(III) dosimeter, which is relatively low. This dosimeter appears to be characterized by a large Ao and a relatively low value of Ea.

4.3. Sonochemical Effects

During the last month of this project some attention was focused on the sonochemical response of a few of the dosimeters, and the results of these studies are presented in Figures 18 through 29, which show replicate measurements at beginning temperatures of 5 and 25 degrees C. As in the case of microwave heating experiments, dosimeters were given a 2 minute delay before the 50 Watt accoustic irradiation was started. During this initial period the dosimeters show the characteristic isothermal decay of chemiluminescence, and the effect is more pronounced at 25 C. The acoustic input power is sufficient to produce heating on the sample, and the more temperature sensitive dosimeters show climbing timecurves after the start of irradiation. In these systems I/Io is defined as the immediate response on applying sonic power; measured values are collected in Table IV. The most interesting result of this set of experiments is the large I/Io observed for the scandium(III) dosimeter, which does

FIGURE 18 SCANDIUM (III) CHLORIDE

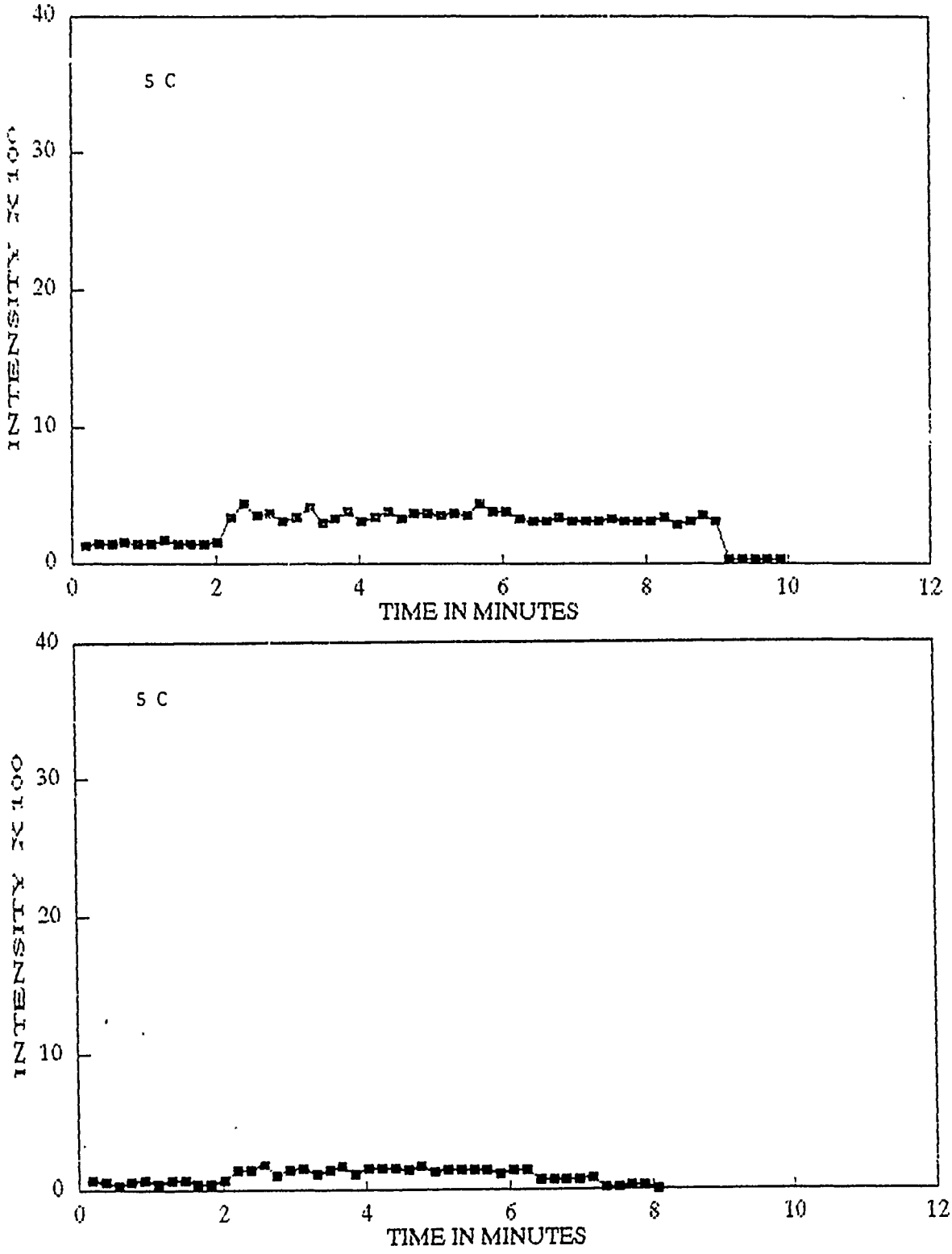
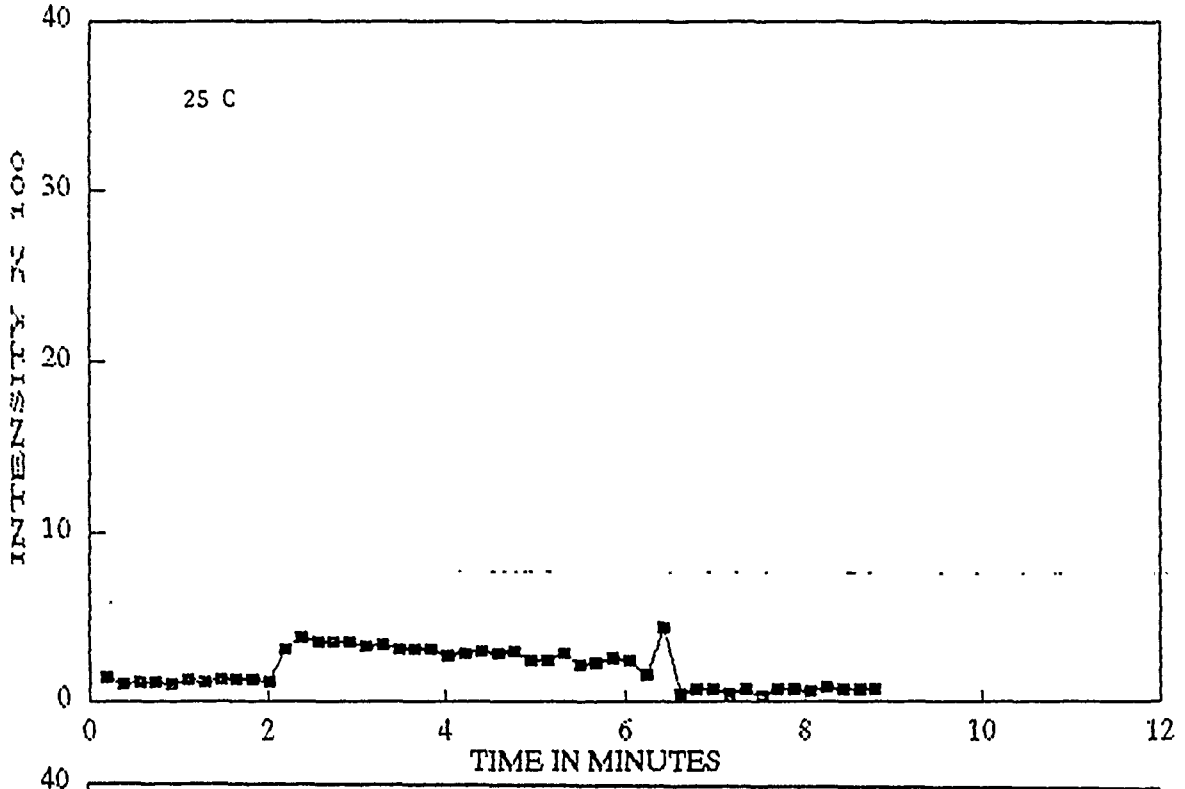


FIGURE 19

SCANDIUM (III) CHLORIDE



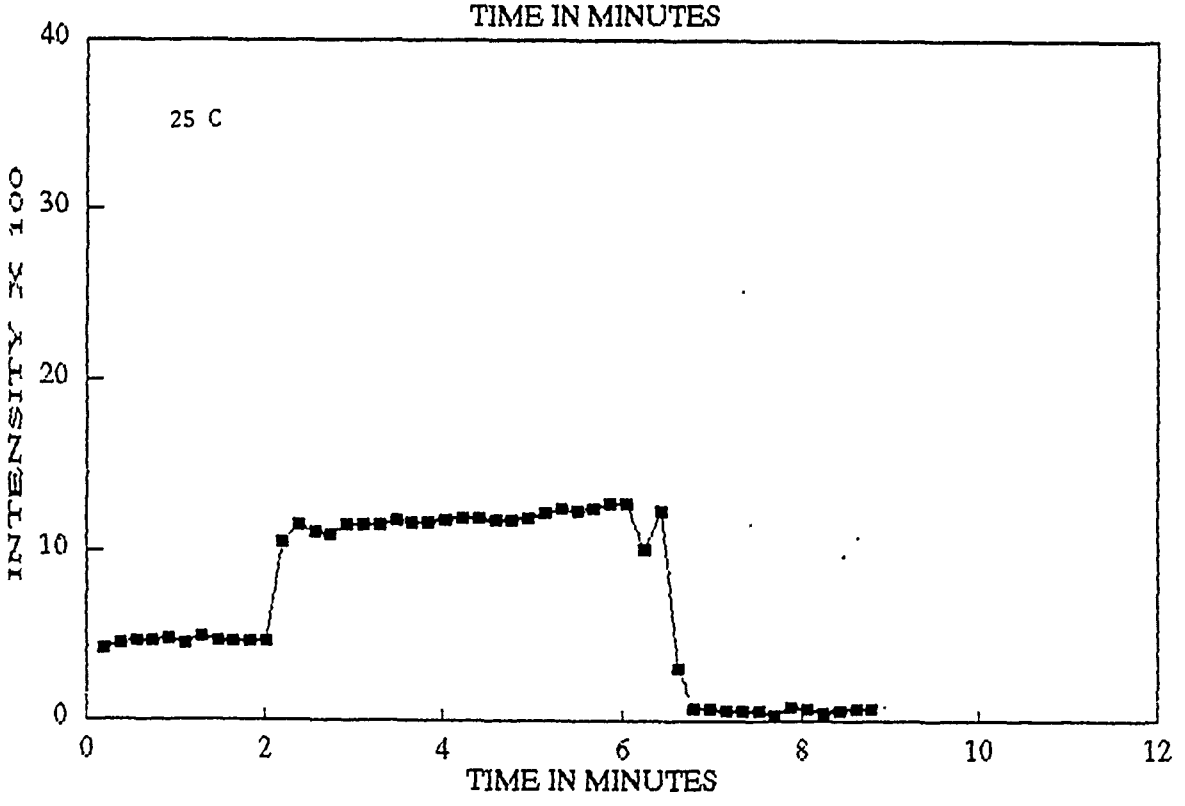
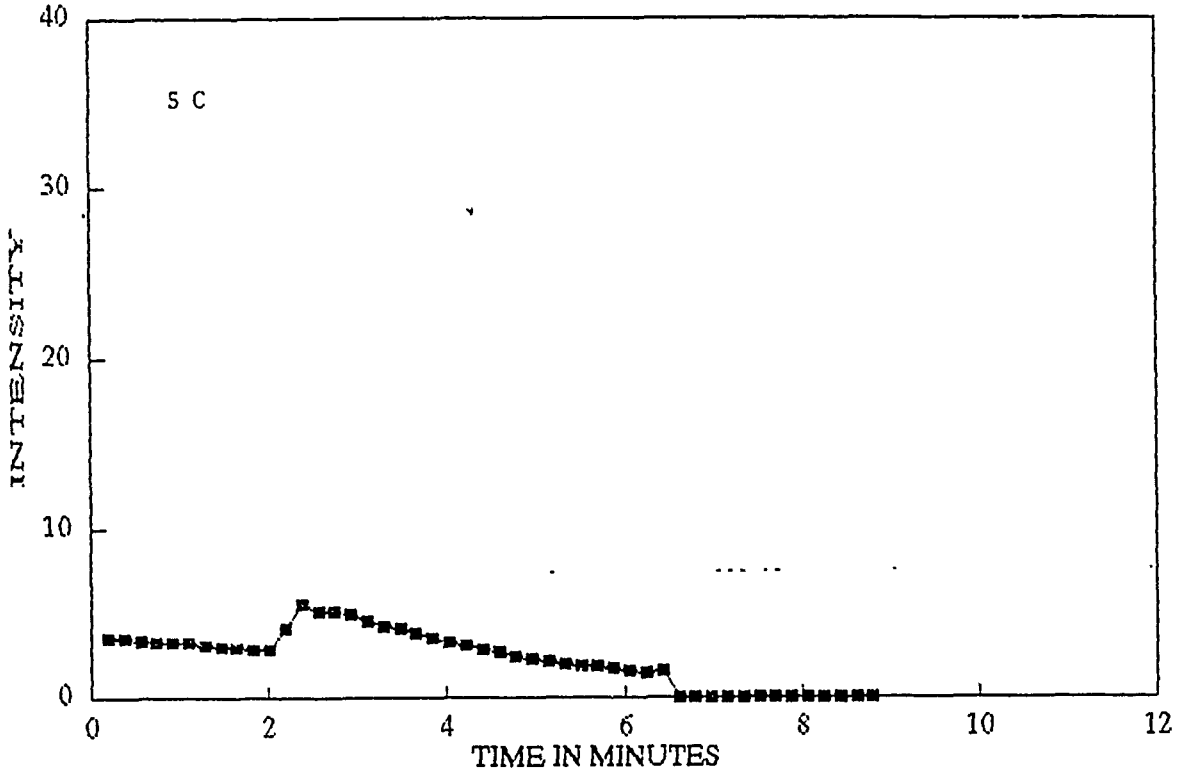


FIGURE 20

COBALT (II) SULFATE



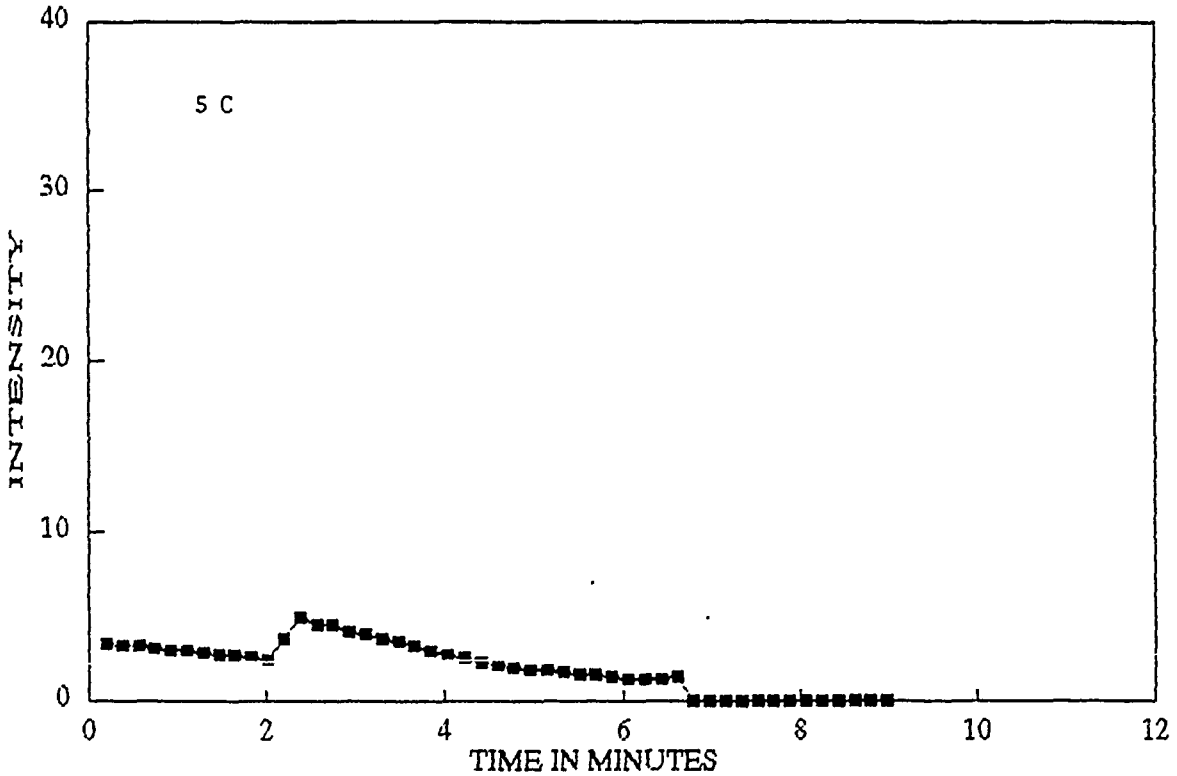
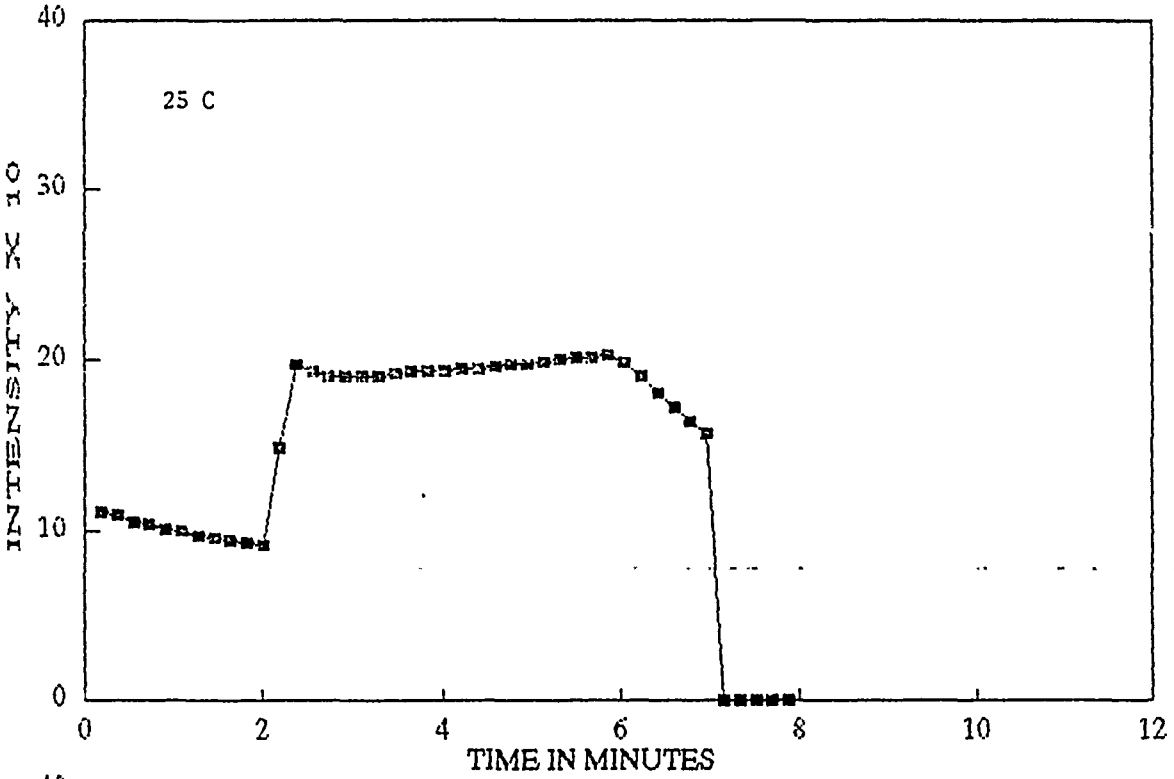


FIGURE 21

COBALT (II) SULFATE



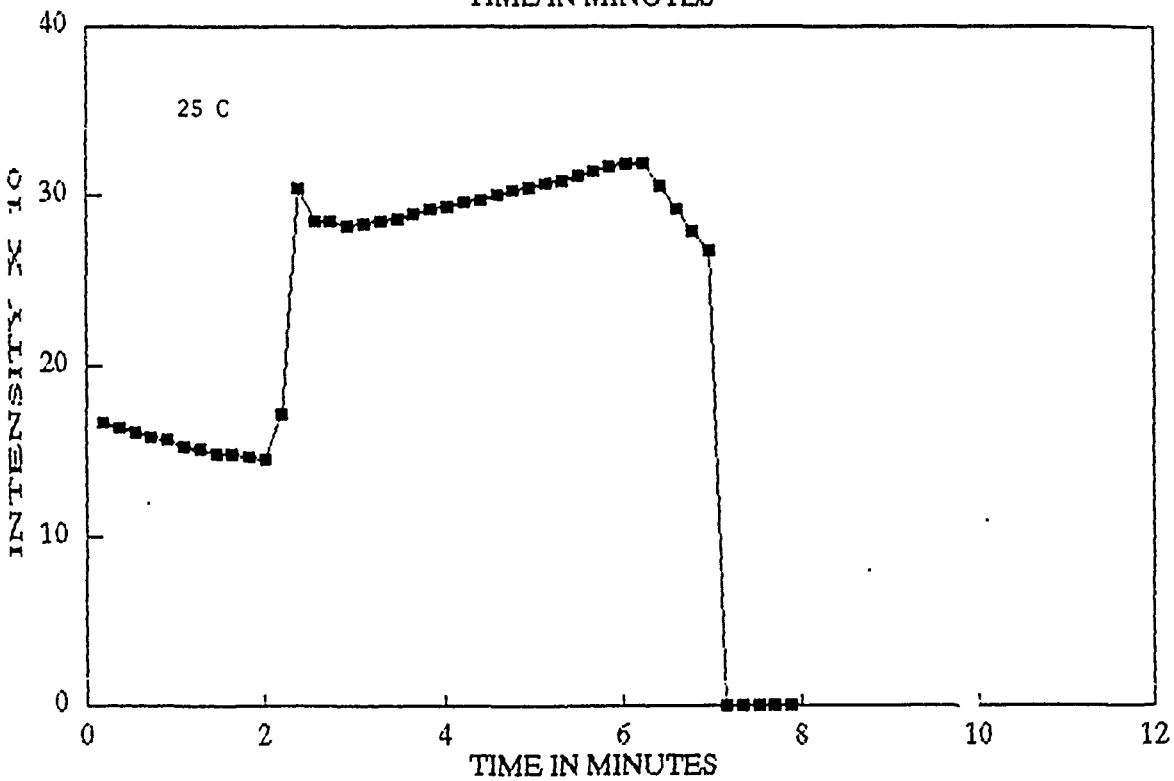
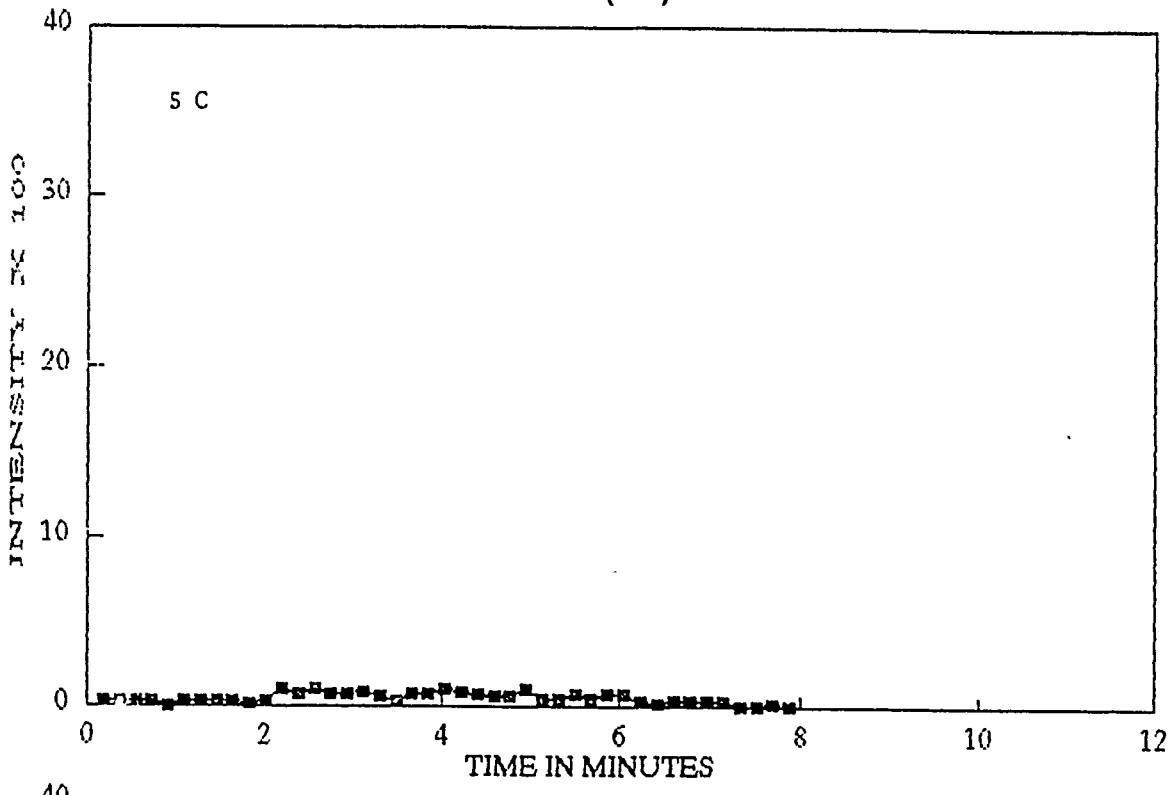


FIGURE 22 ZIRCONIUM (IV) CHLORIDE



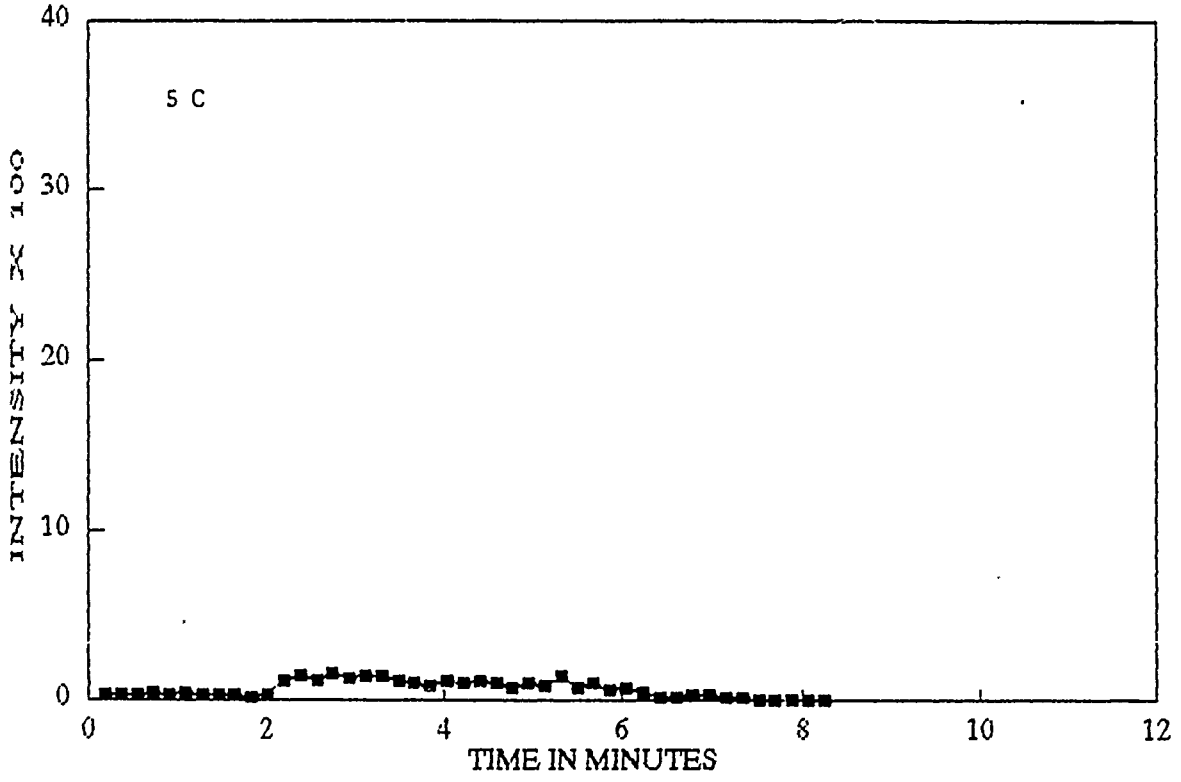


FIGURE 23

ZIRCONIUM (IV) CHLORIDE

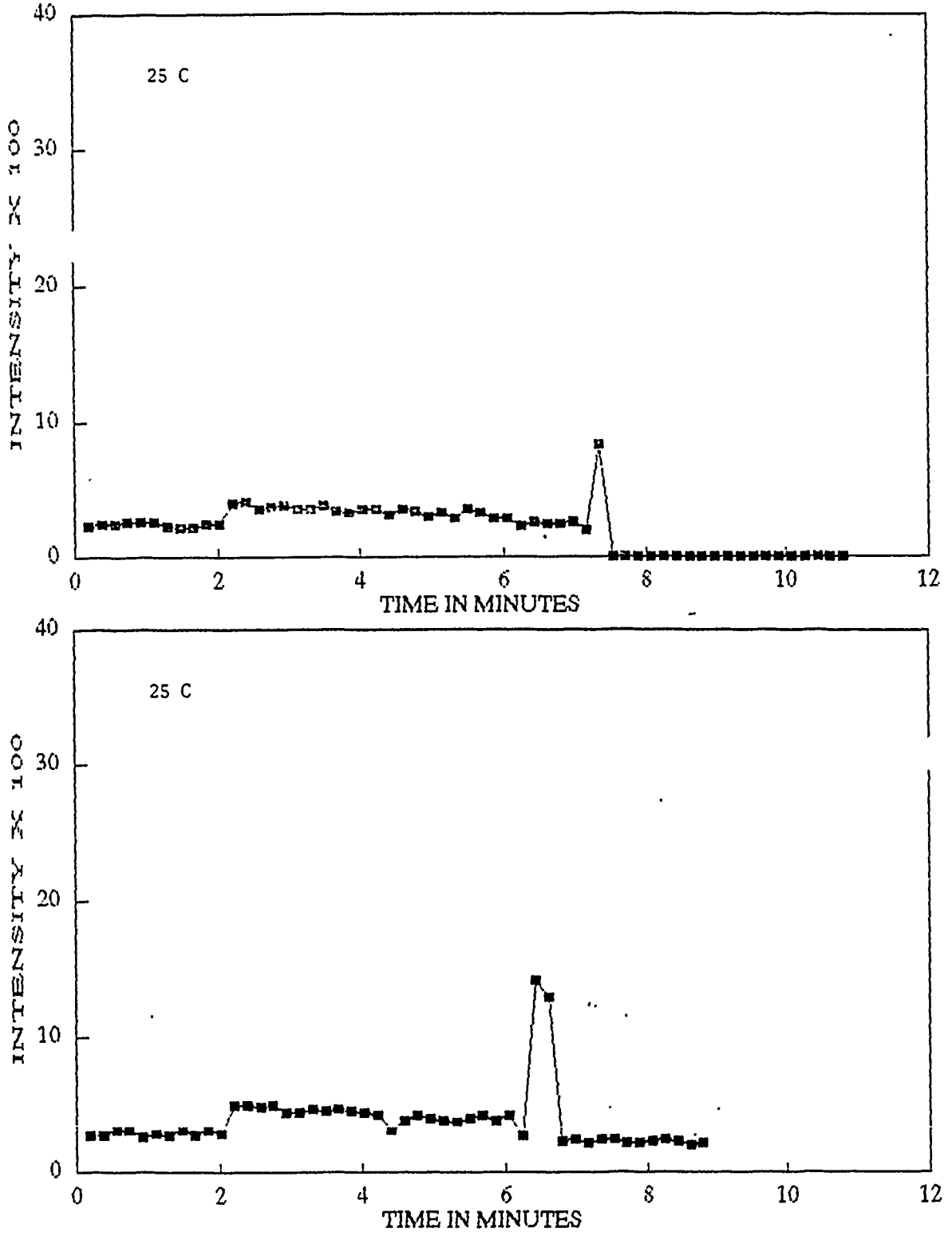
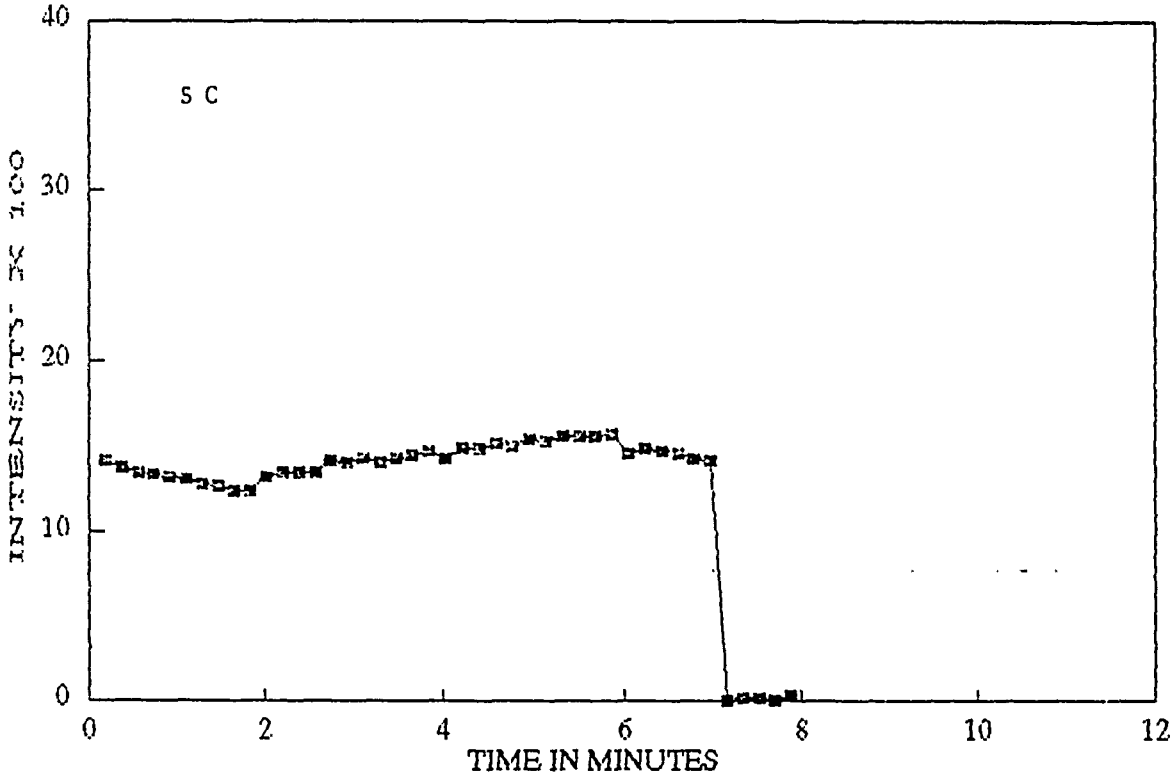


FIGURE 24

IRON (III) SULFATE



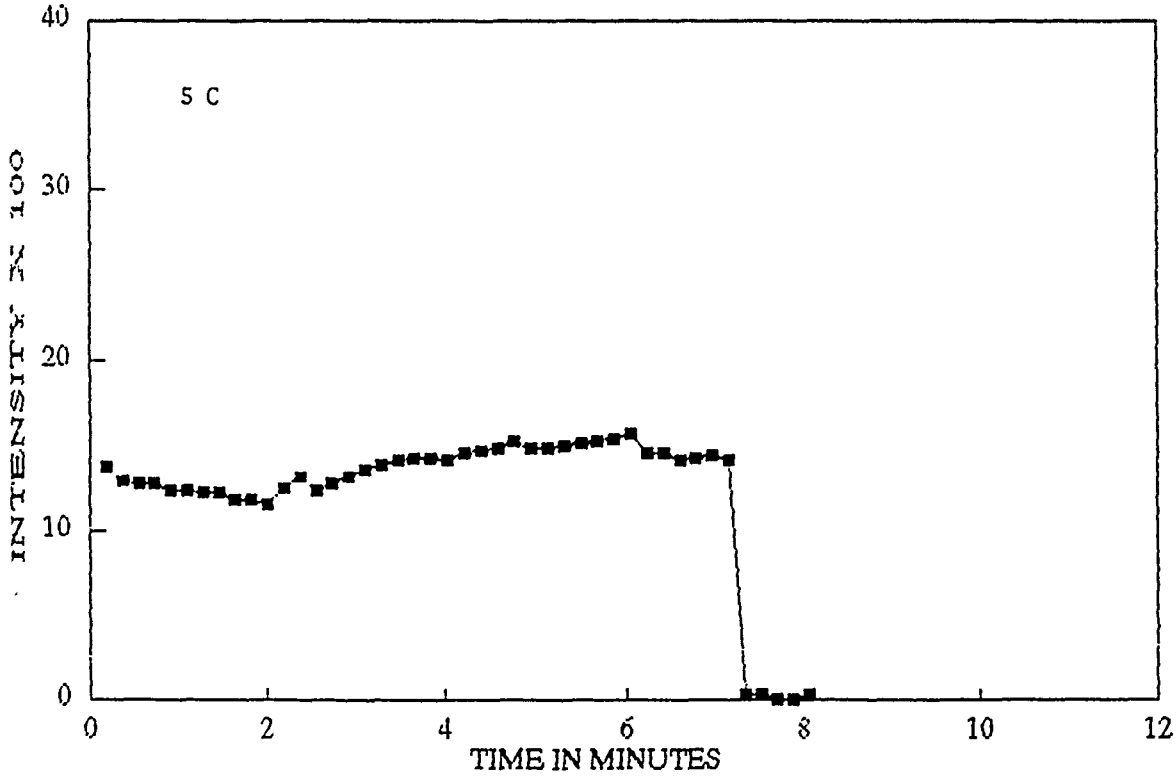


FIGURE 25

IRON (III) SULFATE

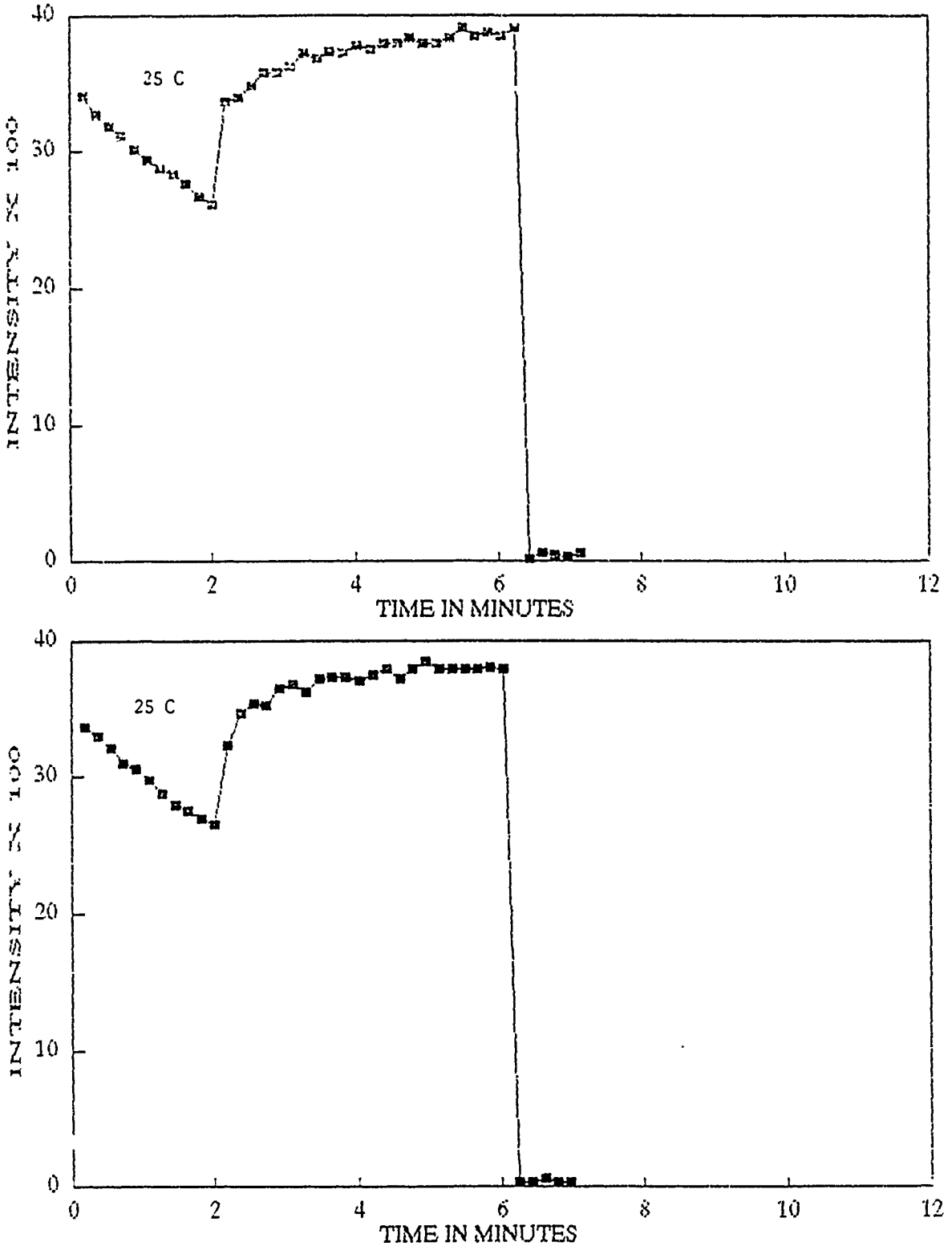
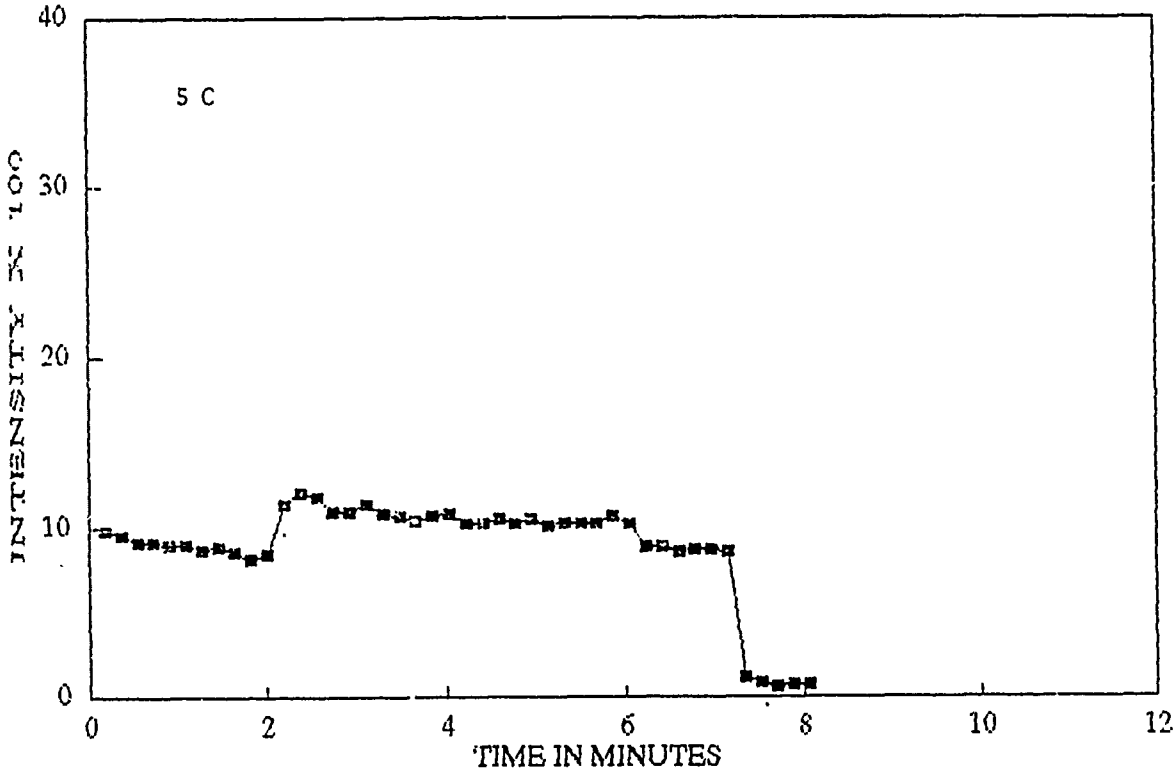


FIGURE 26

COPPER SULFATE



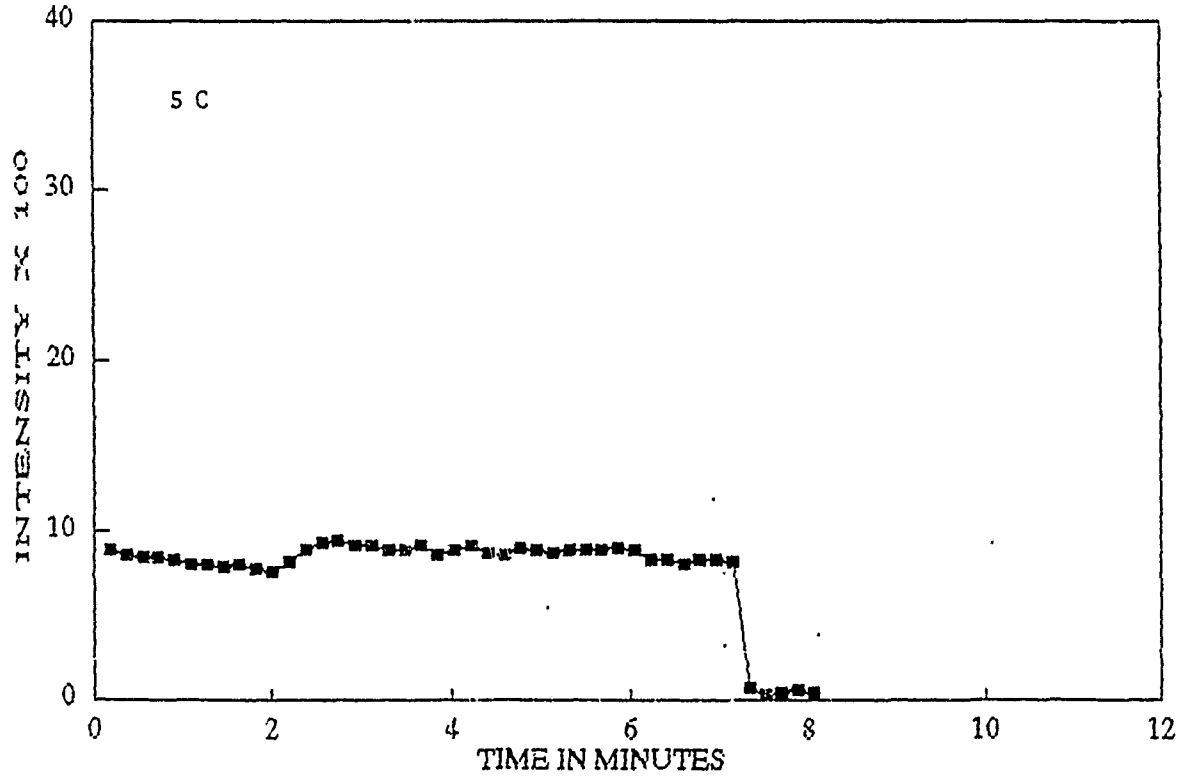
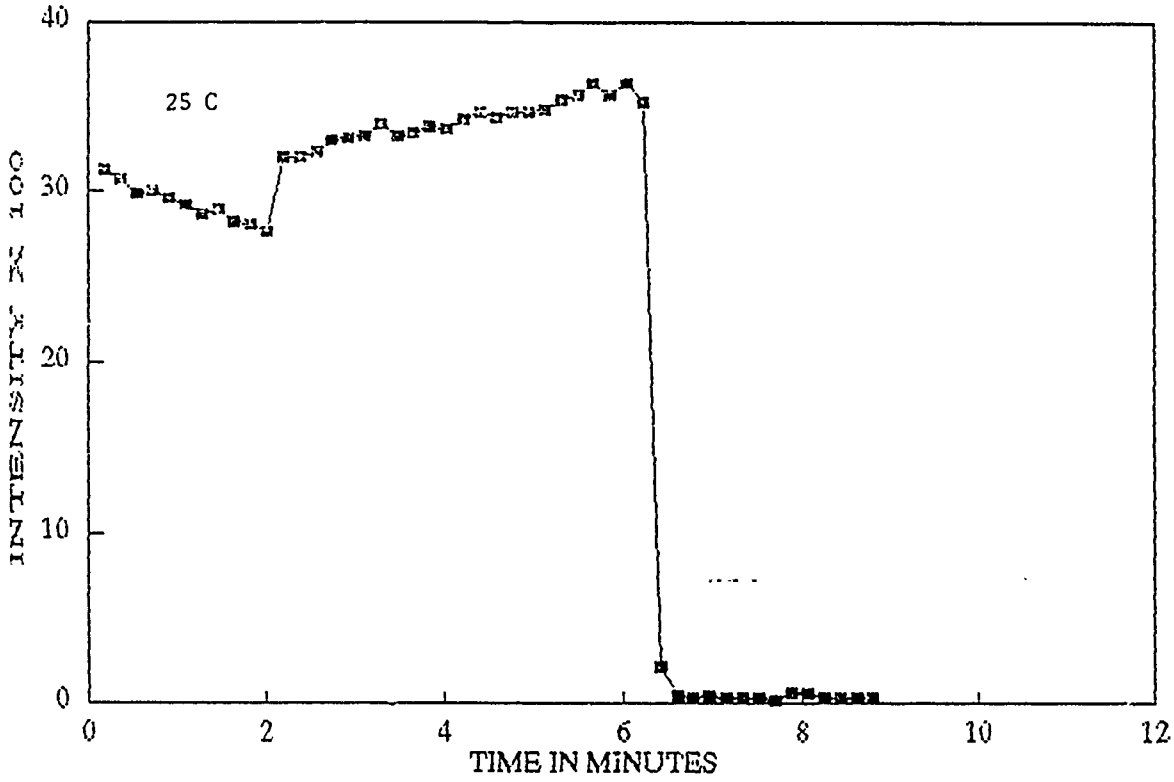


FIGURE 27

COPPER SULFATE



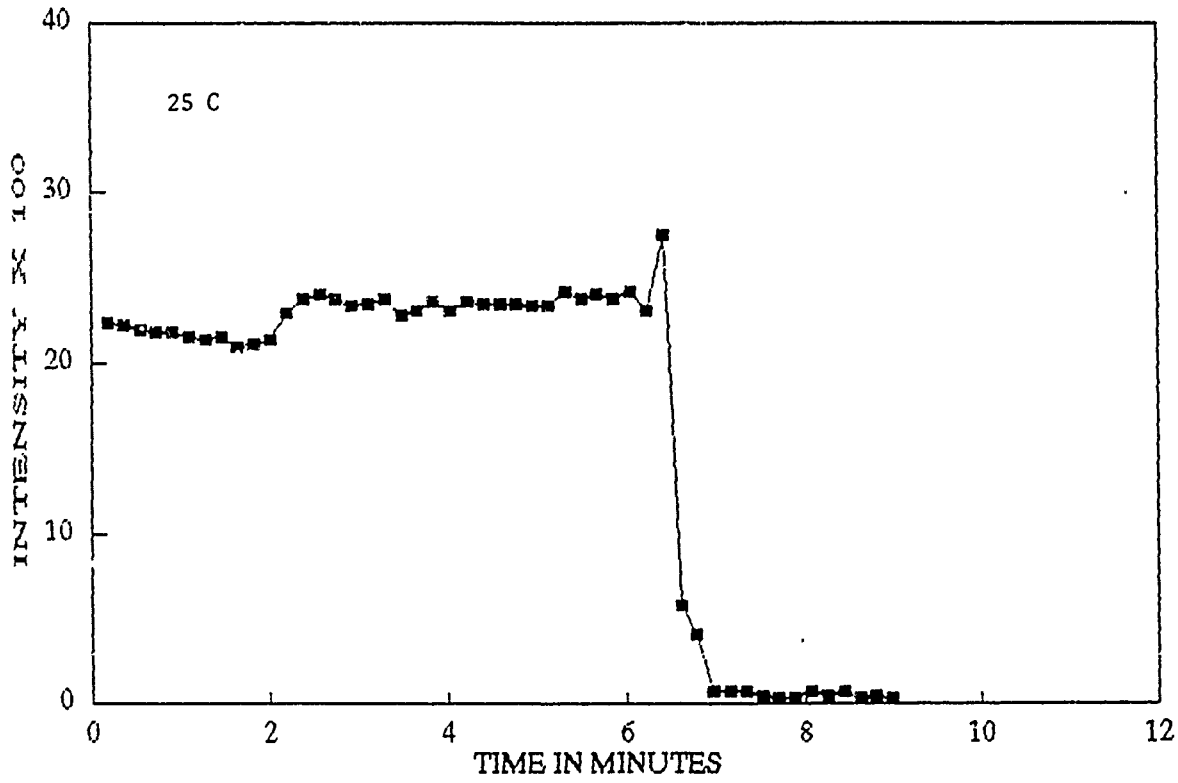
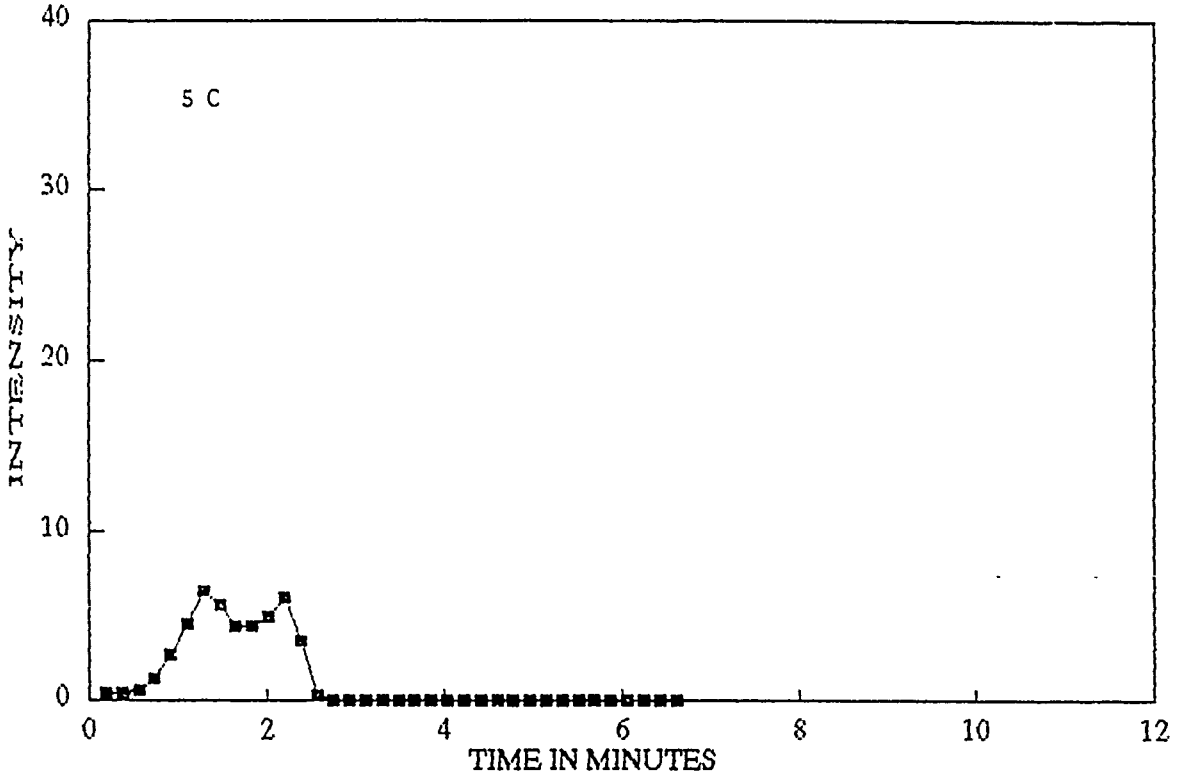


FIGURE 28 RUTHENIUM (III) CHLORIDE



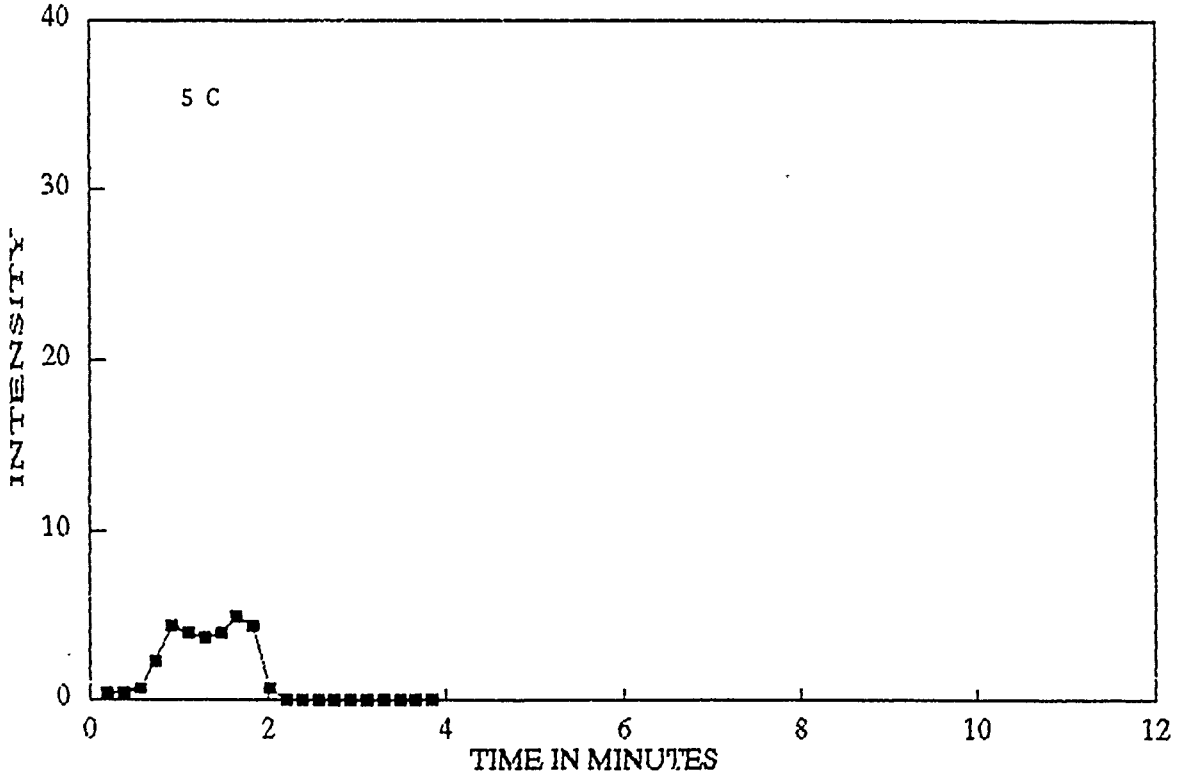
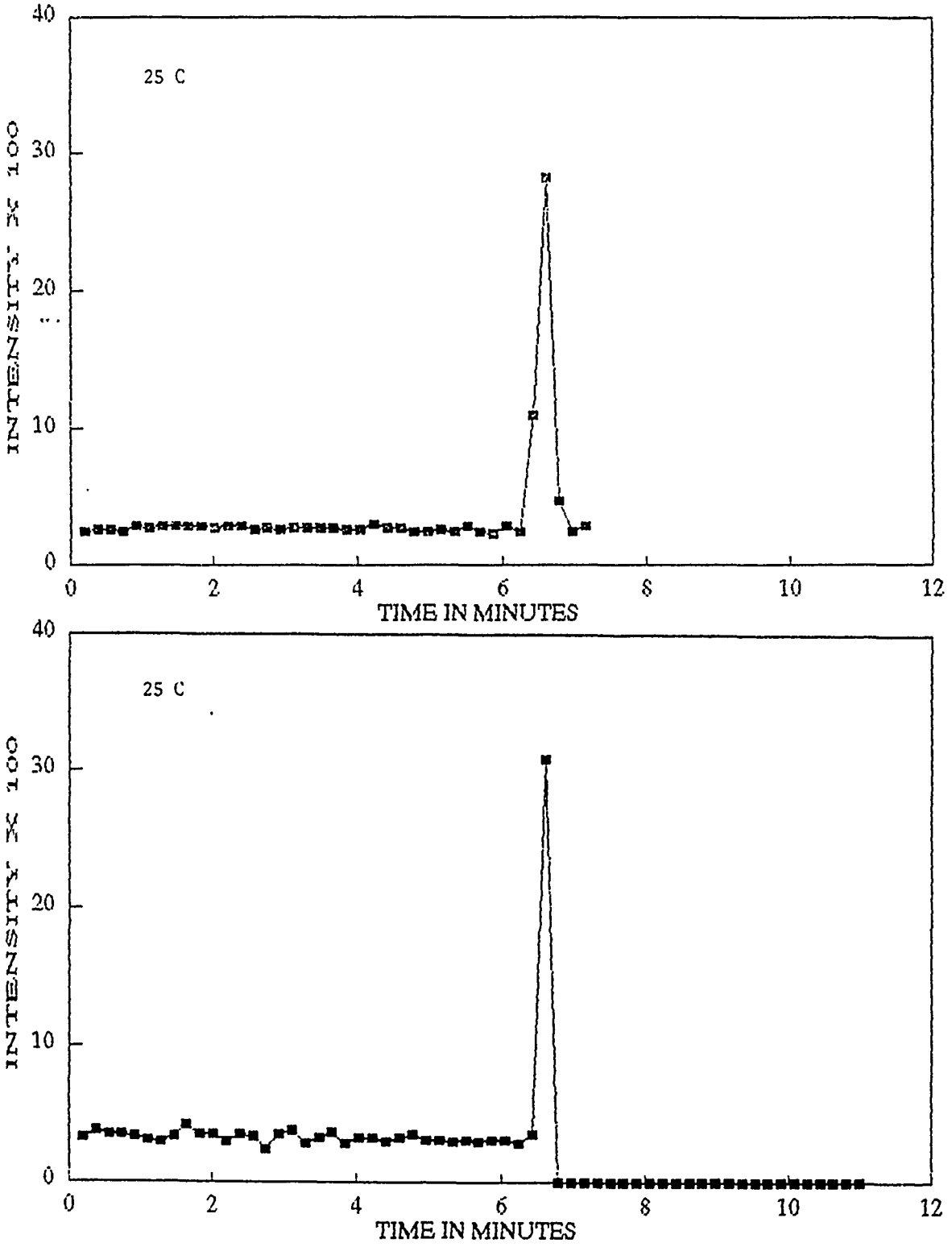


FIGURE 29 RUTHENIUM (III) CHLORIDE



not parallel its temperature sensitivity relative to the others. In view of the small size of this data set, one should have some caution since the method is very sensitive to any variation of geometry; we were very careful in attempting to control geometry, but this factor needs to be more carefully assessed. the trailing edge effects are also interesting. If scandium(III) dosimeters are intrinsically more sensitive to acoustic fields than to heating, they might be used to detect shock wave effects in HPM targets.

4.4. Nmr Studies

A substantial amount of time was spent attempting to detect nmr CIDNP phenomena (chemically induced dynamic nuclear polarization effects) in actively chemiluminescing reaction mixtures containing luminol. There was simply no evidence of this phenomenon. Since these effects depend on pair-wise interactions of reaction intermediate free radicals (8), the absence of a CIDNP effect suggests that luminol radicals, once formed, decompose so rapidly that there is no time to escape from the cage to allow a pairwise interaction, or perhaps the polarization is carried away by nitrogen. One major impediment in this project has been a loss of the nitrogen-15 capability; the broadband probe is still in repair status at this writing. Nitrogen-15 is a difficult isotope, but the CIDNP effect can, under some conditions, greatly enhance the signal strength.

4.5. Other Observations

Some of the dosimeters were sensitive to dissolved carbon

dioxide. Dosimeters of cobalt, vanadium and especially silver (as the fluoride) were substantially sensitive to this dissolved gas. The silver fluoride dosimeter produced a five-fold luminosity increase on bubbling the gas through it. A weaker dependence was seen for the systems of copper, ruthenium, titanium, scandium and silver (as the perchlorate). The iron, zirconium, niobium and nickel dosimeters are apparently indifferent to dissolved carbon dioxide .

5. RECOMMENDATIONS

This investigation has shown that aqueous chemiluminescent dosimeters based on dilute alkaline luminol/hydrogen peroxide/hydrated metal oxide mixtures are suitable for measuring dose distributions in target phantoms. The best dosimeters were those using the metal ions copper(II), cobalt(II) and (III), iron(III) and nickel(II). Ruthenium(III) proved to be the most active catalyst, but fresh dosimeters have a low temperature dependence; however, time ran out before its properties could be examined in aged peroxide solutions. A continuation of the search for temperature-sensitive catalysts using the microwave heating method is recommended.

These dosimeters were also sensitive to acoustic irradiations. The scandiium(III) dosimeter produced the largest acoustic I/Io, which did not correlate with its relatively low temperature sensitivity. A following study should re-examine the sonochemical effects reported here, and if these are confirmed, scandium(III) may be useful for distinguishing between acoustic and heat deposition effects.

The P.I. would like to continue this research project. With the

existing stockpile of metal salts and with instrumentation now available in the lab (especially with modifications which are appropriate for this kind of problem), a continuation of the work would not incur great expense and could be supported by local organized research funds. It is recommended that students enrolling in the research for credit course (CHEM 499X) be assigned to this project. The results from such continuing studies would be reported directly to the interested agency at Brooks AFB (USAFSAM/RZP).

6. REFERENCES CITED

(1) J.R. Wright, "Chemiluminescent Probes Based on Luminol and Luminol Derivatives", Final Report, 1988 USAF-UES Summer Faculty Research Program (USAFSAM/RZP), Brooks AFB, San Antonio, Texas.

(2) K. Van Dyke, ed., Bioluminescence and Chemiluminescence: Instruments and Applications , Vol's I and II, CRC Press, Inc., Boca Raton, Florida, 1985.

(3) E. H. White and D. F. Roswell, "The Chemiluminescence of Organic Hydrazides", Accts. Chem. Res., 3 , p54(1970); M. A. DeLuca and W. D. McElroy, Methods in Enzymology, 133 , 1986.

(4) A. M. Michelson, J. Maral and C. Mony, "Carbonate Anions; Effects on the Oxidation of Luminol, Oxidative Hemolysis, Gamma Irradiation and The Reaction of Activated Oxygen Species with Enzymes Containing Various Active Centers, Biochimie, 65 , p95(1983).

(5) H. K. Florig, "The Future Battlefield: A Blast of Gigawatts?", IEEE Spectrum , March 1988, p54; C. Chou, A. Guy and R. Galambos, "Auditory Perception of Radiofrequency Electromagnetic Fields", J. Acoust. Soc. Amer., 71 , p1321(1982).

(6) R. F. Abdulla, "Ultrasound in Organic Synthesis", Aldrichimica Acta, 21 , p31(1988), and references cited therein.

(7) J. L. Kiel, L. S. Wong and D. N. Erwin, "Metabolic Effects of Microwave Radiation and Convection Heating on Human Mononuclear Leukocytes", Physiol. Chem. Phys. and Medical NMR, 18 , p181(1986), and references cited therein.

(8) R. Kaptein, "Chemically Induced Dynamic Nuclear Polarization. VIII. Spin Dynamics and Diffusion of Radical Pairs" J. Amer. Chem. Soc., 94 , p6251(1972).

Appendices can be obtained from
Universal Energy Systems, Inc.

USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

ENHANCEMENTS TO PC-MAINFRAME INTERFACE FOR DATA ENTRY

Prepared by:	David R. Cecil, Ph.D.
Academic Rank:	Professor
Department and University:	Mathematics Department Texas A & I University
Research Location:	Clinical Investigation Directorate /WHMC, Lackland AFB, San Antonio TX 78235 & Department of Mathematics, Texas A & I University, Kingsville, TX 78363
USAF Researcher:	Cliff Butzin, Ph.D.
Date:	15 December 1989
Contract No:	F49620-88-C-0053/SB5881-0378

ENHANCEMENTS TO PC - MAINFRAME INTERFACE FOR DATA ENTRY

by

David R. Cecil

ABSTRACT

The CALC/CALCHD microcomputer based data entry system developed during my summer 1988 SFRP appointment at Wilford Hall USAF Medical Center was enhanced and enlarged from a new-data-only system to a comprehensive system allowing merging and updating of old data as well as adding new data in any of three possible ways. The new system, named ALL, has provisions for: (1.) data entry to be submitted to Statistical Consulting, (2.) Ridit Analysis and Ridit Analysis graphing, and (3.) means/standard deviations analysis of the data with automatic bar graph generation. All output is saved to floppy diskette for archival purposes of the researcher using the system. Output can also by saved to hard disk and/or a time dated printed-copy of the raw data (or in options (2.) and (3.) the raw or analyzed data or graphs) can be produced.

Using the hardware talents of Mr. Oscar Valdez, a graduate student in Electrical Engineering at Texas A & I University hired under this grant, an interface was devised to enable data from some monitoring equipment to be entered into the ALL system of programs without the necessity of manual input. The software handshaking between the string-generated data of the monitor output and the spreadsheet required data entry form was also developed.

Acknowledgments

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Additional thanks go to Universal Energy Systems for their help in the administrative and directional aspects of this program.

I am very appreciative of the entire staff of Clinical Investigation Directorate at Wilford Hall Medical Center. They are all to be commended for their friendliness and help during my work there. In particular CID's director, Col. John H. Cissik, provided me with encouragement and an excellent working atmosphere while I was there. Dr. Cliff Butzin, chief of the Research Consultation Division of CID, provided stimulating discussions about, support for, and direction to, the project. Captain Patterson and Major Bradley were a pleasure to work with, as was Dr. Bill Ehler, who additionally always provided delightful luncheon conversations.

At Dunn Dentistry Center, Colonel Burgess suggested several modifications to the CALCHD program that added greatly to the end product ALL system.

I. INTRODUCTION:

My research interests over the past decade have been in the applied mathematics areas of combinatorics, numerical methods, graphics, and probability. With the advent of micro-computers (PCs) I have concentrated on developing computer software to assist me in these areas. I have also become aware of the needs of first-time computer users and have published a book in the area of debugging computer programs. This year my book "Debugging BASIC Programs" was translated into Spanish by Paraninfo, SA of Madrid, Spain and is selling there under the title "Depuracion de Programas en BASIC".

The Clinical Investigation Directorate at Wilford Hall Medical Center, located at Lackland AFB conducts research in a large number of areas. Of mayor interest to me were data processing/analysis, and simulation/modeling.

I was most interested in ascertaining how much micro-computers were involved in the data process/analysis area since PC usage frees up mainframe time and memory, and I was curious as to whether researchers were being encouraged to learn to use microcomputers since PCs can be, potentially, of so much use to them. Finally I wanted to see if simulation/modeling using microcomputers was desirable &/or feasible, since large scale processes tax the resources of even large computers.

II. OBJECTIVES OF THE RESEARCH EFFORT:

The preliminary goals were three-fold: (1.) to create a merge/update facility for the CALC/CALCHD programs developed during my summer 1988 SFRP appointment at Wilford Hall USAF Medical Center, Lackland AFB; (2.) to construct a graphics component for the CALC/CALCHD programs and develop informational material on graphs; and (3.) to design a

hardware/software package to enable data to be entered directly into the CALC/CALCHD programs without the necessity of manual input.

Additional goals added during the course of the research effort were (a. (for Col. John Burgess, Chief of Dental Research, Department of General Dentistry, the Dunn Dentistry Center, Lackland AFB)) to include a Ridit Analysis data entry, computation, and automatic graphing of results option to the CALC/CALCHD package of programs, and (b. (for the 1989 Summer Faculty Research Program at CID Analytic Laboratories of University of Miami, Department of Physical Chemistry professor J. Drost, and his student J. Rafferty)) to develop a file consolidation program to use in conjunction with CALCHD and the database features of CALCHD imported from SuperCalc. Also to devise a fast file selection-printing program with current date-time-labeling features for their use with the file consolidation program during fall 1989 while at the University of Miami.

III. LARGER DATA SETS AND A MERGE/UPDATE FACILITY:

The original design of CALC/CALCHD used a maximum data size configuration based on (1.) what was thought to be typical data set sizes, and (2.) constraints imposed by using the programs on microcomputers, such as the Zenith Z100 Series, that have no hard disk drive storage capabilities. As a consequence, using CALC/CALCD on large data sets necessitated having to break the data into a multitude of smaller files each to be run through CALC/CALCHD and then the totality of these files being tediously merged while in the VAX environment.

Since the Z100 Series of microcomputers is rapidly being replaced by the Zenith Z200 Series and since the Air Force is also acquiring larger and faster AT as well as 386 type microcomputers, it was decided that the CALC package (for

microcomputers without hard disk drives) would stay unchanged while the CALCHD package (HD for hard disk) would be modified to provide for larger data set sizes.

This modification was accomplished through design changes to the SuperCalc portion of CALCHD. In addition various overlay informational screen segments (SCREEN1.OVL and SCREEN2.OVL) were revised to inform the user of the new larger sizes CALCHD can now handle.

In addition a merge/update facility was designed to operate in conjunction with CALCHD, since even for moderately sized data, many reseachers have only small chunks of time available for data entry and prefer not to have to enter all their data at one time. Some C language file and string manipulation programs, in particular OLDNEW and OLDNEW1, were developed to facilitate and speed up the merging process and to provide for saving the resultant files to either floppy diskette or hard disk drive. An error detection/correction routine was added in case file names were found to be non-existent or entered incorrectly.

IV. THE RIDIT ANALYSIS PACKAGE:

Ridit Analysis is widely used in many scientific studies in the biological and behavioral sciences where a response variable falls in the borderline between dichotomous classifications and refined measurement systems. One of the early users of the technique was the Cornell Automotive Crash Injury Research Program. Dental researchers use Ridit Analysis in studies such as occlusal isotope penetration and marginal fracture vs mechanical properties of Amalgam.

Dunn Dental Research Center at Lackland asked if I would add a Ridit Analysis procedure to the CALCHD program. Such a procedure was developed using SuperCalc3 as a nucleus. One reference set and a maximum of six other data sets (each

with no more than twenty five (25) observations) can be compared. Ridit means, standard deviations, and t-values (with corresponding degrees of freedom indicated) are computed. The procedure is menu driven with instructions on usage appearing on the screen. Raw data or calculated t-values, etc. can be saved.

If desired, a presentation quality graph of the results can be exhibited on the screen and 35mm slides then obtained. The screen graph is drawn automatically in Harvard Graphics using the new macro commands (only available in the latest version and requiring 640K memory to run). Additional macros and graphical templates (CHARTD and CHARTE) were developed to enable the graphs of the Ridit Analysis to appear without operator assistance. The operator has complete freedom to modify such graphs (e.g. for different font selection, other colors, etc.) at any time after the first graph appears.

V. THE ALL PROGRAM AND MENU SYSTEM PACKAGE:

Major Bradley and Captain Patterson of the Analytic Laboratories at CID heard of the CALCHD/HARVARD GRAPHICS package I had created for Col. Burgess in Dentistry and asked if I would install a similar system on two of their microcomputers. For their purposes the Ridit Analysis procedure was modified so as to calculate means and standard deviations for a maximum of eight (8) variables, each having a maximum of twenty (20) observations. These limits were built in to coordinate with the graphical limitations of Harvard Graphics (eight (8) graphs maximum).

The resulting system was installed as-is on their fast Dell microcomputer. The other PC in their laboratories will be updated to 640K + memory some time in the future, but for now the automatic graph construction feature cannot be utilized, although user generated graphs can be created

manually. By changing the CONFIG.SYS file both older and newer versions of Harvard Graphics run (but without macro graphing capabilities) on this 512K PC.

Thinking that other researchers might also be interested in the graphical features of CALCHD/HARVARD GRAPHICS, I decided to create a new menu driven comprehensive system with a variety of features that could be mixed and matched to the specific desires and needs of a researcher.

The new system is called ALL. The main menu consists of three options: (A.) data entry to be submitted to Statistical Consulting, (B.) means/standard deviations and, if desired, a graph of the means, and (C.) Ridit Analysis and, if desired, a graph of the Ridit means.

Help is always only a keystroke away. New data sets can be entered, or old data sets (previously saved to diskette) can be entered and new data added. Raw data and/or (in options B. and C.) calculated data is saved to hard disk and, optionally, to diskette with either a default file name or a user-defined file name. Hard copy of raw and/or (for B. and C.) calculated data can be obtained with current time and date appearing along with any file description desired. (This date-time-labeling procedure came about as a result of the file consolidation request of Dr. Drost, discussed in the next section.)

VI. THE FILE CONSOLIDATION PROGRAMS:

Dr. Drost of the University of Miami requested a file consolidation program to use in conjunction with my CALCHD system. He, along with his student John Rafferty, compiled an enormous number of data sets during his summer 1989 Faculty Research Program at CID. It was important for each data set to be run using my CALCHD system, but also desirable to be able to extract from each data set only

those observations that satisfied certain constraints. Dr. Drost wondered if CALCHD could somehow automatically create data sets consisting only of those observations satisfying the constraints; and then combine selected ones of these data sets into one large consolidated file.

Using the database features of CALCHD imported from SuperCalc3, I instructed the two gentlemen in the procedures needed to be taken to select only data satisfying certain desired characteristics. This can't be programmed since the number of possible sets of characteristics that the two men might be interested in is too large for any kind of manageable menuing system that might be built into a computer program.

I did write a series of program files (FIRST, CLEANUP, and DOIT) that, taken together, allow a researcher to selectively view, &/or choose &/or not include data sets for consolidation into a new file. This series of programs can also be used to selectively create a new file from old files; whether the original files were data sets, or program files, or strings, or any combination of these.

VII. THE HARDWARE/SOFTWARE DIRECT ENTRY PACKAGE:

In order to develop the electronics portion of a hardware/software system capable of directly entering data into CALCHD without the necessity of manual input, I needed someone with a E.E. microprocessor backgrond. Part of the money for this grant was used to employ (for one-quarter time during the Fall 1989 Semester) an Electrical Engineering graduate student at Texas A & I University. The faculty of the Department of Electrical Engineering and Computer Science recommended Mr. Oscar Valdez, one of their outstanding students. Mr. Valdez was subsequently hired to study, design, and implement an interface between the patient monitoring equipment in Veterinary Surgery at CID

and an AT clone microcomputer.

Mr. Valdez conferred with Sgt. Jarred Ross and Dr. William Ehler at CID during June 1989, before his employment began, in order to assimilate the product requirements of the monitoring equipment and of the microcomputer chosen to be used (a Zenith 248 AT class machine). During the 1989 fall semester Mr. Valdez postulated various design configuration theories. Best interface design options, whether to construct new hardware or enhance commercially available equipment, and what language to handshake the hardware interface to the software CALCHD program system were debated.

Progress was not as swift as I had hoped for. Mr. Valdez also had, unknown to me until November, a grant from NASA for the fall semester.

It was finally decided to purchase a commercially available interface and design around it. This interface is not a stock item but, rather, is built to specifications when ordered. Unfortunately this interface will not be ready until late January 1989. At that time Mr. Valdez and I will, at no further cost to the grant, install the interface in Veterinary Surgery at CID, Lackland AFB. We will have the automatic data entry system working as soon as possible. I sincerely regret the delay in finishing this part of my grant.

VIII. RECOMMENDATIONS:

The enhanced version of my CALC/CALCHD programs, now called ALL, is being used at CID and at Dunn Dentistry. Other researchers will probably also use ALL since it will be available through Dr. Butzin in Statistical Consulting at CID.

Follow-up research to this grant might take a variety of paths. I have three areas in particular that I think would produce interesting and useful reslts.

(1.) Coordinate the use of the ALL system with some microcomputer based statistical package such as Microstat (which is available at several sites at Wilford Hall and Lackland) or SOLO. This would be useful for the generation of elementary statistics from the data sets entered into ALL and/or CALCHD. I do not suggest that any elaborate microcomputer software package be used with ALL/CALCHD since more involved statistical analysis is best left to Statistical Consulting and to the capabilities of the SPSS package on the VAX.

(2.) If future data sets exceed the current capacity constraints imposed on ALL by SuperCalc3 (26 variables by 500 patients, or 66 variables by 164 patients, or 127 variables by 30 patients) then a revision of ALL using version 5 of SuperCalc or using one of the more powerful spreadsheet programs such as Excel or Quattro Pro would be in order. These other programs allow for the use of extended and/or expanded memory.

(3.) The use of a handy scanner, in conjunction with optical character recognition software, might provide another method of automatic data entry. This is not feasible for the monitoring requirements mentioned in section VII., since accurate timed data entry is required there. But for hand-entered numerical data sets submitted to Statistical Consulting optical scanning would save an enormous amount of data entry time. Unfortunately, commercially available optical scanning software can only recognize, with approximately 95% accuracy, non-proportional typewritten text. Software would need to be developed that could recognize hand-scribed numbers.

REFERENCES

Manuals:

1. Harvard Graphics, Version 2.0, User's Guide, Mountain View, California 94039-7210, Software Publishing Corporation, 1987.

2. SuperCalc3, Release 2.1, User's Guide and Reference Manual, Benton Harbor, Michigan 49022, Heath Company, Zenith Systems, 1985.

Books:

3. Somerson, Paul, PC-DOS Power Tools, Techniques, Tricks, & Utilities, New York, NY 10103, Bantam Computer Books, 1988.

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

EFFECT OF HYPEROXIA ON THE PERMEABILITY OF THE BLOOD-BRAIN BARRIER IN SEVERAL LABORATORY SPECIES AND ON ORGANOTYPIC EXPLANT TISSUE CULTURES OF HAMSTER BRAIN

Prepared by: Donald W. Welch, Ph.D

Academic Rank: Research Scientist

Research Location: Texas A&M University

Hyperbaric Laboratory

College Station, Texas 77840

Contract No: F49620-85-C-0013/SB5851-0360

I. SUMMARY

The investigations outlined in this proposal have been designed to expand on the studies conducted at the Wilford Hall Medical Center in cooperation with the Clinical Investigative Facility (CIF). This project successfully established the surgical and neuropathologic techniques required to effectively evaluate alterations in blood-brain barrier (BBB) permeability using micromolecular tracers. However, due to the high degree of morphologic integrity of the rat BBB the investigators were unable to establish a dose/response relationship between increased partial pressures of oxygen and increased BBB permeability in that animal. Preliminary observations in other laboratory species (rabbit and guinea pig) indicated that alterations in BBB permeability subsequent to increased inspired partial pressures of oxygen did occur and these observations provided the basis for the present project.

The objectives of these studies were three-fold:

Objective One: To determine if a dose/response relationship between oxygen partial pressure and increased BBB permeability could be effectively established in several laboratory animal species.

Objective Two: To comparatively examine the effect of increased partial pressures of oxygen on BBB permeability in several laboratory species-rabbit, mouse, rat and guinea pig.

Objective Three: To study the direct effects of increased partial pressures of oxygen on the cellular components of an *in vitro* organotypic explant tissue culture system of hamster brain.

The animals were divided into one control group and three experimental groups with each experimental group having three different exposure times. All experimental groups breathed 100% oxygen for the specified times and pressures as follows:

Oxygen Exposures

Pressure (ATA)	Times		
1.0 (760 mmHg)	6.0 hrs	12.0 hrs	24.0 hrs
2.0 (1520 mmHg)	60 min	120 min	180 min
3.0 (2280 mmHg)	30 min	60 min	120 min

Within each of the species studied air controls were subjected to the same procedures as the experimental groups with the exception of being exposed to increased partial pressures of oxygen.

The guinea pig and mouse-rat were concluded to be the most resistant with the guinea pig having the most clear cut point (prominence of dye staining of tissue) at which breakdown occurred. The rabbit was concluded to be the most susceptible of the group since all groups subjected to hyperbaric oxygen had some evidence of BBB disruption as indicated by dye penetration. In addition, the effect of the different hyperbaric pressures used were quite difficult to evaluate in this species because of the lack of clear-cut differences between the control and experimental CNS tissue of animals examined.

Although only limited numbers of cultures were used in this preliminary study, it appeared that the susceptibility of myelin to 3.0 ATA was equal to, or even possibly greater than, that detected following exposure to 6.0 ATA of pressure and was detectable after 24 hours. In contrast, there was essentially no difference between general explant growth and neurons (cell bodies and processes) after 24 hours. After 72 hours the degeneration of myelin was essentially the same and explants were slightly more granular than after 24 hours. In contrast to the results after 24 hours, at 72 hours neurons had degeneration of cell bodies and processes. Further, compared to cultures exposed to 6.0 ATA of pressure to those exposed to 3.0 ATA for 72 hours, general explant integrity was much less affected.

Reasons for the differences in susceptibility of the BBB to hyperbaric oxygen pressures among the different species could not be explained based on the work conducted to date. However, further investigation of the mechanisms involved in the differences encountered could contribute valuable insight into the function of the BBB, and is considered to be warranted. In addition, the exact mechanisms by which pressure and oxygen are involved in causing degeneration of the glial tissue, myelin sheaths and neurons in tissue culture explants of hamster brain could not be determined, but further pursuit of their identification would represent a very fertile area for future investigation.

II. INTRODUCTION

The studies outlined in this report were designed to expand on investigations conducted at the Wilford Hall Medical Center in cooperation with the Clinical Investigative Facility (CIF). The overall objective of the initial research project was to examine, on a preliminary basis, the effect of increased partial pressures of oxygen on the blood-brain barrier (BBB) in the rat. Preliminary observations, previously carried out, noted an increase in the permeability of the BBB with HBO in the rabbit (Ehler, unpublished observations, Chambi, unpublished observation).

In brief review of this project the investigators successfully established the surgical procedures required to chronically implant jugular catheters in the rat so that multiple samplings and injections could be performed (Addendum A). Also developed were the neuropathologic techniques to quickly and accurately evaluate alterations in BBB permeability in a large number of animals (Addendum B). The final objective, which was to establish a dose/response relationship between oxygen partial pressure and an increase in BBB permeability in the rat, was not achieved. After being exposed to 100% oxygen at 3.0 atmospheres absolute (ATA) for a total of 210 minutes no appreciable breakdown of the BBB was recorded with Trypan blue or the micromolecular tracer, sodium flourescein (NaFl).

Our preliminary studies revealed that a dose/response relationship could not be established because of the unfortunate choice of the rat as the experimental model. The majority of studies conducted to date utilizing rats as a BBB model system focussed on permeability changes induced by mechanical factors such as bubbles or emboli. However, none of the past investigations examined the effect of increased partial pressures of oxygen on the BBB of the rat. Due to the resistance of the rat BBB to oxygen, it appeared essential to examine on a very preliminary basis, the differences in sensitivity to hyperbaric oxygen (HBO)

between several other species of laboratory animals.

These preliminary observations indicated that in both rabbits and guinea pigs, there was an increase in BBB permeability when exposed to 100% oxygen at 3.0 ATA for a total of 90 minutes. There was a definite increase in permeability in the cerebrum that included the cortex, neostriatum, thalamus and periaqueductal areas. In addition, the cerebellar cortex had comparable staining. In the case of the guinea pig, it was found that animals exposed to 100% oxygen at 3.0 ATA for 30 minutes showed a definite increase in permeability in the cortex, while animals exposed to 100% oxygen at 2.0 ATA for 60 minutes, showed a minimal amount of staining.

These results suggested a significant difference in the susceptibility of the BBB between these species. It also demonstrated that the permeability of the BBB in the guinea pig could be altered by HBO and presented the possibility of establishing a dose/response relationship between oxygen partial pressure and BBB permeability.

Based on these preliminary observations the following objectives were proposed as a continuation and expansion of the Summer Faculty Research Program. The objectives of the proposed studies were three-fold:

Objective One: To determine if a dose/response relationship between oxygen partial pressure and increased BBB permeability could be effectively established in several laboratory animal species.

Objective Two: To comparatively examine the effect of increased partial pressures of oxygen on BBB permeability in several laboratory species-rabbit, mouse, rat and guinea pig.

Objective Three: To study the direct effects of increased partial pressures of oxygen on the cellular components of an *in vitro* organotypic explant tissue culture system of hamster brain.

III. BACKGROUND

Investigations which first revealed the presence of a diffusion barrier between the vascular compartment and the brain were first published by Ehrlich in 1885 (5). Ehrlich systemically injected vital dyes into laboratory animals and noted that the dyes were taken up by virtually all organs of the body with the exception of the brain. However, Ehrlich did not interpret his results in the context of a barrier, but did comment that the brain probably did not contain the necessary chemical constituents to absorb the dye. Approximately 30 years later Goldmann conducted additional experimentation which amplified Ehrlichs' work and subsequently hypothesized the existence of a blood-brain barrier (6). Over the period of the next 50 years several controversies developed as to the nature and location of the barrier. In the early 1960's the controversies were basically put to rest with the acceptance of two distinct barrier systems--the BBB and the blood-CSF barrier. The BBB is found at the majority of the capillaries and the blood-CSF barrier is localized at the small number of capillaries perfusing the choroid plexus and other circumventricular organs (2). Studies in the late 1960's established that the barrier phenomenon is comprised of the endothelial tight junctions and the paucity of pinocytosis or fenestrations in the CNS endothelial lining (3, 15).

With the establishment of the basic nature of the barrier, investigators began focusing attention on quantifying and characterizing the role played by BBB nutrient transport in the regulation of brain metabolism. These investigations expanded to studies aimed at determining the nature of compounds that are able and unable to pass through the barrier, characterization of disease processes which alter the barrier, selective alteration of the barrier, and a multitude of studies establishing the physical and chemical parameters of the BBB. Several of these investigations addressed barrier alterations under hyperbaric conditions. These included alterations induced by increased pressure (4), decom-

pression sickness (3,4), gas embolism (9, 11, 12), and increased partial pressures of oxygen (10, 14, 18). While the majority of the published studies indicate an increase in BBB permeability due to increased inspired partial pressure of oxygen there is still some controversy in the literature (1, 7). Standard medical thinking dictates that this "breakdown" would be detrimental to the patient or subject. However, due to the fact that thousands of hyperbaric oxygen treatments have been conducted without a noticeable detrimental effect, the nature of the "breakdown" must be questioned. Is the induced alteration specific or nonspecific? Is it selective or nonselective? Do extended periods of oxygenation under normobaric or hypobaric conditions (ie. four hours of denitrogenation required prior to an astronaut going extra-vehicular in space) alter the permeability of the BBB? Alterations of the BBB under the conditions of increased partial pressures of oxygen could have far-reaching implications in such areas as the treatment of neurological diseases, drug delivery, and in the treatment of acquired immunodeficiency syndrome (AIDS). The U.S. Air Force has become a recognized leader in the use of HBO therapy for a variety of medical disorders, as well as extensively using oxygen to enhance operational effectiveness. It is easily conceivable that the characterization of BBB alterations due to increased partial pressures of oxygen would accelerate the utilization of HBO therapy in returning injured personnel to active duty status as well as increasing operational effectiveness.

IV. MATERIALS AND METHODS

a. Micromolecular Dye Studies. In this series of experiments the micromolecular dye sodium flourscein (NaFl) was used as an indicator of BBB permeability. These experiments were designed to determine the feasibility of establishing a dose/response relationship between oxygen partial pressure, exposure time in the mouse, guinea pig and rat (Objective One). These studies also allowed for the comparative evaluation of BBB alterations between several laboratory species (Objective Two).

The animals were divided into one control group and three experimental groups with the experimental groups having three different exposure times. All experimental groups will breathe 100% oxygen for the specified time and pressure.

All animals were anesthetized with Nembutal (30 mg/Kg, IP) before any surgery was performed such as placement of indwelling catheters. At the time of sacrifice, they were again anesthetized as before and the chest was opened. A canula was placed in the left ventricle and an opening was made in the right ventricle. A 10% buffered formalin infusion was begun in the left ventricle, replacing the blood in the entire body. Care was taken to avoid a high hydrostatic pressure to avoid a mechanical alteration of the anatomical structure of the brain capillaries. Care also was taken to assure that the brain received adequate formalin perfusion, not only to assure fixation, but to wash out any fluorescein which might remain in the blood within the capillaries. Preliminary studies showed that this procedure did not wash out fluorescein which already had crossed the BBB.

The skull then was opened and the brain carefully removed intact except for the pituitary gland which remained in the sella turcica. Each brain then was sectioned coronally into 0.5 to 1.0 cm slices and placed in order on a strip of black plexiglass. They were then placed under the UV light and were photogra-

phed with color film in a dark room.

A veterinary neuroanatomist evaluated each slide, blinded as to protocol, using a grading system that involved the assigning of a numerical value of 0.5 to 4.0 to the degree of staining intensity detected in the various parts of the brain examined. The assigned scores were totaled and a final score was given to each animal. The degree of dye penetration into the various parts of the cerebral cortex was recorded separately so that a more cortical evaluation of this part of the central nervous system could be made. This evaluation involved determining whether penetration involved: 1) meninges and underlying cortex, 2) full width cortex (deep involvement) and, 3) whether primary cortical involvement was dorsally or ventrally located in the cerebral hemisphere.

Areas of the brain that were evaluated included the following: meninges and penetration of underlying areas (eg. olfactory bulb, cerebral cortex and ventral brain); frontal lobe (including cortex); parietal lobe (including cortex); occipital lobe (including cortex); temporal lobe (including cortex); caudate nucleus; septum; pre-optic area; thalamus; hypothalamus; mammillary body; hippocampus; infundibulum; anterior colliculus; posterior colliculus; pons; cerebellum and medulla oblongata.

In addition to the above described flourescein dye experiments, four animals subjected to inspired oxygen partial pressures causing significant alterations in BBB permeability were euthanatized immediatedly following exposure and perfused with 10% buffered formalin for microscopic examination. The brains were sectioned coronally at representative levels, stained with hematoxylin and eosin stain and examined using conventional light microscopy for lesions.

When necessary animals were surgically instrumented with an indwelling jugular vein catheter according to the procedure detailed in Addendum A. All animals were allowed a recovery time of at least 24-48 hours after surgery before being included in an experimental group. Following recovery from surgery the

animals were subjected to the appropriate hyperbaric oxygen exposure as follows:

Oxygen Exposures

Pressure (ATA)	Times		
1.0 (760 mmHg)	6.0 hrs	12.0 hrs	24.0 hrs
2.0 (1520 mmHg)	60 min	120 min	180 min
3.0 (2280 mmHg)	30 min	60 min	120 min

Within each of the species studied, air controls were subjected to the same procedures as the experimental groups with the exception of being exposed to increased partial pressures of oxygen. In addition, extensive controls were run to determine if the surgical manipulations used to implant the jugular catheters resulted in any significant alterations in BBB permeability.

Immediately upon the decompression of all animals, NaFl (10%) was injected (1.0 ml/kg body weight) via the indwelling catheter or the femoral vein. Upon completion of the infusion the tracer was allowed to circulate for 15 minutes. The animals were then be anesthetized, the chest cavity opened, and the dye flushed from the vascular compartment by intracardic perfusion using the procedure described above.

b. Explant Tissue Culture Studies. Neonatal offspring from a colony of disease-free Syrian hamsters, supplied by Charles River Breeding Laboratories and maintained locally in isolation, were used for preparation of explant cultures. The basic methods for culture preparation and examination was the same as that previously described (16, 17). Briefly, the procedure involved growing explants (1 mm3) prepared from neonatal hamster brain on coverslips. Mature cultural components include neurons with myelin sheaths and functional synapses, oligodendroglia, astrocytes and microglia. Cultures were examined by light microscopy in the living state and following special staining or impregnation

for myelin sheaths, neurons, and neuroglia. Cultures subjected to hyperbaric conditions were maintained in tubes (with cloth filter caps) or in plastic tissue chambers which allowed for the uniform equilibration of the pressure and the gas mixture being employed.

The elevated carbon dioxide pressure was employed to maintain optimal medium pH and tissue culture growth.

The control cultures grown at 1.0 ATA were compared to cultures grown at 3.0 ATA and 6.0 ATA according to the following protocol:

Oxygen Exposures

Pressure	Oxygen	Carbon Dioxide
ATA (mmHg)	%, (mmHg)	%, (mmHg)
1.0, (760)*	21, (160)	2-5, (15-38)
3.0, (2280)	98-95, (2234-2166)	2-5, (46-114)
6.0, (4560)	98-95, (4469-4332)	2-5, (91-228)

* The remainder of this mixture was comprised of nitrogen.

Hyperbaric Oxygen Administration. Hyperbaric oxygen was administered in a steel chamber which could treat up to 16 animals simultaneously. It contained an environmental control system which removed carbon dioxide, and allowed the delivery of any partial pressure of oxygen desired.

Animals were placed in the chamber and the door sealed. The chamber then was pressurized with the desired gas mixture and flushed for 5 minutes to assure removal of the air which was in the chamber when the door was sealed. Temperature was maintained at 22 degrees C, throughout the exposure.

When tissue cultures were exposed, the same basic procedure was followed, except that the chamber temperature was elevated to 37 degrees C, which was the standard temperature for tissue culture. Care also was taken to avoid dehydra-

tion of the cultures by maintaining a saturated condition in the chamber.

V. RESULTS

A. Animals exposed to hyperbaric oxygen pressures and subsequently given fluorescein dye intravenously. The scoring of the photographs taken of the brains of the experimental animals examined, and the hyperbaric pressures to which they were subjected, are given in the following tables.

TABLE 2. LABORATORY SPECIES-MOUSE

Cerebral Cortical Evaluation

Animal Number	Score	Meninges and Underlying Cortex Affected	Deep Cortex Affected	Dorsal Ventral Cortex Affected	Hyperbaric Oxygen Exposure
219	19.25	+	-	D,V	Control
220	10.25	+	-	D	Control
221	25.25	+	-	D,V	Control
222	8.50	+,-	-	D	Control
217	12.00	+		D,V	1.0 ATA, 6.0 hrs.
217A	13.25	+	+	D,V	1.0 ATA, 6.0 hrs.
218	6.25	+	-	D	1.0 ATA, 6.0 hrs.
215	11.75	+,-	-	D,V	1.0 ATA, 12 hrs.
216	13.50	+	-	D,V	1.0 ATA, 12 hrs.
213	19.00	+	+	D,V	1.0 ATA, 24 hrs.
214	19.00	+	-	D,V	1.0 ATA, 24 hrs.
203	23.50	+	+	D	2.0 ATA, 60 min.
204	12.00	+	-	D,V	2.0 ATA, 60 min.
201	9.50	+	-	D	2.0 ATA, 120 min.
202	5.50	+	-	D	2.0 ATA, 120 min.
205	23.50	+	-	D,V	2.0 ATA, 180 min.
211	36.00	+	+	D,V	3.0 ATA, 30 min.
212	32.25	+	-	D,V	3.0 ATA, 30 min.
210	43.50	+	+	D,V	3.0 ATA, 30 min.

D - Dorsal cortex most prominently affected
V - Ventral cortex most prominently affected

TABLE 3. LABORATORY SPECIES-GUINEA PIG

Cerebral Cortical Evaluation

Animal Number	Score	Meninges and Underlying Cortex Affected	Deep Cortex Affected	Dorsal Ventral Cortex Affected	Hyperbaric Oxygen Exposure
304	8.00	+	-	D,V	Control
313	7.00	+	-	D	Control
318	8.75	+	-	D	1.0 ATA, 6.0 hrs.
319	10.75	+	-	D,V	1.0 ATA, 6.0 hrs.
316	10.50	-	-	-	1.0 ATA, 12 hrs.
317	9.00	-	-	-	1.0 ATA, 12 hrs.
320	22.00	+	+	D	1.0 ATA, 24 hrs.
321	41.75	+	+	D	1.0 ATA, 24 hrs.
309	15.00	-	-	-	2.0 ATA, 60 min.
310	11.00	+,-	-	V	2.0 ATA, 60 min.
311	13.75	+	-	D,V	2.0 ATA, 120 min.
312	7.00	+	-	V	2.0 ATA, 120 min.
314	20.25	+	+	D,V	2.0 ATA, 180 min.
315	1.00	-	-	-	2.0 ATA, 180 min.
301	4.75	+	-	D,V	3.0 ATA, 30 min.
302	6.75	+,-	-	D	3.0 ATA, 30 min.
307	2.00	-	-	-	3.0 ATA, 60 min.
308	4.00	-	-	-	3.0 ATA, 60 min.
305	10.75	+,-	-	D	3.0 ATA, 120 min.
306	8.75	+,-	-	D	3.0 ATA, 120 min.

D - Dorsal cortex most prominently affected
V - Ventral cortex most prominently affected

TABLE 4. LABORATORY SPECIES-RABBIT

Cerebral Cortical Evaluation

Animal Number	Score	Meninges and Underlying Cortex Affected	Deep Cortex Affected	Dorsal Ventral Cortex Affected	Hyperbaric Oxygen Exposure
403	51.10	-	+	D	Control
404	51.50	-	+	D	Control
415	66.78	-	+	D	1.0 ATA, 6.0 hrs.
416	54.82	-	+	D,V	1.0 ATA, 6.0 hrs.
413	31.50	-	+	D	1.0 ATA, 12 hrs.
414	58.50	-	+	D,V	1.0 ATA, 12 hrs.
417	45.00	-	+	D,V	1.0 ATA, 24 hrs.
418	49.00	-	+	D,V	1.0 ATA, 24 hrs.
411	49.75	-	+	D	2.0 ATA, 60 min.
412	33.50	-	+	D	2.0 ATA, 60 min.
401	25.00	-	+	D	2.0 ATA, 120 min.
402	48.25	-	+	D	2.0 ATA, 120 min.
419	45.85	-	+	D	2.0 ATA, 180 min.
420	48.50	-	+	D,V	2.0 ATA, 180 min.
405	38.70	-	+	D	3.0 ATA, 30 min.
406	39.93	-	+	D	3.0 ATA, 30 min.
407	47.00	-	+	D	3.0 ATA, 60 min.
408	60.75	-	+	D	3.0 ATA, 60 min.
409	50.00	-	+	D	3.0 ATA, 120 min.
410	74.37	-	+	D	3.0 ATA, 120 min.

D - Dorsal cortex most prominently affected
V - Ventral cortex most prominently affected

Controls were run with each of the species studied. These controls were maintained under the same conditions as the experimental animals and subjected to the identical experimental conditions prior to evaluation of BBB permeability. It was determined that within the experimental parameters there was no significant alteration within the control group when compared with the experimental groups. Should any alterations in BBB permeability be the result of conditions other than increased oxygen tension it was considered that these conditions were uniform among all experimental and control animals.

Surgical implantation of chronic jugular catheters was only used in the rat and it was determined that the procedure presented in Addendum A did not trigger any significant alterations in BBB permeability in this species.

Of the species studied, the most distinct differences were detected with the guinea pig. The majority of the brains of these animals were essentially negative with the few positive ones (animals numbers 314, 320, 321) being prominently stained. Areas of intense staining (not necessarily all areas present in each animal) included the deep cerebral cortex of the frontal, parietal, and occipital lobes; temporal cerebral cortex; thalamus; anterior colliculus; tegmentum; cerebellum and medulla oblongata.

The mouse and rabbit groups were the most difficult to evaluate. With the mouse species the majority had cortical penetration from the meninges with none of them having the degree of deep staining prominence that was seen in the positive guinea pigs and many of the rabbits. Also, staining with the mouse group was usually characterized by varying shades of green-yellow, while in the other two groups positive staining was more prominent, and often yellow, orange-yellow or dark brown in color.

The problem with the rabbit group was that they all had varying degrees of rather prominent staining and the scores reflected this result. It was concluded that some of the brain tissue (eg. various nuclei of the brain) may have autofluoresced although the common deep staining of the cerebral cortex of most animals was difficult to explain.

B. Animals exposed to hyperbaric oxygen pressures and subsequently examined microscopically for lesions. No lesions were detected in the animals subjected to partial pressures of oxygen resulting in BBB alterations and examined immediately following exposure by conventional light microscopy.

C. Effect of Hyperbaric Oxygen on Myelinated Explant Cultures of Hamster Cerebellum. Preliminary studies were made to determine the effect of different

hyperbaric oxygen pressures on myelinated explant cultures of hamster cerebellum in order to definitively determine the effects on the components of central nervous tissue by eliminating the various influences that would be present in the intact animal. The studies completed to date should be considered preliminary in nature, but at the same time should be considered of potential significance in that they definitely suggest that further investigative efforts are warranted.

The results of myelinated explant tissue cultures which were exposed to two different hyperbaric oxygen pressures (3.0 and 6.0 ATA) as detailed above under Materials and Methods, were as follows:

a. Cultures exposed to hyperbaric oxygen at 6 ATA. After 24 hours, living cultures exposed to hyperbaric oxygen were more granular, particularly in the explant proper, with minor differences in the outgrowth surrounding the explants when compared with the controls. There was also a significantly a greater degree of myelin degeneration in the hyperbaric oxygen exposed cultures than in the controls. Myelin lesions consisted of undulations and focal outpouc ng of the sheath. Also, in contrast to the controls, there was an increase in the amount of myelin debris in some experimental cultures. The neurons of these cultures, as evaluated following silver impregnation, had changes that ranged from mild degeneration to reduction or loss of neuronal cell bodies in some instances. Mild degeneration of neuronal cell processes (consisting of slight beading and fraying, to reduction in length) also occurred.

Seventy-two hours following hyperbaric oxygen exposure, the living cultures were, in general, more granular, particularly in the explant proper, with less differences in the outgrowth, although some increased granularity in this area was also detected in some instances. Myelin was reduced in amount in experimental cultures and of the myelin remaining there was evidence of an increased degree of degeneration as seen by the irregularity of the sheath and free myelin

debris. Experimental cultures impregnated for the evaluation of neuronal cell bodies and axons had in all instances a loss of tissue in the center of the explant and, thus, it was not possible to evaluated neuronal cell bodies that are generally located within this area of the explant. Explants did have detectable evidence of axonal degeneration or loss, at the edge of the areas of tissue loss of some explants, although intact axons were still detectable in some instances.

b. Cultures Exposed to Hyperbaric Oxygen at 3.0 ATA. Twenty-four hours following hyperbaric treatment living explant cultures were either not noticeably different or were only slightly more granular than the controls. Also, there was essentially no difference between the hyperbaric oxygen treated and the control cultures with regard to the outgrowth surrounding the explants. Myelin was, however, essentially absent in experimental cultures. One explant had one small strand of degenerated myelin that was very thin. Examination of silver impregnated cultures for neurons revealed that there were no dramatic differences between the experimental and control cultures after 24 hours post-exposure.

Seventy-two hours post-exposure, the changes involving myelin as detected in living cultures were essentially the same as after 24 hours. The central aspects of the cultures were slightly more granular than the controls with little difference in the outgrowth. In addition, foci of beating ependymal cilia were also detected in two of the experimental explants. There was definitely a reduction (total or almost complete absence) in the amount of myelin in the experimental cultures when compared with the controls. Evidence of undulations of still intact myelin sheaths and myelin debris was also detected in hyperbaric treated cultures. Following silver impregnation, experimental cultures had evidence of degeneration of neuronal cell bodies and axons when compared to the controls. Although some explants had minor granularity, they were sti

intact and were not as severely altered as cultures exposed to hyperbaric oxygen at 6.0 ATA, in which the central portion was completely missing as a result of degeneration.

VI. DISCUSSION

The existence of the BBB has long been viewed with mixed feelings by the medical community. This intrinsic barrier has been considered to be the major protective barrier of the central nervous system (CNS) which is responsible for the selective filtering of nutrients and metabolites required for maximal neural cell function and for the exclusion of harmful or unnecessary substances. Unfortunately, in certain disease and pathological conditions the same characteristics which make it an effective barrier become a nemesis in the subsequent treatment of many bacterial infections and neurological disorders because it excludes many antibiotics and other drugs from the brain parenchyma. Classical medical thinking dictates that a breakdown of the BBB is detrimental to brain homeostasis. However, if in fact, the permeability of the BBB is significantly altered by clinically accepted levels of hyperbaric oxygen, this generality may not be true in all cases.

This investigation provided valuable insight into the effects of hyperbaric oxygen: 1) on the BBB of several laboratory species, and 2), on specific components of the central nervous system by using an explant tissue culture model of hamster brain.

Hyperbaric oxygen studies (using a variety of pressure regimens) indicated that definite differences in susceptibility to breakdown of the BBB occurred between the guinea pig, mouse-rat, and rabbit. The guinea pig and mouse-rat were concluded to be the most resistant with the guinea pig having the most clear cut point (prominence of dye staining of tissue) at which breakdown occurred. The rabbit was concluded to be the most susceptible of the group since all groups subjected to hyperbaric oxygen had some evidence of BBB disruption as indicated by dye penetration. In addition, the effect of the different hyperbaric pressures used were quite difficult to evaluate in this species because of the lack of clear-cut differences between the control and experimen-

tal CNS tissue of animals examined.

Reasons for the differences in susceptibility of the BBB to hyperbaric oxygen pressures among the different species could not be explained based on the work conducted to date. However, further investigation of the mechanisms involved in the differences encountered could contribute valuable insight into the function of the BBB, and is considered to be warranted.

The absence of lesions in the brains of animals that were examined by light microscopy immediately following exposure to hyperbaric oxygen treatment were concluded to be due to the lack of sufficiently high inspired partial pressures of oxygen and/or an insufficient time period between exposure and euthanasia of experimental animals to permit morphologic alterations of affected components to develop.

Preliminary results of hyperbaric oxygen effects on the components of the central nervous system in vitro were fairly distinct. At 6.0 ATA of pressures there was degeneration of both the explant, myelin sheaths and neurons (cell bodies and processes). Explant and myelin sheath lesions were detectable as early as 24 hours post-exposure while there was mild degeneration of neurons themselves. After 72 hours, explant degeneration was more severe in experimental cultures. Demyelination and degeneration of neuronal cell processes were also detected. Evaluation of neuronal cell bodies was not possible because of loss of the central parts of the experimental explants (due to degeneration) following processing for silver impregnation.

Although only limited numbers of cultures were used in this preliminary study, it appeared that the susceptibility of myelin to 3.0 ATA was equal to, or even possibly greater than, that detected following exposure to 6.0 ATA of pressure and was detectable after 24 hours. In contrast, there was essentially no difference between general explant growth and neurons (cell bodies and processes) after 24 hours. After 72 hours the degeneration of myelin was essen-

tially the same and explants were slightly more granular than after 24 hours. In contrast to the results after 24 hours, at 72 hours neurons had degeneration of cell bodies and processes. Further, compared to cultures exposed to 6.0 ATA of pressure to those exposed to 3.0 ATA for 72 hours, general explant integrity was much less affected.

The exact mechanisms by which pressure and oxygen are involved in causing degeneration of the glial tissue, myelin sheaths and neurons in tissue culture explants of hamster brain could not be determined, but further pursuit of their identification would represent a very fertile area for future investigation.

VII. REFERENCES

1. Balentine, J.D. (1975)
Ultrastructural pathology of hyperbaric oxygenation in the central nervous system: Observations in anterior horn gray matter.
Lab. Invest. 31(6): 580-592.

2. Brightman, M.W. (1977)
Morphology of blood-brain barrier interfaces.
Exp. Eye Res. 25, Suppl. 1-25.

3. Chryssanthou, C., B. Graber, S. Mendelson, and G. Goldstein (1979).
Increased blood-brain barrier permeability to tetracycline in rabbits under dysbaric conditions.
Undersea Biomed. Res. 6(4): 319-328.

4. Chryssanthou, C., M. Springer, and S. Lipschitz (1977)
Blood-brain and blood-lung barrier alteration by dysbaric exposure.
Undersea Biomed. Res. 4(2): 117-129.

5. Ehrlich, P. (1885)
Das sauerstoff-bedurfnis des organismus: eine farbenanalytische studie.
Berlin: Hirschward

6. Goldmann, E.E. (1913)
Vitalfarbung am Zentrulnervensystem.
Berlin: Eimer

7. Gruenau, S.P., M.T. Folker, and S.I. Rapoport (1981).
Lack of hyperbaric O2 effect on blood-brain barrier permeability in conscious rats.
Aviat. Space Environ. Med. 5(3): 162-165.

8. Hossmann, K. and Y. Olsson (1974).
The effect of transient cerebral ischemia on the vascular permeability to protein tracers.
Acta Neuropathol. (Berl) 18: 103-112.

9. Johansson, B. (1975).
Blood-brain barrier dysfunction in experimental gas embolism. In: Program and Abstracts, Sixth Symposium on Underwater Physiology, San Diego, p., 27.

10. Lanse, S.B., J.C. Lee, E.A. Jacobs, and H. Brody (1978).
Changes in the permeability of the blood-brain barrier under hyperbaric conditions.
Aviat. Space Environ. Med. 49(7): 890-894.

11. Lee, J.C., and J. Olszewski (1959).
Effect of air embolism on permeability of cerebral blood vessels.
Neurol. 9: 619-625.

12. Persson, L.I., B.B. Johansson, and H.-A. Hansson (1978).
Ultrastructural studies on blood-brain barrier dysfunction after cerebral air embolism in the rat.
Acta Neuropathol. (Berl) 44: 53-56.

13. Povlishock, J.T., D.P. Becker, H.G. Sullivan and J.D. Miller (1978).
Vascular permeability alterations to horseradish peroxidase in experimental brain injury.
Brain Res. 153: 223-239.

14. Rattazzi, M.C., S.B. Lanse, R.A. McCullough, J.A. Nester, and E.A. Jacobs. Towards enzyme replacement in GM2-gangliosidosis: Organ disposition and induced CNS uptake of human B-hexosaminidase in the cat.
Second Intl. Symp. Enzyme Therapy and Genetic Disease. Eds. R.J. Desnick, Birth Defects: Orginal Article Series 16: 179-193. New York, Alan R. Liss, 1980.

15. Reese, T.S. and M.J. Karnovsky (1967)
Fine structural localization of blood-brain barrier to exogenous peroxidase.
J. Cell Biol. 34: 207-217.

16. Storts, R.W. and A. Koestner (1968).
General cultural characteristics of canine cerebellar explants.
Am. J. Vet. Res. 29(12): 2351-2364.

17. Storts, R.W. and A. Koestner (1969).
Development and characterization of myelin in tissue culture of canine cerebellum.
Z. Zellforsch. 95: 9-18.

18. Terro, T., J. Legtosalo, V-P. Lehto, M. Heino, I. Kantola, and L.A. Laittinen, (1980).
Permeability changes in cerebral, iridic, and retinal vessels during experimental decompression sickness in the rat.
Aviat. Space Environ. Med. 51(2): 137-141.

19. Westergaard, E. (1977).
The blood-brain barrier to horseradish peroxidase under normal and experimental conditions.
Neuropathol. (Berl) 39: 181-187.

20. Wolman, M., I. Klatzo, E. Chui, F. Wilmes, K. Nishimoto, K. Fujiwara, and M. Spatz (1981).
Evaluation of the dye-protein tracers in pathophysiology of the blood-brain barrier.
Acta Neuropathol. (Berl) 54: 55-61.

VIII. EXPERIMENTAL SUBJECTS

All animals will be used and cared for in compliance with DOD Directive 3216.1 and AFR 169-2, "The Use of Animals in DOD Programs", and NIH Publication #85-23, "Guide for the Care and Use of Laboratory Animals". Appropriate consideration will be given to the policies, standards, and guidelines for the proper use, care, handling, and treatment of animals. This type of experiment cannot be done in an *in vitro* model as it is not an accurate representation of the various systems operant in the live animal. The techniques and model outlined under the section entitled "MATERIALS AND METHODS" are well established in the literature and lend themselves very well to this experiment.

IX. ADDENDUM A

Surgical Procedure for Catheter Implantation

At least 48 hours prior to surgery, silastic catheters were constructed as follows. A 2.5 inch sleeve of 0.03" x 0.065" silastic tubing was soaked in xylenes for approximately one hour. This served to expand the tubing to facilitate the insertion of the smaller bore tubing. A 5.5 in. section of silastic tubing (0.02" x 0.037") was threaded into the larger bore tubing so that one end was flush. This was accomplished by threading a piece of 1-0 braided silk through the large bore tubing, tying it to the end of the small tubing and working the smaller bore tubing through the shaft of the larger tube. These catheters were allowed to dry for several hours. Following drying, a ring made with silastic medical adhesive was positioned approximately one inch from the free end of the small bore tubing. This ring was fashioned so that the tubing could be easily inserted into a small opening in the vessel. The ring was molded in a cone shape with the pointed end positioned to enter the vessel first. The entire diameter of the cone was approximately the same as the outside diameter of the large bore tubing (0.067"). The silastic adhesive was allowed to cure for at least 24 hours before being used. The end of the catheter inserted in the vessel was trimmed to a short bevel to facilitate vessel entry.

All instruments, catheters and other supplies were autoclaved prior to surgery. The surgical procedure described can only be considered "clean" surgery because of the necessity of dorsally exteriorzing the catheter between the shoulder blades of the animal. However, extreme care was taken to maintain a sterile environment in the area of the incision and at the point of entry of the catheter into the vascular system.

Because of its proximity to the surface and its relatively large diameter, the jugular vein was chosen as the vessel for catheter implantation.

Animals were anesthetized with Nembutal or acepromazine-ketamine and the area under the jaw and neck, and the scapula area were shaved free of hair. The clipped areas were thoroughly cleaned with betadine, and an incision (right of midline) was made from the area of the thoracic inlet towards the right jaw for approximately one inch. This incision only penetrated the skin. Just prior to this procedure a small incision was made over the scapula of the animal. A subcutaneous tunnel was then formed with curved hemostats, and the catheter passed through to the area of the vein. The jugular vein was gently isolated just anterior to its entry into the thoracic cavity and two lengths of 4-0 silk were positioned under the vein.

The catheter was prepared for vessel entry by inserting a 22g tubing adapter into the scapula end, and filling the catheter with sterile saline. The cranial end of the vessel was ligated with one of the lengths of 4-0 silk. The ends of the tie were used at a later time to help secure the catheter. A loose knot was made around the caudal end of the vessel. With slight tension on the vein a small perpendicular cut was made and then enlarged with a small longitudinal cut extending caudally from the first cut. The catheter was inserted and the caudal suture tied just above the silastic adhesive ring. Care was taken to insure that the vessel was not occluded by the caudal tie. The catheter was then further secured by the ends of the rostral tie. This also secured the two sections of the vessel. The wound was closed with a non-absorbable suture. A heparin lock was placed in the tubing adaptor and the entire unit secured with several sutures along the midline of the back extending caudally. At least one suture was taken at the point of catheter entry to insure that the opening remained as small as possible.

Post-operative care included a daily application of bactericidal salve at the point of catheter entry and a daily flushing using sterile saline. It was not necessary to use heparinized saline because clotting is not a problem with

silastic tubing.

Using this technique catheters remained patent for 8-14 days.

X. ADDENDUM B

Neuropathologic Techniques

1. Sodium Flourscein Studies. The micromolecular dye sodium flourscein (NaFl) will be used throughout these as the indicator of BBB permeability. Sodium flourscein has been used extensively in BBB studies (8) and is considered an indicator of BBB permeability. In order to obtain a comprehensive picture of CNS changes, ten levels of tissue will be examined in each animal.

Each animal will be appropriately anesthetized (pentobarbital, I.P., 40 mg/kg), the hair removed from the femoral area by a depilatory and the area swabbed with an antiseptic solution. A small incision will be made, the femoral vein isolated and the dye injections will be made at this point. Following perfusion with buffered formalin, the brain will be removed and cut into seven sagital sections. The spinal cord will be examined at the cervical, thoracic and lumbar levels. Each tissue section will be examined under a dissecting scope for NaFl staining. In order that sections can be viewed with the microscope while being exposed to ultraviolet (UV) light a special light box has been constructed. All tissue will be evaluated under blind conditions, one to three hours after staining.

Photographic records will be kept of the tissue. In addition, the pathologist, Dr. Ralph Storts, DVM, Ph.D, will record his findings for each animal.

2. Light and Electron Microscopy Studies. Animals designated for microscopic study will be anesthetized and whole body fixation will be accomplished via cardiac perfusion. The perfusion fixatives will consist of 4% buffered formalin-1% gluteraldehyde mixture. The brain and spinal cord will be removed and maintained in the same fixative until examined. Preselected areas of the brain and spinal cord will be examined microscopically and ultrastructurally. For light microscopic examination, nervous tissue will be processed and embedded in paraffin for subsequent staining with hemotoxylin and eosin and other selected

methods (for myelin, neurons, and neuroglia). Tissue for electron microscopic (EM) studies will be post-fixed in 1% osmium tetroxide, dehydrated, cleared and embedded in Epon. For orientation, 1 micron sections will be cut and stained with Paragon stain and examined by light microscopy. Selected areas will then be thin sectioned, placed on uncoated grids and stained for observation by electron microscopy.

Several of the animals will be subjected to horseradish peroxidase and microperoxidase prior to processing. Peroxidases are commonly used as tracers in BBB investigations (9, 10, 11). In addition, several of the animals designated for subsequent EM analysis will be injected with ferritin-labelled immunoglobulin prior to tissue processing. Tissue will be examined for passage of the peroxidases and ferritin across the BBB of these animals.