

UNITED STATES AIR FORCE
SUMMER RESEARCH PROGRAM -- 1998
GRADUATE STUDENT RESEARCH PROGRAM FINAL REPORTS

VOLUME 7
ARMSTRONG LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES
5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Colonel Jan Cervený

Program Manager, RDL
Scott Licoscas

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly-Clemmons

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1998

20010319 037

AQM01-06-1215

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

This document is one of a set of 15 volumes describing the 1998 AFOSR Summer Research Program. The following volumes comprise the set:

<u>VOLUME</u>	<u>TITLE</u>
1	Program Management Report
	<i>Summer Faculty Research Program (SFRP) Reports</i>
2	Armstrong Laboratory
3	Phillips Laboratory
4	Rome Laboratory
5A & 5B	Wright Laboratory
6	Arnold Engineering Development Center, Air Logistics Centers, United States Air Force Academy and Wilford Hall Medical Center
	<i>Graduate Student Research Program (GSRP) Reports</i>
7	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10	Wright Laboratory
11	Arnold Engineering Development Center, and Wilford Hall Medical Center
	<i>High School Apprenticeship Program (HSAP) Reports</i>
12	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A, 15B & 15C	Wright Laboratory

REPORT DOCUMENTATION PAGE			Form Approved No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including 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 Project, Washington, DC 20503.			AFRL-SR-BL-TR-00-	
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE December, 1983		
4. TITLE AND SUBTITLE 1998 Summer Research Program (SRP), Graduate Student Research Program (GSRP), Final Reports, Volume 7, Armstrong Laboratory			5. FUNDING NUMBERS F49620-93-C-0063	
6. AUTHOR(S) Gary Moore				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research & Development Laboratories (RDL) 5800 Uplander Way Culver City, CA 90230-6608			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research (AFOSR) 801 N. Randolph St. Arlington, VA 22203-1977			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the GSRP participants at the Armstrong Laboratory.				
14. SUBJECT TERMS Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to *stay within the lines* to meet *optical scanning requirements*.

Block 1. Agency Use Only (*Leave blank*).

Block 2. Report Date. Full publication date including day, month, and year, if available
(e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	PR - Project
G - Grant	TA - Task
PE - Program Element	WU - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es).
Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es).
Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (*If known*)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with....; Trans. of....; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."

DOE - See authorities.

NASA - See Handbook NHB 2200.2.

NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.

DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.

Leave blank.

NASA - Leave blank.

NTIS -

Block 13. Abstract. Include a brief (*Maximum 200 words*) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (*NTIS only*).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

GSRP FINAL REPORT TABLE OF CONTENTS

i-vi

1. INTRODUCTION	1
2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM	2
3. RECRUITING AND SELECTION	3
4. SITE VISITS	4
5. HBCU/MI PARTICIPATION	4
6. SRP FUNDING SOURCES	5
7. COMPENSATION FOR PARTICIPATIONS	5
8. CONTENTS OF THE 1995 REPORT	6

APPENDICIES:

A. PROGRAM STATISTICAL SUMMARY	A-1
B. SRP EVALUATION RESPONSES	B-1

GSRP FINAL REPORTS

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
MR Jason M Henry	Ohio University , Athens , OH Phantom/Merlin Force-Reflecting Telcooperation:Implementation	AFRL/HEC _____	7- 1
MR Keith S Jones	University of Cincinnati , Cincinnati , OH Issues in Steady-State Visual Evoked Response Based Control	AFRL/HEC _____	7- 2
MR Christian A Kijora	Arizona State University , Mesa , AZ Research Techniques A Search for Crew Resource Management Documents	AFRL/HEA _____	7- 3
MS Vanessa D Le	Univ of Texas at Austin , Austin , TX Permeability Characteristics of an Endothelial Cell Model of the Blood-Brain barrier	AFRL/HED _____	7- 4
MR Peter D Naegele	University of Toledo , Toledo , OH Anisotropies in Visual Search Performance Across The Upper and Lower Visual Fields as a Function of	AFRL/HEP _____	7- 5
MS Nicole L Proulx	University of Dayton , Dayton , OH Nefotiation at a Distance: Why you Might want to use the Telephone	AFRL/HES _____	7- 6
MS Mary K Sheehan	Texas A & M Univ-College Station , College Station , TX Understanding Disagreement Across Rating Sources An Assessment of the Measurement Equivalence of Rat	AFRL/HEJ _____	7- 7
MR Brian D Simpson	Wright State University , Dayton , OH The Effect of Spatial Separation and Onset Asynchrony on the Detectability adn Intelligibility of a	AFRL/HES _____	7- 8
MS Julie A Stiles-Shipley	Bowling Green State University , Bowling Green , OH The Rate of Skill Acquisition for males and females on space Frotress	AFRL/HEJ _____	7- 9
MR Michael E Stiso	University of Oregon , Eugene , OR Weighing The Importance of Spatial Ortanization and Priority of Targets on UAV Mission Planning Whic	AFRL/HEJ _____	7- 10

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
MR Benjamin J Bernocco	Pennsylvania State University, University Park, PA A study of Optimal finite-Thrust Spacecraft Trajectories for the Techsat 21 Mission	AFRL/VSS _____	8- 1
MR Robert J Fuentes	Univ of Colorado at Boulder, Boulder, CO Stable Controller Design For Deployable Precision Structures Using Perturbation Theory	AFRL/VSD _____	8- 2
MR Jeffery M Ganley	University of New Mexico, Albuquerque, NM Determination of the Residual Stress Profile in a thin Composite Part	AFRL/VSD _____	8- 3
MR David A Joiner	Rensselaer Polytechnic Instit, Troy, NY A Study of the Effects of Novae on The Infrared Celestial Background	AFRL/VSF _____	8- 4
MS Johnelle L Koriath	Univ of Texas at Dallas, Richardson, TX A Computational Analysis of Stacked Blumleins Used in Pulsed Power Devices	AFRL/DEH _____	8- 5
MS Elizabeth M Monaco	Holy Cross College, Worcester, MA Approximating Morse Potentials Numerically and Analytically	AFRL/VSF _____	8- 6
MR Tyrone A Ospino	University of Puerto Rico, San Juan, PR Conducting fluid Real-Time 2-D Electronic Interpolator and Spatial Filter For wavefront Sensor-To-Co	AFRL/DEB _____	8- 7
MR Kenneth F Stephens II	University of North Texas, Denton, TX Simulation of Plasma-Wall mixing in a Magnetized Target Fusion Concept	AFRL/DEH _____	8- 8
MR Michael V Wood	Pennsylvania State University, University Park, PA Characterization of Spatial Light Modulator for Aberration compensation of Severely Distor Primary	AFRL/DEB _____	8- 9

SRP Final Report Table of Contents

Author	University/Institution Report Title	Rome Laboratory Directorate	Vol-Page
MR Sang H Bae	Univ of Calif, Los Angeles , Los Angeles , CA A Simulation Study of the Vulnerabilities in Commercial Satellite Constellations	AFRL/IFG _____	9- 1
MR Keith R Buck	Colorado State University , Fort Collins , CO Near-Optimal Routing of Unmanned Surveillance Platforms	AFRL/IFE _____	9- 2
MR Kevin P Crossway	SUNY OF Tech Utica , Utica , NY	AFRL/IFT _____	9- 3
MR Brian R Waterhouse	Syracuse University , Syracuse , NY Empirical and Theoretical Foundations for a Two-Dimensional Non-Homogeneity Detector for Radar	AFRL/SNR _____	9- 4

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
MR Jeffrey C Anderson	Clemson University , Clemson , SC Growth and Characterization of 3-inch Nitride Semiconducting Epitaxial Films	AFRL/SNH _____	10- 1
MR Erik L Antonsen	Northern Illinois University , Urbana , IL Modified herriott Cell Interferometr for Pulsed Plasma Thruster Neutral Density Measurements	AFRL/PRR _____	10- 2
MR Daniel J Bodony	Purdue University , West Lafayette , IN MDICE Analysis of An F-18C Wing	AFRL/VAA _____	10- 3
MR Gregory C Harding	Florida Inst of Technology/Gen. Washingt , Melbourne , FL Interactions between Weakly Ionized Gas Plasmas and Shock Waves A Revoew	AFRL/PRT _____	10- 4
MR John L Hazel	Western Michigan University , Kalamazoo , MI The Physical Basis of Boid And Crotaline Infrared Detection	AFRL/MLP _____	10- 5
MR Timothy J Leger	Wright State University , Dayton , OH Enhancements to a Driect Aeroelastic Stability Computational Model	AFRL/VAS _____	10- 6
MR Ronald O Nelson	University of Idaho , Moscow , ID A Detailed Study of the Numerical properties of FDTD Algorithms for Dispersive Media	AFRL/VAA _____	10- 7
MR Andrew D Panetta	Rensselaer Polytechnic Instit , Troy , NY The Design of a Double Strut Support System for Low Speed wind Tunael Testing of Rotating, Axisymmet	AFRL/PRS _____	10- 8
MR Eduardo L Pasilião	University of Florida , Gainesville , FL A Greedy Randomized Adaptive Search Procedure for the Multi-Criteria Radio Link Frequency Assignment	AFRL/MN _____	10- 9
MR Craig A Riviello	Wright State University , Dayton , OH In-SituSynthesis of Discontinuously Reinforced Titanium alloy Composites Via Blended Eldemental Powd	AFRL/ML _____	10- 10
MS Lisa A Schaefer	Arizona State University , Tempe , AZ Evaluation of Using Agents for Factory Layout Affordability	AFRL/ML _____	10- 11

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
MS Katherine J Schafer	University of Detroit, Detroit, MI Synthesis of 7-Benzothiazol-2YL-9,9-Didecylfluorene-2Ylamine: A Versatile Intermediate For a New Ser	AFRL/ML	10- 12
MR Aboubakar Traore	Stevens Inst of Technology, Hoboken, NJ Theory of Envelope-Function within 6*6 Luttinger Model in holes Subband States of Si/Ge Quantum Well	AFRL/SNH	10- 13
MR Robert A Weisenseel	Boston University, Boston, MA MRF Segmentation for Feature Extraction in Sar Chip Classification	AFRL/SNA	10- 14
MR Jerry M Wohletz	Massachusetts Inst of Technology, Cambridge, MA Parameter Estimation for the Tailless Advanced Fighter Aircraft (TAFA)	AFRL/VAC	10- 15

SRP Final Report Table of Contents

Author	University/Institution	Arnold Engineering Development Center Directorate	Vol-Page
MR Gregory M Laskowski	Stanford University, Stanford, CA	AEDC	11- 1
	Wind Validation: Incompressible Turbulent Flow Past A Flat Plate		
MS Donna M Lehman	Univ of Texas Health Science Center, San Antonio, TX	WHMC	11- 2
	Relationship Between Growth Hormone and Myelin Basic Protein Expression In Vivo		

1. INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from \$20,000 to \$25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1997 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435
1997	148	98	140	427
1998	85	40	88	213

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education*, *Winds of Change*, and *IEEE Spectrum*. Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1998 are shown in this table.

1998 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	382	85	13
(HBCU/MI)	(0)	(0)	(0)
GSRP	130	40	7
(HBCU/MI)	(0)	(0)	(0)
HSAP	328	88	22
TOTAL	840	213	42

4. SITE VISITS

During June and July of 1998, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/MIs annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCU/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1998 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1998 SRP selected participants are shown here.

1998 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	67	38	75
USAF Laboratory Funds	17	2	13
Slots Added by AFOSR (Leftover Funds)	0	0	0
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	0	0	N/A
TOTAL	84	40	88

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1998 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996	1997	1998
Faculty Members	\$690	\$718	\$740	\$740	\$740	\$770	\$770	\$793
Graduate Student (Master's Degree)	\$425	\$442	\$455	\$455	\$455	\$470	\$470	\$484
Graduate Student (Bachelor's Degree)	\$365	\$380	\$391	\$391	\$391	\$400	\$400	\$412
High School Student (First Year)	\$200	\$200	\$200	\$200	\$200	\$200	\$200	\$200
High School Student (Subsequent Years)	\$240	\$240	\$240	\$240	\$240	\$240	\$240	\$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of \$52/day for faculty and \$41/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and GSRP associates, 65 % claimed travel reimbursements at an average round-trip cost of \$730.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-three percent (85 out of 188) of faculty associates took orientation trips at an average cost of \$449. By contrast, in 1993, 58 % of SFRP associates elected to take an orientation visits at an average cost of \$685; that was the highest

percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

This is the third year of using direct deposit for the SFRP and GSRP associates. The process went much more smoothly with respect to obtaining required information from the associates, about 15% of the associates' information needed clarification in order for direct deposit to properly function as opposed to 7% from last year. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1998 REPORT

The complete set of reports for the 1998 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1998 associates, as indicated below:

1998 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, USAFA, WHMC	6	11	

APPENDIX A -- PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. Selectees represent 44 states.

GSRP - Applicants came from 44 states. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	85
GSRP	40
HSAP	88
TOTAL	213

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	83	0	83
Master's	1	3	4
Bachelor's	0	22	22
TOTAL	186	25	109

SFRP Academic Titles	
Assistant Professor	36
Associate Professor	34
Professor	15
Instructor	0
Chairman	0
Visiting Professor	0
Visiting Assoc. Prof.	0
Research Associate	0
TOTAL	85

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	177	47
Colleague familiar with SRP	104	24
Brochure mailed to institution	101	21
Contact with Air Force laboratory	101	39
<i>IEEE Spectrum</i>	12	1
<i>BIIHE</i>	4	0
Other source	117	30
TOTAL	616	162

APPENDIX B -- SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	100
HSAPs	75
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

- A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).
- B. Faculty Associates suggest higher stipends for SFRP associates.
- C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1998 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

Lab	Evals Recv'd	How Many Associates Would You Prefer To Get ?								(% Response)			
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
		0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
USAF	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	84	32%	50%	13%	5%	80%	11%	6%	0%	73%	23%	4%	0%

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
 - a. Time available for evaluation of applications:
 - b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5.
 - a. Benefits of associate's work to laboratory:
 - b. Benefits of associate's work to Air Force:
6.
 - a. Enhancement of research qualifications for LFP and staff:
 - b. Enhancement of research qualifications for SFRP associate:
 - c. Enhancement of research qualifications for GSRP associate:
7.
 - a. Enhancement of knowledge for LFP and staff:
 - b. Enhancement of knowledge for SFRP associate:
 - c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10.
 - a. Your working relationship with SFRP:
 - b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
 b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Responses to above questions

	<i>AEDC</i>	<i>AL</i>	<i>USAF</i>	<i>PL</i>	<i>RL</i>	<i>WHMC</i>	<i>WL</i>
	<i>A</i>						
<i># Evals Recv'd</i>	0	7	1	14	5	0	46
<i>Question #</i>							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1998 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 120 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1998 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13.
 - a. Would you apply again?
 - b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
 - a. At Home:
 - b. With Friend:
 - c. On Local Economy:
 - d. Base Quarters:
17. Value of orientation visit:
 - a. Essential:
 - b. Convenient:
 - c. Not Worth Cost:
 - d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1997 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1998 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1998 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 23 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

PHANToM/MERLIN FORCE-REFLECTING TELEOPERATION: IMPLEMENTATION

Jason M. Henry
Graduate Student
Department of Mechanical Engineering

Ohio University
257 Stocker Center
Athens, OH 45701-2979

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1998

PHANTOM/MERLIN FORCE-REFLECTING TELEOPERATION: IMPLEMENTATION

Jason M. Henry
Graduate Student
Department of Mechanical Engineering
Ohio University

Abstract

This report focuses on implementation of *PHANTOM/Merlin* force-reflecting teleoperation. The purpose of the research was to enhance telepresence in human-centered teleoperation of remote manipulators using haptic feedback. The commercially available 3-degree-of-freedom PHANTOM haptic interface was used to command translational rates to a 6-degree-of-freedom Merlin robot. A force/torque sensor on the remote manipulator enabled a force/moment accommodation algorithm (FMA), and it also allowed the user to feel the forces that the remote manipulator exerted on the environment. The FMA allowed the Merlin to orient itself so that minimum contact wrench was exerted at the end effector. This allowed for general spatial motion of the Merlin despite the fact that the user could only command translations. This method of rate control and FMA enabled on all axes at all times is unique, and it has been named the Naturally-Transitioning Rate-to-Force Controller (NTRFC). This setup was ideally suited to peg-in-the-hole tasks using the Fitt's Law taskboard in the Human Sensory Feedback Lab. Preliminary data has shown the NTRFC to be an effective method of control.

PHANTOM/MERLIN FORCE-REFLECTING TELEOPERATION: IMPLEMENTATION

Jason M. Henry

INTRODUCTION

Teleoperation of remote manipulators is greatly enhanced by using a force-reflecting input device. This haptic feedback increases the sense of telepresence by enabling the operator to feel the forces and moments exerted by the slave manipulator on the environment. Rate control can be an effective method of teleoperation, but it presents a problem when the manipulator contacts the environment. If a rate is commanded while the slave is in contact with the environment the slave will try to push through the environment and the contact forces integrate to unacceptably large values. This can result in damage to the slave or the environment. This problem can be overcome by the use of a force/moment accommodation (FMA) algorithm working simultaneously with the rate algorithm. This method of control has been named the Naturally Transitioning Rate to Force Controller (NTRFC). When NTRFC is enabled the slave moves through free space using rate control, but when the slave contacts the environment the FMA algorithm commands the slave to move such that minimum contact wrench is exerted. The result is stable contact where the force the slave exerts on the environment is proportional to the rate commanded by the master. The NTRFC is similar to impedance control (Hogan, 1985) with only the damping term when in contact with the environment but in free motion it is pure rate control. The NTRFC requires no mode changes, logical switches, or gain changes in the controller software or hardware. It is important to note that the rate control and FMA are enabled on all axes at all times.

The NTRFC control method was implemented using the 3-degree-of-freedom PHANTOM haptic interface as the master, and a Modular Expandable Robot Line (Merlin) as the slave. The PHANTOM could only command translational motion, but the NTRFC allowed for general spatial motion when the Merlin came into contact with the environment. Other methods used to control the PHANTOM/Merlin combination include: pure rate control with force reflection, pure rate control without force reflection, position and orientation (POSE) control with force reflection, and POSE control without force reflection. It is important to note that when using POSE control the orientation of the end effector was fixed since only translations could be commanded with the PHANTOM. To the extent of the author's knowledge this was the first time that the PHANTOM was used for force-reflecting teleoperation of a kinematically dissimilar remote manipulator. The only other known teleoperation was done by SensAble Inc. using one PHANTOM to drive another. This report will concentrate on the NTRFC algorithm since the other modes of control are already well documented using different master/slave combinations.

METHODOLOGY

The PHANToM and Merlin are shown in figures 1 and 2 respectively.

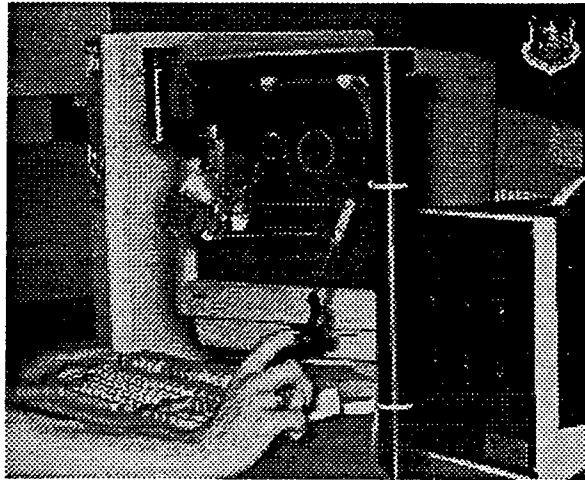


Figure 1. The PHANToM haptic interface used as the master device.

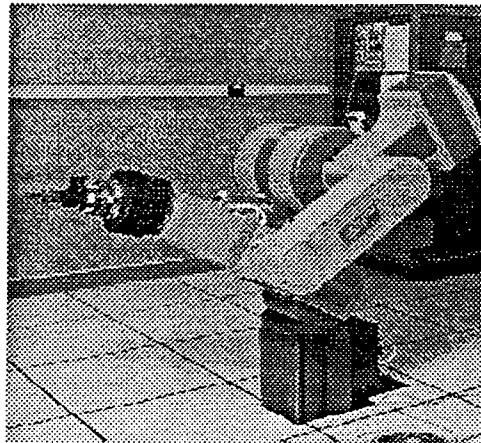


Figure 2. The Merlin slave manipulator.

The telerobotic control architecture presented in this report requires kinematics transformations, which relate Cartesian and joint variables within the master and slave devices. Specifically, this section presents the *DH* parameters, forward kinematics transformation, and Jacobian matrices for the *PHANToM* master and *Merlin* slave. The *PHANToM* forward kinematics solution is required for determining operator input commands and the Jacobian matrix is required for force-reflection to the operator. The *Merlin* forward kinematics solution is required for on-line control and the Jacobian matrix is essential for resolved-rate control.

DH Parameters

The Denavit-Hartenberg (DH) parameters provide a standard means to describe the joint/link geometric relationships in a serial manipulator (Craig, 1989). Figure 3 shows the *Merlin* and Figure 5 the *PHANToM* kinematic diagrams, from which the DH parameters of Tables 1 and 2 are derived. All angular units are *degrees*. If the θ_1 angular offset of row 3 is included for the *Merlin*, Figure 3 shows the zero-joint-angle configuration. Nominal *Merlin* and measured *PHANToM* joint angle limits are also given. Nominal values for the *Merlin* lengths are: $a_2 = 17.375$, $d_2 = 119$, $d_4 = 17.25$, and nominal values for the *PHANToM* lengths are: $L_1 = L_2 = 8.25$. All linear units are *inches*.

Table 1. *Merlin* DH Parameters

i	α_{i-1}	a_{i-1}	d_i	θ_i	Limits
1	0	0	0	θ_1	± 147
2	-90	0	d_2	θ_2	+56,-230
3	0	a_2	0	$\theta_3 - 90$	+56,-230
4	-90	0	d_4	θ_4	± 360
5	90	0	0	θ_5	± 90
6	-90	0	0	θ_6	± 360

Table 2. *PHANToM* DH Parameters

I	α_{i-1}	a_{i-1}	d_i	θ_i	Limits
1	0	0	0	θ_1	± 90
2	-90	0	0	θ_2	+110,-70
3	0	L_1	0	θ_3	+135,+45

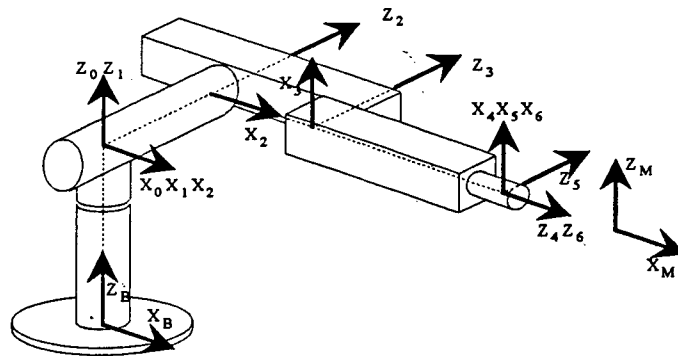


Figure 3. Merlin kinematic diagram.

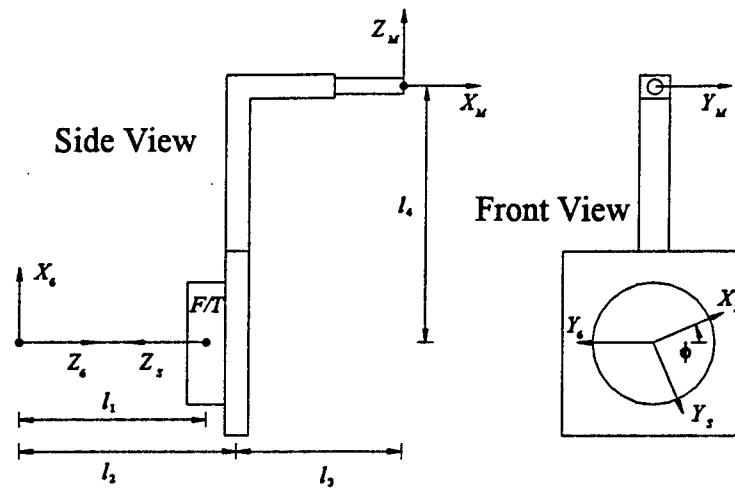


Figure 4. Merlin peg details.

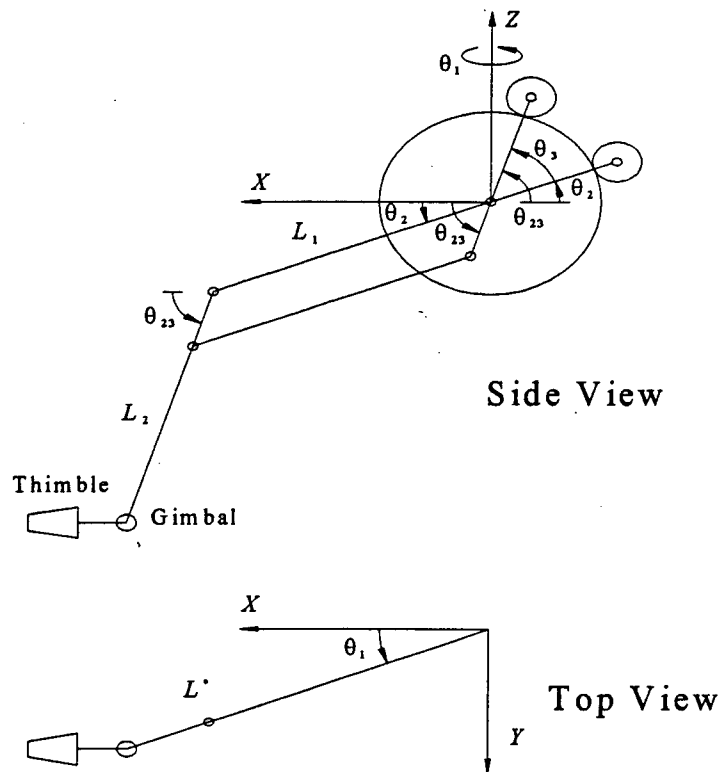


Figure 5. PHANToM kinematic diagram.

Figure 4 shows the details for the wrist {6}, F/T sensor {S}, and MRF {M} frames in the HSF lab, where the peg for insertion is bolted to the F/T sensor via the L-shaped handle. Three required homogeneous transformation matrices are given below. ${}^6_M T$ is required in forward kinematics and ${}^S_M T$ is required in the force/moment accommodation algorithms. Nominal end-effector values are $l_1=4.5$, $l_2=5.5$, $l_3=5$, $l_4=8$ (inches).

$${}^6_M T = \begin{bmatrix} 0 & 0 & 1 & l_4 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & l_2 + l_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad {}^S_S T = \begin{bmatrix} s\phi & -c\phi & 0 & 0 \\ -c\phi & -s\phi & 0 & 0 \\ 0 & 0 & -1 & l_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

$${}^M_S T = {}^M_6 T {}^6_S T = {}^6_T^{-1} {}^S_T \quad (2)$$

Forward Kinematics

The forward kinematics transformation gives the position and orientation (pose) of the moving frame of interest n with respect to the kinematic base frame 0 (Craig, 1989):

$${}^0_n T = \begin{bmatrix} {}^0_n R & {}^0P_n \\ 0 & 1 \end{bmatrix} \quad {}^0X_n = \{x \ y \ z \ \gamma \ \beta \ \alpha\}^T \quad (3)$$

The pose can be represented by ${}^0_n T$ (the 4x4 homogeneous transformation matrix with the 3x3 orientation matrix ${}^0_n R$ and the 3x1 position vector ${}^0P_n$) or ${}^0X_n$ (whose first 3 components are ${}^0P_n$ and second 3 are orientation numbers extracted from ${}^0_n R$, e.g. Z-Y-X Euler convention, see Eq. 15). Given one row in a *DH* parameter table, the homogeneous transformation matrix relating the pose of neighboring frames in a serial chain is (Craig, 1989):

$${}^{i-1}_i T = \begin{bmatrix} c\theta_i & -s\theta_i & 0 & a_{i-1} \\ s\theta_i c\alpha_{i-1} & c\theta_i c\alpha_{i-1} & -s\alpha_{i-1} & -d_i s\alpha_{i-1} \\ s\theta_i s\alpha_{i-1} & c\theta_i s\alpha_{i-1} & c\alpha_{i-1} & d_i c\alpha_{i-1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

where $c\theta_i = \cos(\theta_i)$, $s\theta_i = \sin(\theta_i)$, etc. The forward kinematics transformation for active joints is:

$${}^0_n T = \prod_{i=1}^n {}^{i-1}_i T = {}^0_1 T {}^1_2 T \dots {}^{n-1}_n T \quad (5)$$

The overall forward kinematics for the *Merlin* is given below:

$${}^W_0 T = {}^W_0 T {}^B_0 T {}^S_B T {}^M_S T \quad (6)$$

where M , B , O , and W stand for the *MRF*, *Base*, O , and *Wrist* frames. The world frame is Wo .

For the *PHANToM*, the forward kinematics solution is simpler since orientation of the finger thimble cannot be determined using *HSF* hardware. The symbolic forward kinematics solution is derived from Fig. 6 (we have no need to use the *DH* parameters), recognizing from the top view that $L^* = L_1 c_2 + L_2 c_{23}$. The solution is (where X , Y , and Z are expressed in the *PHANToM* base coordinates):

$$\begin{aligned} X &= c_1 (L_1 c_2 + L_2 c_{23}) \\ Y &= s_1 (L_1 c_2 + L_2 c_{23}) \\ Z &= -(L_1 s_2 + L_2 s_{23}) \end{aligned} \quad (7)$$

where $c_i = \cos(\theta_i)$, $s_i = \sin(\theta_i)$, $c_{23} = \cos(\theta_2 + \theta_3)$, and $s_{23} = \sin(\theta_2 + \theta_3)$. Care must be taken to distinguish between the *Merlin* joint angles θ_i and the *PHANToM* joint angles θ_i . The simplified forward kinematics solution on the right makes use of the equal link lengths $L_1 = L_2 = L$. It is important to note that the third *PHANToM* angular encoder directly reads $(\theta_2 + \theta_3)$ rather than just θ_3 .

Jacobian Matrices

The Jacobian matrix ${}^k J$ for a serial chain maps joint rates $\dot{\Theta} = \{\dot{\theta}_1 \ \dot{\theta}_2 \ \dots \ \dot{\theta}_n\}^T$ into Cartesian rates ${}^k \dot{X} = \{\dot{x} \ \dot{y} \ \dot{z} \ \omega_x \ \omega_y \ \omega_z\}^T$ of the frame of interest with respect to the base, expressed in any frame k : ${}^k \dot{X} = {}^k J \dot{\Theta}$. The i^{th} column of ${}^k J$ is the Cartesian velocity of the point of interest due to joint rate i alone (with $\dot{\theta}_i$ factored out). This fact leads to the following formula for the i^{th} column of ${}^k J$, where ${}^i z_i = \{0 \ 0 \ 1\}^T$:

$${}^k J_i = \begin{Bmatrix} {}^k R \left({}^i z_i \times {}^i P_n \right) \\ {}^k R^i z_i \end{Bmatrix} \quad (8)$$

Equation 8 is applied for each moving joint to yield the configuration-dependent 6x6 *Merlin* translational and rotational Jacobian matrix, relating the velocity motion of the *Wrist* with respect to the *Base* frames, expressed in k (k is chosen as the *Merlin* {0} frame in this report).

The *PHANToM* Jacobian matrix is obtained symbolically from the forward kinematics solution and the general definition of the Jacobian matrix:

$$J = \left[\frac{\partial G_i}{\partial q_j} \right] \quad (9)$$

where i is the row index and j the column index; the vector of dependent functions are $G = \{X \ Y \ Z\}^T$ and the vector of independent variables is $q = \{\theta_1 \ \theta_2 \ \theta_3\}^T$.

Using Eqs. 7 and 9, the configuration-dependent 3x3 *PHANTOM* translational Jacobian is:

$${}^0J = \begin{bmatrix} -Ls_1(c_2 + c_{23}) & -Lc_1(s_2 + s_{23}) & -Lc_1s_{23} \\ Lc_1(c_2 + c_{23}) & -Ls_1(s_2 + s_{23}) & -Ls_1s_{23} \\ 0 & -L(c_2 + c_{23}) & -Lc_{23} \end{bmatrix} \quad (10)$$

Resolved Rate Control

The algorithm implemented was based on Whitney's method (1969), and figure 6 shows the general algorithm. This section assumes a static *Base* and *CRF*; the method can be extended to handle moving *Base* and Control Reference Frame (*CRF*) frames for dynamic tasks. The time-varying manipulator Jacobian matrix maps joint rates to Cartesian rates of the *Wrist*: ${}^k\dot{X}_W = {}^kJ\dot{\Theta}$. The Cartesian rates ${}^k\dot{X}_W = {}^k\{\underline{v}_W \ \underline{\omega}_W\}^T$ express the translational and rotational velocities of the *Wrist* with respect to the *Base*, expressed in the coordinates of any frame k . Common choices are $k = \text{Wrist}$, 0, or *Base*; simplest symbolic terms for the Jacobian matrix result when k is the frame midway between the *Base* and *Wrist*, often the *Elbow* frame. The equation ${}^k\dot{X}_W = {}^kJ\dot{\Theta}$ must be inverted (or, more efficiently, solved by Gaussian elimination) at each control step. First, the input velocity $\dot{X}_{HC}$ must be converted to the resolved rate input ${}^k\dot{X}_W$ (the equivalent Cartesian velocities of the *Wrist* frame to produce $\dot{X}_M$). This rigid-body velocity transformation and coordinate transformation is given in Eq. 11 (Craig, 1989).

$${}^k\dot{X}_W = \begin{bmatrix} \underline{v}_W \\ \underline{\omega}_W \end{bmatrix} = \begin{bmatrix} {}^k_jR & {}^k_WR^W P_{MRF} \times {}^k_jR \\ 0 & {}^k_jR \end{bmatrix}^j \begin{bmatrix} \underline{v}_{MRF} \\ \underline{\omega}_{MRF} \end{bmatrix} \quad (11)$$

$\dot{X}_M$ always gives the six-dof velocity of *MRF* with respect to *Base*, but can be expressed in any coordinates j (e.g. *CRF*, *Base*, 0, or *World*). Now the rate equation is inverted to calculate the instantaneous joint rates necessary to obtain the commanded ${}^k\dot{X}_W$:

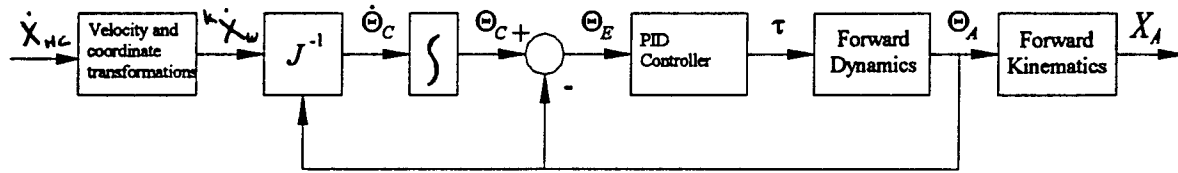


Figure 6. Resolved rate control algorithm

$$\dot{\Theta}_C = {}^k J^{-1} {}^k \dot{X}_w \quad (12)$$

The commanded joint rates are numerically integrated to commanded joint angles Θ_C . These angles are commanded to the Merlin and achieved using linear independent *PID* control already implemented in the Merlin software. Joint encoder feedback Θ_A is used to form the errors for servo control.

This algorithm is sensitive to kinematic singularities, where the manipulator loses freedom to move in one or more Cartesian direction. In the neighborhood of singularities, extremely high joint rates are theoretically required to satisfy a finite Cartesian command. To deal with this problem, the determinant of the Jacobian matrix ${}^k J$ must be monitored. When the determinant approaches zero, the matrix inverse (or Gaussian elimination) in Eq. 12 is replaced by a matrix pseudoinverse based on Singular Value Decomposition (*SVD*). Near singularities, the exact Cartesian command ${}^k \dot{X}_w$ cannot be satisfied, but the *SVD* will yield bounded joint rates which will move the manipulator through the singular neighborhood until Eq. 12 can take over again.

For teleoperation, the displacement of the operator's hand with the master device is interpreted to be the rate $\dot{X}_{HC}$.

Constant Force Return to Center with Virtual Walls

For Cartesian rate input commands, the manipulator will move with a commanded velocity when the PHANToM Cartesian pose is different from its reference pose G_0 . Therefore, a return-to-center (*RTC*) force should be provided to assist the operator's hand in finding the zero input pose. A constant-force return-to-center (*CFRTC*) approach was developed. Figure 7 shows the *CFRTC* force as a function of scalar displacement ΔX_{HCl} from the zero reference for one axis. The magnitudes in figure 7 are arbitrary and must be determined for specific master devices based on performance requirements. It was found that gains ranging from 50-100 were ideal depending on the users preference. The *CFRTC* is symmetric about $\Delta X_{HCl} = 0$; each side displays three distinct but continuous regions. The first is the deadband, serving two purposes: a) providing a small region of zero input surrounding the zero pose; and b) providing a parabolic virtual wall which the operator must overcome if an input is to be commanded in that particular Cartesian axis. The second largest zone is the working range, which provides the *CFRTC*. The third zone provides a stiff virtual spring to alert the operator when the edge of the PHANToM workspace is encountered. When implemented it was found that this stiff spring was unnecessary so the flat *CFRTC* zone was extended to the workspace boundary. In this case, the operator must be aware of the workspace boundaries.

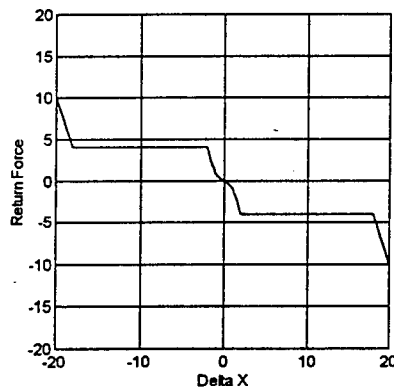


Figure 7. CFRTC for One Cartesian Axis

Force Control

An active force controller has been implemented in the resolved-rate scheme to command forces to the environment with the manipulator. This active force controller is basically a general impedance controller (Hogan, 1985) with only the damping term. A six-dof force/torque sensor (with frame S) mounted after the last joint reads

the contact wrench $F_S = \{f_S \ m_S\}^T$. The weight and gravity-moment of the end-effector mounted outboard of the *F/T* sensor (transformed to *S*) must be subtracted from the sensor reading. This modified sensor reading in *S* must be transformed by rigid body transformations and coordinate rotations (Craig, 1989) to the equivalent *MRF* wrench:

$$F_M = \begin{Bmatrix} f_M \\ m_M \end{Bmatrix} = \begin{bmatrix} {}^M_S R & 0 \\ {}^M_{P_S} \times {}^M_S R & {}^M_S R \end{bmatrix} \begin{Bmatrix} f_S \\ m_S \end{Bmatrix} \quad (17)$$

A wrench error vector $F_E = F_C - F_M$ is formed from the difference of the sensed and commanded wrenches in the *MRF*. The wrench error is converted to a rate $\dot{X}_F = K_F F_E$ which is added to the user-commanded rate. This rate drives the manipulator motion so the desired force is achieved continuously. The diagonal gain matrix K_F has units *m/Ns* and *rad/Nms* for translational and rotational terms, respectively. Since 0 moment is always commanded by the PHANToM the Merlin end effector will automatically orient itself for minimal Cartesian contact wrench and misalignments. This is called *force/moment accommodation (FMA)* (Williams et.al., 1996; Williams and Murphy, 1997b).

The PHANToM allows the user to feel the contact forces exerted on the environment by the Merlin. The required transformation is (Craig, 1989):

$$\tau_T = J_{HC}^T F_W \quad (18)$$

where τ_T is the vector of PHANToM joint torques required to feel the task forces and J_{HC} is the PHANToM Jacobian matrix. The task forces F_M are scaled by matrix gain K_{FR} and sent to the PHANToM.

Naturally-Transitioning Rate-to-Force Controller (*NTRFC*)

The Naturally-Transitioning Rate-to-Force Controller (*NTRFC*) is a combination of rate control and force control enabled on all axes at all times. The *NTRFC* can be applied to control any manipulator(s) with wrist-mounted force/torque sensor, rate inputs, and contact with the environment. The concept was developed heuristically at NASA Langley Research Center (Williams et.al., 1996) and demonstrated to be very effective in experiments (Willshire et.al., 1992). The system behaves as a rate controller in free motion and as a force controller in contact. The transition requires no mode changes, logical switches, or gain changes in the controller software or hardware and thus is termed a natural transition. The *NTRFC* concept was improved during summer 1998 at the HSF lab. Rigorous modeling was performed and design procedures were developed. The *NTRFC* control diagram is given in

Fig. 10. The *NTRFC* has been implemented in the *PHANToM/Merlin* system and human-factors evaluation experiments are planned to test its effectiveness.

In free motion, the displacement of the operator's hand on the master device is proportional to the Cartesian rate of the manipulator end-effector. In contact, the displacement of the operator's hand on the master device is proportional to the force exerted by the manipulator end-effector on the environment. No change in control mode is necessary since rate control and the *FMA* algorithm act simultaneously on all Cartesian axes. If force reflection is enabled then the force of the operator's hand on the master device is proportional to the force exerted by the manipulator end-effector on the environment. This enables very effective *telepresence*. A journal article submission has been prepared describing the *NTRFC* (Williams et.al., 1998).

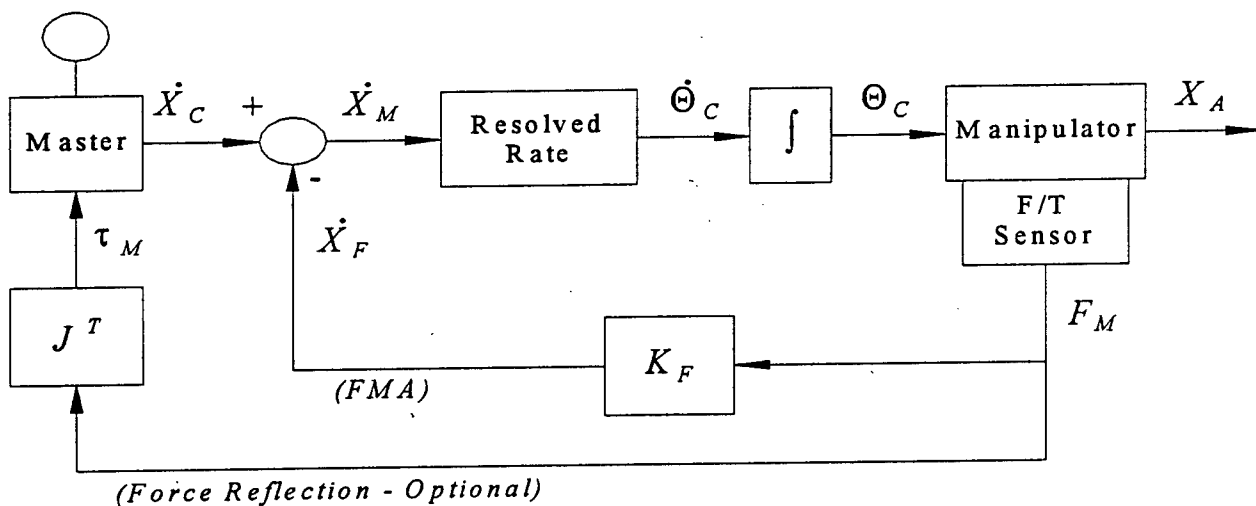


Figure 8. The NTRFC algorithm.

IMPLEMENTATION

The PHANToM was connected to a standard pc through an ISA card. The ISA card read the encoders and software was provided that changed these values into radians. Software was also provided that calculated the

motor torques required to provide the desired combination of CFRTC and force-reflection to the user. The Merlin was connected to a controller that used linear independent PID joint control given a commanded angle. The pc was equipped with two serial lines, one connected to the Merlin high-speed host interface, and the other connected to the JR3 force/torque sensor. The force/torque sensor used an interrupt to ensure that the most current readings were being sent to the pc. A special timing program was written that increased the accuracy of the internal timer from 54msec to 1msec. This was done for smooth control of the Merlin.

When the program was first run the user had the choice of rate or pose control, and each of these could be done with or without force reflection. If rate control was chosen then force/moment accommodation could also be used. The user can also choose whether they wanted CFRTC or not. This will focus on rate control with force reflection, force/moment accommodation, and CFRTC. The Merlin would then move from the calibrate position of straight out to the start position of an elbow down position similar to that shown in figure 3. Then the mass and center of gravity of the L-shaped peg holder were found by taking readings of the force/torque sensor at specific orientations. This is crucial because the mass of the peg holder and sensor bias must be subtracted from the current sensor reading to ensure that the sensor is reading only the interaction forces of the Merlin on the environment.

A while loop would then be entered that could be exited when the user hit any key. The programs in the loop were:

Get_JR3_reading- This gets the most recent force/torque reading from the buffer.

Pose_diff- This calculates the difference in PHANToM pose from the calibrated position to the new position.

M_fwdkin- This calculates the position of the Merlin based on the DH parameters.

Aux_force- This calculates the CFRTC and sums all of the forces that are to be felt by the user.

Force_ref- This calculates the force exerted by the Merlin on the environment in MRF coordinates.

Force_ctrl_ This calculates the rate commands due to the FMA algorithm.

Mer_jac- This calculates the Merlin jacobian.

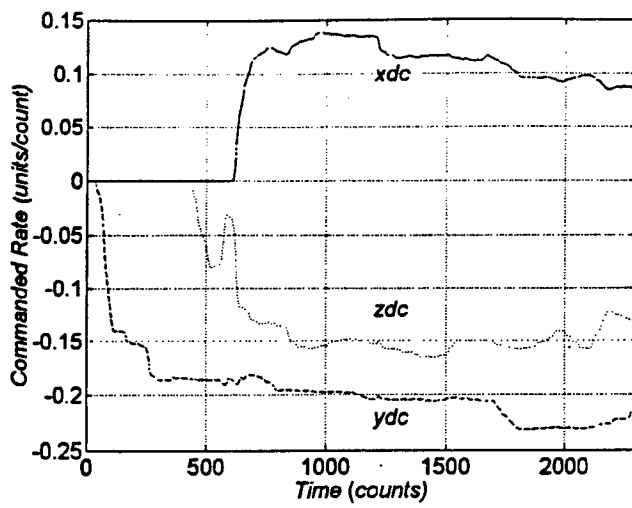
Mer_rate- This calculates the rate commanded by the user including the deadband and sums it with the rate commanded by the Force_ctrl program. This is the total rate that is commanded to the Merlin. The outputs are the 6 commanded joint angles.

Merlin_move- This takes the 6 commanded joint angles and sends them to the Merlin via the serial connection.

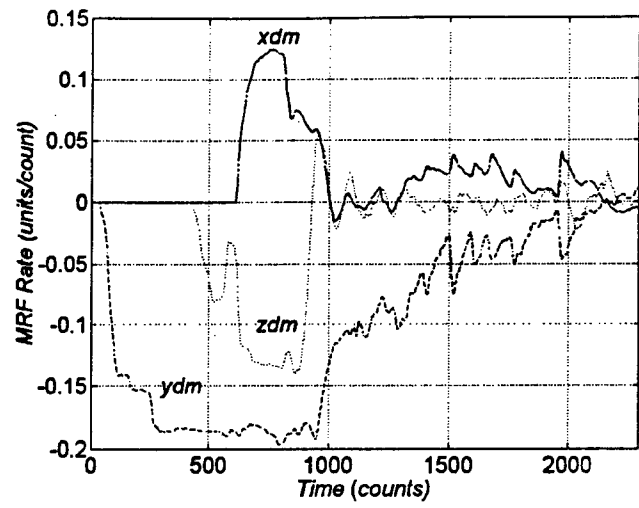
Preliminary Data

Some preliminary data has been collected, and the initial results look promising. Figures 9a-d show experimental results for the NTRFC in the *PHANToM/Merlin* system. Figure 9a shows the human-commanded input rates $\dot{X}_C$.

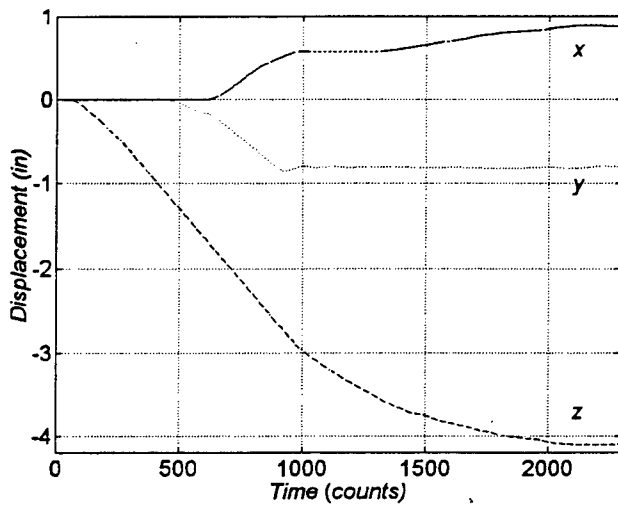
Figure 9b shows the MRF rates $\dot{X}_M$. Figure 9c shows the resulting Cartesian displacements of the tip of the peg (origin of the *MRF*), and Figure 9d shows the associated sensed forces, transformed to the peg tip. This motion involved driving the tip of the peg into the task board corner, input rates held approximately constant by the operator were required in all three Cartesian directions. At different times, the peg contacted the task board in different directions. These experimental results demonstrate the natural transition between rate control in free motion and force control in contact, for all three axes. Though some of the data looks unsteady, the *feel* of the task was quite smooth.



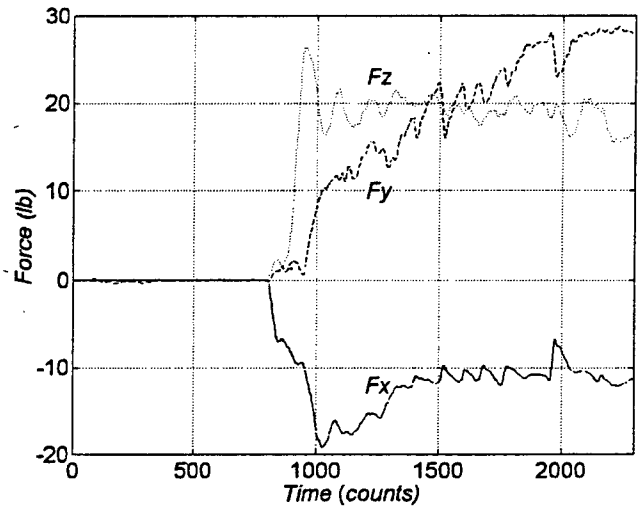
a. Commanded Rates



b. MRF Rates



c. Cartesian Displacements



d. MRF Forces

Figure 9. NTRFC Experimental Results

CONCLUSION

A powerful and general control architecture is presented for real-time, sensor-based, rate-based, shared control of general telerobotic systems including force-reflecting hand controllers. This architecture has been implemented in hardware and the work will continue. This work is the first time the *PHANToM* haptic interface has been used for force-reflecting teleoperation of a kinematically dissimilar slave manipulator to increase telepresence in remote operations. The initial results from the NTRFC look to be promising. Human-factors experiments are planned to fully evaluate the *PHANToM/Merlin* system in remote tasks in the HSF Lab. Simulations of both a 1-P and 3-R robot controlled by the NTRFC were also created for the HSF lab. A method for optimizing the NTRFC gains has also been established. For further information on the simulations and optimization routines see (Williams et al, 1998)

ACKNOWLEDGMENTS

The author gratefully acknowledges support for this research from the Air Force Office of Scientific Research and Armstrong Laboratory in the Summer Graduate Research Program. Many thanks also to Major Tim Adam for support and facilities in the Human Sensory Feedback Lab at Wright-Patterson AFB. The following researchers helped make the summer work more productive and enjoyable: Lieutenant Kurtis Johnson, Dr. Dan Repperger, Jim Berlin, and Lieutenant Mike Krier. The author thanks Dr. Robert Williams II for the guidance provided.

References

- American Robot Corporation, 1985, "Service manual for the System II Merlin Intelligent Robot".
- M.D. Bryfogle, 1990, "Force Reflection Algorithms for Exoskeleton Controllers", AAMRL-TR-90-090.
- J.J. Craig, 1989, **Introduction to Robotics: Mechanics and Control**, Addison Wesley, Reading, MA.
- C.J. Hasser, 1995, "Tactile Feedback for a Force-Reflecting Haptic Display", AL/CF-SR-1996-0134.
- J.M. Henry, 1998, "PHANToM/Merlin Force-Reflecting Teleoperation: Implementation", Final Report, AFOSR Summer Graduate Student Research Program, AFRL/HECP.
- N. Hogan, 1985, "Impedance Control: An Approach to Manipulation", ASME Journal of Dynamic Systems, Measurement, and Control.
- M.Z. Huang, 1993, "Efficient Coordination of an Anthropomorphic Telemanipulation System", AL/CF-TR-1995-0120.
- T.R. Kane, P.W. Likins, and D.A. Levinson, 1983, Spacecraft Dynamics, McGraw-Hill.
- M.A. Murphy, 1997, "Implementation of FreFlex/Merlin Teleoperation", Final Report, AFOSR Summer Graduate Student Research Program, Armstrong Laboratory.
- R.P. Paul, 1981, **Robot Manipulators**, MIT Press, Cambridge, MA.
- M. Raibert and J.J. Craig, 1981, "Hybrid Position/Force Control of Manipulators", ASME Journal of Dynamic Systems, Measurement, and Control.
- D.W. Repperger, E.L. Scarborough, and T.L. Chelette, 1991, "Construction of a Dual Axis Force Reflection Stick and Test Station", AL-TR-1992-0041.
- D.W. Repperger, 1991, "Active Force Reflection Devices in Teleoperation", IEEE Control Systems Magazine, 52-56, January.
- D.W. Repperger, C.A. Phillips, and T.L. Chelette, 1995, "A Study on Spatially Induced "Virtual Force" with an Information Theoretic Investigation of Human Performance", IEEE Transactions on Systems, Man, and Cybernetics, 25(10):1392-1404.
- D.W. Repperger, 1995, "Biodynamic and Spasticity Reduction in Joystick Control via Force Reflection", AL/CF-TR-1995-0152.
- L.B. Rosenberg, 1992, "The Use of Virtual Fixtures as Perceptual Overlays to Enhance Operator Performance in Remote Environments", AL/CF-TR-1994-0089.
- L.B. Rosenberg, 1993, "The Use of Virtual Fixtures to Enhance Operator Performance in Time Delayed Teleoperation", AL/CF-TR-1994-0139.
- SensAble Devices, Inc., 1995, "PHANToM Haptic Interface: Installation & Technical Manual".
- D.E. Whitney, 1969, "Resolved Motion Rate Control of Manipulators and Human Prostheses", IEEE Transactions on Man-Machine Systems.
- R.L. Williams II, J.M. Henry, D.W. Repperger, and M.A. Murphy, 1998, "Naturally-Transitioning Rate-to-Force Control", submitted to the ASME Journal of Dynamic Systems, Measurement, and Control.
- R.L. Williams II, 1997, "Force-Reflecting Teleoperation Control Architecture", Final Report, AFOSR Summer Faculty Research Program, Armstrong Laboratory.
- R.L. Williams II, 1997, "Cartesian Control of Force-Reflecting Hand Controllers", Conference on Applied Mechanisms and Robotics, Cincinnati, OH.
- R.L. Williams II, 1997, "Cable-Suspended Haptic Interface", International Journal of Virtual Reality, in press.
- R.L. Williams II, F.W. Harrison, and D.I. Soloway, 1997, "Shared Control of Multiple-Manipulator, Sensor-Based Telerobotic Systems", 1997 IEEE Intl Conference on Automation and Robotics, Albuquerque, NM.
- R.L. Williams II, F.W. Harrison, and D.I. Soloway, 1996, "Naturally-Transitioning Rate-to-Force Controller for Manipulators", 1996 IEEE International Conference on Automation and Robotics, Minneapolis, MN.
- K. Willshire, F.W. Harrison, E.F. Hogge, R.L. Williams II, and D.I. Soloway, 1992, "Results of Telerobotic Hand Controller Study Using Force Information and Rate Control", AIAA Paper 92-1451, AIAA Space Programs Conference, Huntsville.

ISSUES IN STEADY-STATE VISUAL EVOKED RESPONSE BASED CONTROL.

**Keith S. Jones
Doctoral Student
Department of Psychology**

**University of Cincinnati
Cincinnati, OH 45221-0627.**

**Final Report for:
Graduate Student Research Program
Alternative Control Technology Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC
and
Alternative Control Technology Laboratory**

September, 1998

ISSUES IN STEADY-STATE VISUAL EVOKED RESPONSE BASED CONTROL.

Keith S. Jones
Doctoral Student
Department of Psychology
University of Cincinnati.

Abstract

Electroencephalographic (EEG)-based control devices are one of several emerging technologies that will provide operators with a variety of new hands-free control options. In general, EEG-based control translates brain electrical activity into a control signal. This report describes two experimental evaluations of a new steady-state visual evoked response-based control. With this control, the luminance of selectable items on a computer display was modulated at different frequencies. The operator's choice between these items was identified by detecting which frequency pattern was dominant in the visual evoked brain activity. The first study was designed to characterize the performance of this system, in addition to evaluating two candidate control frequencies. The results are encouraging. Participants were able to use this form of EEG-based control and their performance was stable. In fact, participants averaged over 90 percent correct selections.

Having established the efficacy of this type of control, it would be reasonable to develop and test enhancements to the current system. One possible improvement was an automated calibration procedure. Prior to its development, system parameters had to be adjusted manually to tune aspects of the system to particular characteristics of the user's electrical brain activity. With this automated procedure, an algorithm would perform these adjustments. A second experiment was developed to address the usefulness of this procedure. It was found that the automated procedure consistently set the two calibration parameters below where the manual calibrator would have set them. Due to the specific nature of the parameters, this could have resulted in faster selection times (a positive result) and/or an increased false selection rate. Selection times were found to be faster, however, the false selection rate was not increased. Due to the positive benefit and lack of negative consequences associated with its use, the automated calibration procedure was deemed useful.

ISSUES IN STEADY STATE VISUAL EVOKED RESPONSE BASED CONTROL.

Keith S. Jones

Introduction¹

A great deal of research has focused on developing alternative control technologies. These efforts have produced control devices whose input stems from electroencephalographic (EEG) or electromyographic (EMG) data, eye and head position data, along with recognition of the spoken word. These devices promise to provide operators with a variety of new control options (McMillan, Eggleston and Anderson, 1997) and may be particularly suited for hands-busy applications or when conventional controls are less accessible. In fact, these hands-free control devices may be especially useful to exploit the unique advantages afforded by wearable computers. For example, aircraft maintenance technicians working on a noisy flight line could access schematics on head-mounted displays using EEG-based control, while keeping their hands dedicated to the repair task.

The present research has focused on development of an EEG-based control. In general, EEG-based control translates selected aspects of the brain's electrical activity into a control signal. However, there are a number of different ways this can be accomplished. One approach requires operators to bring these electrical signals under conscious control (Nasman, Calhoun and McMillan, 1997). Successful use of this type of voluntary control requires large amounts of biofeedback-based operator training. A different approach employs naturally occurring brain responses that correspond to sensory processing, cognitive activity or motor control. One example of this approach capitalizes on the "P300", a brain response that varies with stimulus probability and task relevance. With careful design of the task format and procedures, it is possible to use the natural variance of the P300 for task control (Farwell and Donchin, 1984).

The approach evaluated in this study is also based on naturally occurring brain responses. Specifically, this method employs a brain response known as the steady-state visual evoked response (SSVER). The SSVER is typically generated by a luminance modulated (flickering) stimulus. It is characterized by increased power at the frequency of that stimulus (Regan, 1989). To implement this control, the luminance of selectable items on a computer display is modulated each at a specific frequency. The system monitors the brain's electrical activity for a SSVER at the frequency of each selectable item while the operator fixates on the desired flickering item. A selection occurs if the system detects a SSVER at the frequency of a particular item.

¹ Portions of text for the Introduction and the description of Study I are drawn from a paper to be presented at the 42nd annual meeting of the Human Factors and Ergonomics Society (1998).

Study I: Evaluation of SSVER-based Control

This study had three objectives. First, this study was to confirm that control with this device is indeed based on the SSVER. Second, performance of operators using this control device for a two choice selection task was characterized. Finally, performance with two candidate control frequencies was examined, since SSVER amplitude has been shown to vary with evoking stimulus frequency (Regan, 1989).

Method

Participants

Nine individuals participated (4 male, 5 female). One participant (male) was excused. Pre-experimental testing suggested that this individual was incapable of generating a SSVER at either of the control frequencies. Prior experience with brain-actuated control was not a selection factor since only naturally occurring SSVER responses were employed. All participants were required to have normal or corrected to normal vision.

Electrical Recording

Silver-chloride-coated, plastic surface electrodes were used to acquire signals from occipital sites O1 and O2 (in accordance with the 10-20 International System). Oz served as the ground. These electrodes were incorporated into a headband. The bipolar signal (O1-O2) was used for control input.

Apparatus

An S75-01 biological signal amplifier (Coulbourn Instrumentation Inc.) filtered and amplified the electrical signals. A 486 desktop PC operating at 120 MHz executed the task and display software. A 15 inch color monitor served as the display. The software timing and data acquisition were controlled by an analog input/output board, manufactured by Scientific Solutions, Inc.

Control Algorithm

Four criteria had to be met in order to select an item. The SSVER amplitude had to be above a threshold value, larger than the average amplitude of the frequencies 1 Hz above and below the item's frequency and larger than the amplitude of the other selectable item's frequency. If these three criteria were met for 0.3 seconds, then a selection occurred.

Experimental Design

Each participant was tested with each of two control frequencies (17.56 and 23.42 Hz). Performance at each of these frequencies was evaluated under two trial types: trials where the

items were luminance modulated and trials where they were not. The only difference between these two trial types was the lack of flicker in the latter condition. The non-flickering condition was intended to confirm that control was based on the SSVER, as opposed to non-evoked electrical activity.

Each participant completed two testing sessions (one per day, at least 24 hours apart). Each testing session consisted of 15 blocks of trials. Selectable items were luminance modulated in 10 of these blocks. The order of the two types of trial blocks was random with the constraint that the participant was presented with the appropriate number of each trial type.

Within each block of trials, participants were asked to attempt 10 selections (trials), five with each of the two control frequencies. Aside from this constraint, selection requests between the two control frequencies were randomized.

Performance was evaluated with two variables: percent correct selections ($\# \text{ correct} / \# \text{ of attempts}$) and the mean time to correctly select an item. Subjective data were collected with a questionnaire.

Procedure

Each testing session lasted approximately one hour. Alcohol was used to clean the participant's scalp and aloe vera gel was applied to the electrodes prior to their positioning. During test sessions, participants were seated in front of a computer display (Figure 1). The top of the screen was positioned approximately at eye level. Participants were familiarized with the experimental task and were instructed to minimize blinking and physical movements during a trial. Prior to data collection, the control algorithm parameters were calibrated for the particular characteristics of each participant's SSVER.

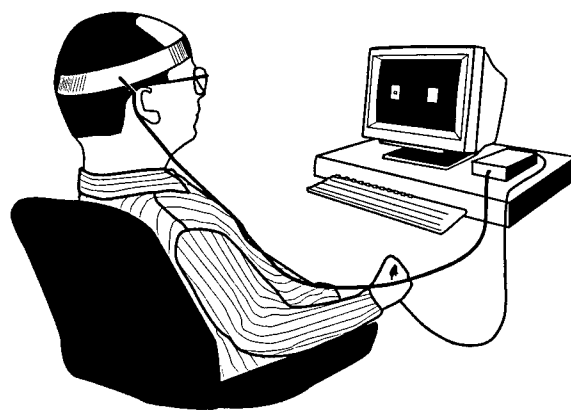


Figure 1. Schematic of experimental apparatus.

Prior to each trial, two vertically oriented rectangular selectable items were presented on the display. Each selectable item was 2.9 cm wide by 3.8 cm tall and they were separated by 10.3 cm. The participant pressed a hand-held button to initiate the trial. For luminance modulated trials, pressing the button initiated the flicker. The left and right items flickered at 23.42 and 17.56 Hz, respectively. For non-flickering trials, the items remained static after the button press. Prior to each trial, a small yellow square appeared inside one of the two items indicating which item should be selected. The participant's task was to select the item that contained the yellow box by fixating upon it. A red border surrounded the selectable item when the system detected that a SSVER satisfied the amplitude requirements of the control algorithm. A selection occurred if the participant maintained the required SSVER for .3 seconds. If a selection occurred (correct or incorrect), or the participant failed to make a selection within 10 seconds, the trial ended and the system reset itself for the next trial. Progression through the trials was self-paced, with the participants pressing a button to initiate each trial. The minimum delay between trials was 3 seconds, however, more time was allowed if needed. At the completion of each block of trials, participants received feedback on their performance (percent correct and average selection time for the previous trial block).

Results and Discussion

Prior to analysis, the data were collapsed across trials. This produced an average for each frequency within each trial type. These averages were computed per session for each dependent variable.

Data for each dependent variable were subjected to separate 2 (trial type) x 2 (session) x 2 (control frequency) repeated measures ANOVAs. The sole significant source of variation in both analyses was the main effect for trial type ($F(1,7) = 87.7, p < .05$ and $F(1,7) = 75.5, p < .05$, for percent correct and selection time, respectively). These results will be addressed in more detail in the following sections.

Confirmation of SSVER-Based Control

Prior to characterizing performance with this new control device, it would be prudent to confirm the locus of operator performance. This device has been referred to as SSVER-based control. However, it could be argued that operator performance is due to non-evoked electrical activity. The following analyses were intended to address this issue.

Figure 2 presents mean percent correct selections versus mean selection time for both the flickering and non-flickering conditions. As can be seen in the figure, performance with flickering items was characterized by high accuracy and relatively low selection times. On the

other hand, non-flickering trials were characterized by low accuracy and relatively high selection times. These results, together with the significant Trial Type effect, suggest that performance under the flickering and non-flickering trials was qualitatively different.

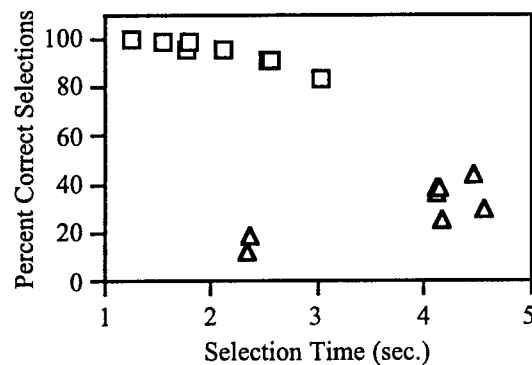


Figure 2. Mean percent correct selections versus mean time to correctly select an item. Flickering and non-flickering trials are represented by squares and triangles, respectively. Each data point corresponds to one participant.

As further confirmation of the SSVER as the locus of performance, the each participant's recorded brain electrical activity was subjected to Fast Fourier Transformations. This provided an estimate of amplitude across frequencies. Recall that the SSVER is characterized by increased amplitude at the flicker frequency. Exemplars from one participant are presented in Figure 3. Each plot represents the average of five two-second epochs.

It is evident in these plots that SSVER-like responses are present at the two control frequencies in the two luminance flicker conditions (top and middle panels). However, this response is absent in the data from the non-flickering condition (bottom panel).

These results suggest that control with this device was dependent on the participants' SSVER. Without the flickering items, and the resultant SSVERs, participants demonstrated very poor control.

Performance with SSVER-Based Control

During flickering trials, participants averaged 92 percent correct selections (range = 83 to 99%) with an average selection time of 2.1 seconds (range = 1.24 to 3.02). The ANOVAs indicated that performance was consistent across sessions. The Session main effects and associated interactions were not significant, $p > .05$ in each case. This was true for both dependent variables (see Figure 4).

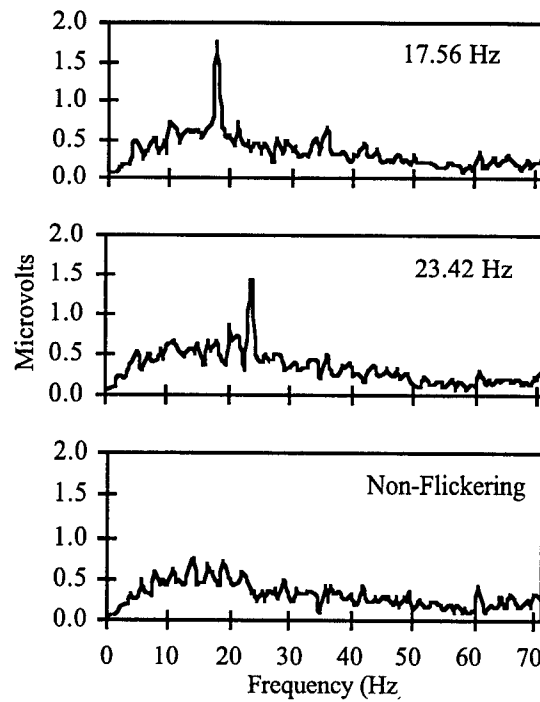


Figure 3. Exemplars of amplitude versus frequency for flickering and non-flickering trials from one participant.

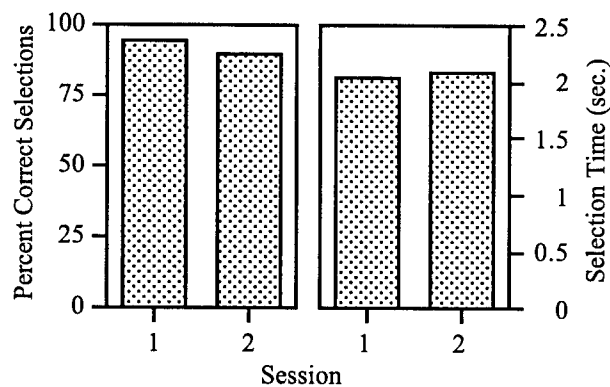


Figure 4. Mean percent correct selections and mean time to correctly select an item versus session.

The ANOVAs also indicated that performance did not significantly differ depending on which of the two control frequencies was used ($p > .05$, for each main effect and associated interactions).

Participants' Opinion of SSVER-Based Control

After the second testing session, participants completed a debriefing questionnaire. This was aimed at assessing the participants' impression of their ability to utilize the SSVER-based control to select items on a computer display.

Participants were asked to rate their overall ability to make selections with the flickering items on a 11-point scale ranging from 0 ("no control") to 10 ("total control"). Participants' ratings ranged from 7 to 10, with a mean of 8.75. This suggests that, on average, participants felt that they had a relatively high level of control. Their impressions seem consistent with the performance results.

In addition, participants were asked to rate how often the task required hard concentration. This rating was made on a 5-point scale, ranging from 1 ("always") to 5 ("never"). All participants responded either 3 or 4, indicating that the task "sometimes" or "rarely" required hard concentration. It may be that with additional exposure to this type of control, participants would find less concentration required.

Participants were also asked to comment on the use of the red feedback border. Recall that the red border surrounded the selectable item when the required SSVER was detected. Participants responded that the red border was helpful when it surrounded the correct item, but distracting around the incorrect item.

Conclusions

The results of this study are encouraging. Overall, participants were able to perform at a high level of accuracy with relatively low selection times. In addition, it appears that their ability to use this controller was stable across days.

Comparisons of SSVER-based control with other types of controls suggest that speed and accuracy values are in similar ranges. Admittedly, this is a coarse comparison since there are differences in the details of the selection tasks. For example, Calhoun, Janson and Arbak (1986) asked participants to select switches located on the front panel of a cockpit simulator. This task was performed manually, as well as with eye-line-of-sight-based control. They found that eye-line-of-sight-based switch selection averaged 1.78 seconds while manual selection averaged 1.72 seconds. Epps, Snyder and Muto (1986) found times ranging from 1.1 to 2.7 seconds for mouse-driven selection of targets at various sizes and distances. Card, English and Burr (1978) found that the time required to select a text element with a mouse averaged 1.66 seconds with a 5% error rate.

In comparison, our participants averaged over 90 percent correct selections with average selection times slightly over 2 seconds. Therefore, the speed and accuracy of our participants

compares favorably with that of other types of controls. In addition, this is the initial version of this new control device. We are optimistic that performance improvements can be realized with further research and development. To that end, our laboratory is investigating improved SSVER detection methods using advanced signal processing and neural network solutions. We are also examining stimulus parameters (size, shape, color and frequency) that may influence SSVER strength.

It is important to note that any performance limitations may be offset by the unique advantages of SSVER-based control. For example, this type of device would allow hands-free control where manual control is not possible or less convenient. Training is not a design issue since this control is based on naturally occurring brain responses. In addition, SSVER-based control can be implemented with inexpensive components on a desktop PC-based system. This and other forms of fixation-based selection are a more "direct way" of selecting displayed objects since an operator usually looks at an item before initiating a control input (Glenstrup and Engell-Nielsen, 1995). In effect, through SSVER-based control, the advantages of eye gaze-based control can be realized with less expensive and obtrusive components.

Despite these potential advantages, several integration issues need to be addressed. First, just as in eye-tracker-based systems, control input is "on" all the time. Thus, either a consent response or some means of engaging the control is required. This may slightly increase selection time (e.g., 400 msec for a manual consent; Calhoun, Arbak and Boff, 1984). Ideally, a hands-free consent response or flickering mode on/off input could be developed to eliminate unintended selections produced by ongoing electrical activity. Second, SSVER-based control requires selectable items to be modulated in some fashion. Therefore, the system needs to be designed to minimize the distracting effects of this modulation. Finally, the number of selectable items that can be generated and reliably activated and how spatially separated they must be needs further evaluation.

In summary, these results demonstrate the feasibility of a SSVER-based control approach. The advantages of a hands-free device that does not require operator training or expensive, bulky hardware make this an attractive candidate controller, worthy of further development and evaluation.

Study II: Evaluation of SSVER-based Control Automated Calibration Software

The control evaluated in Study I utilizes a selection algorithm that requires three parameters to be met (the threshold, amplitude ratio and relative amplitude factors; see previous discussion for details). Prior to use, two of these parameters must be calibrated to characteristics of each individual user. Additionally, a fourth value, used in the adaptive threshold logic, needs to be calibrated (e.g. the threshold adjustment factor). Historically, the calibration of these parameters has been performed manually. However, it would be ideal to automate this

procedure. To this end, software was developed that will perform this calibration. This study evaluated the effectiveness of this automated calibration software. Participants were asked to perform a two item selection task. This was done separately with calibration settings based on the automated calibration software and a manual calibrator. Data was collected over a two-day period with each calibration technique tested twice per day; once with and without a visual fixation point (labeled "box on" and "box off" in the following figures). The issues revolving around the fixation point will not be addressed in this report.

Judgment Criteria

In order to evaluate the effectiveness of this procedure, a comparative standard was needed. That is, in order to determine if the automated calibration software was performing adequately, we needed to specify what the appropriate parameter settings should be. In the absence of a clear indisputable criterion, the settings derived from the manual calibrations were used in this capacity. Therefore, the efficacy of the automated software was evaluated with respect to how closely they matched the manual calibration settings. Prior to this, the manual calibration settings were examined to determine if the decision to use them as a comparative standard was appropriate.

Manual Settings Evaluation

The manual calibrator is required to monitor characteristics of the user's EEG and make algorithm parameter adjustments based on this information. If the calibrator is accurately performing this task, their parameter adjustments should converge on a given value that would be optimal for a particular participant. To examine this, the three algorithm parameters were plotted separately for each participant (see Figures 1 - 3). As can be seen in the figures, the calibration values did tend to converge over trials. Given this, the choice of the manual calibration settings as a comparative standard seemed warranted.

Automated Calibration Evaluation

With the manual calibration values as a standard, the efficacy of the automated calibration software was evaluated. Initially, the data were examined for consistency. Data from each session were subtracted to provide a difference score. These difference scores were calculated for the manual and automated conditions separately. The variability of automated scores was generally less than or equal to the variability of the manual scores (see Figure 4). Such a result implies that the automated procedure produces sufficiently consistent results.

Following this, differences between the manual and automated values were calculated for each pair within session and fixation point conditions. In this case, a positive result indicates that the manual values were set higher than the automated values. A negative result suggests

that the automated values were set higher than the manual settings. As can be seen in Figure 4, the automated values were generally lower than the manual settings with regard to the threshold and threshold adjustment factor. The amplitude ratio values fell roughly symmetrically around the zero difference point.

This pattern of results has potentially positive and negative consequences. One might expect that the lower threshold and threshold adjustment factor values might result in faster selection times. However, it is also possible that this combination could result in more incorrect selections and more time spent in the incorrect state (i.e. selecting the wrong stimulus). Each of these possibilities will be explored in the following sections.

As a result of the aforementioned pattern of results, selection times might be faster with the automated as opposed to the manual calibration procedure. Figure 5 presents percent correct and average selection time data for the two conditions. As can be seen in the figure, the selection times were consistently faster for the automated calibration. Therefore, it appears that the incongruous nature of the automated and manual calibration settings has a positive effect on selection time.

The lower automated calibration settings also could result in an increase in false selections. Across both fixation point conditions and sessions, the automated trials averaged 6.75 incorrect selections, while the manual trials averaged 7 incorrect selections. Therefore, it doesn't appear that the automated calibration's lower settings resulted in an increased false selection rate.

The final possibility concerns an increase in the amount of time spent in the incorrect state (i.e. activating the wrong stimulus). Across both fixation point conditions and session, the time spent in the incorrect state for the automated trials averaged 2.94 seconds while manual trials averaged 3.14 seconds. Therefore, it appears that the aforementioned pattern of results did not adversely increase the amount of time spent in the incorrect state.

The results of this experiment indicate that the automated calibration procedure consistently set the threshold and threshold adjustment factor values lower than the manual calibrator. This had several possible consequences for performance: it could result in faster selection times, an increased incorrect selection rate and/or an increase amount of time spent in the incorrect state. Analyses targeting these issues suggested that selection time was in fact decreased (a positive result). Additionally, the levels of incorrect selections or time spent in the incorrect state were not increased. Due to the lack of negative consequences associated with the differences between the manual and automated calibration settings, no major changes to the automated calibration algorithms seem warranted.

Conclusions

The two studies described in this report support the ongoing design and testing of a new control modality. Study I tested the efficacy of this newly developed technology. The results of

Figure 1

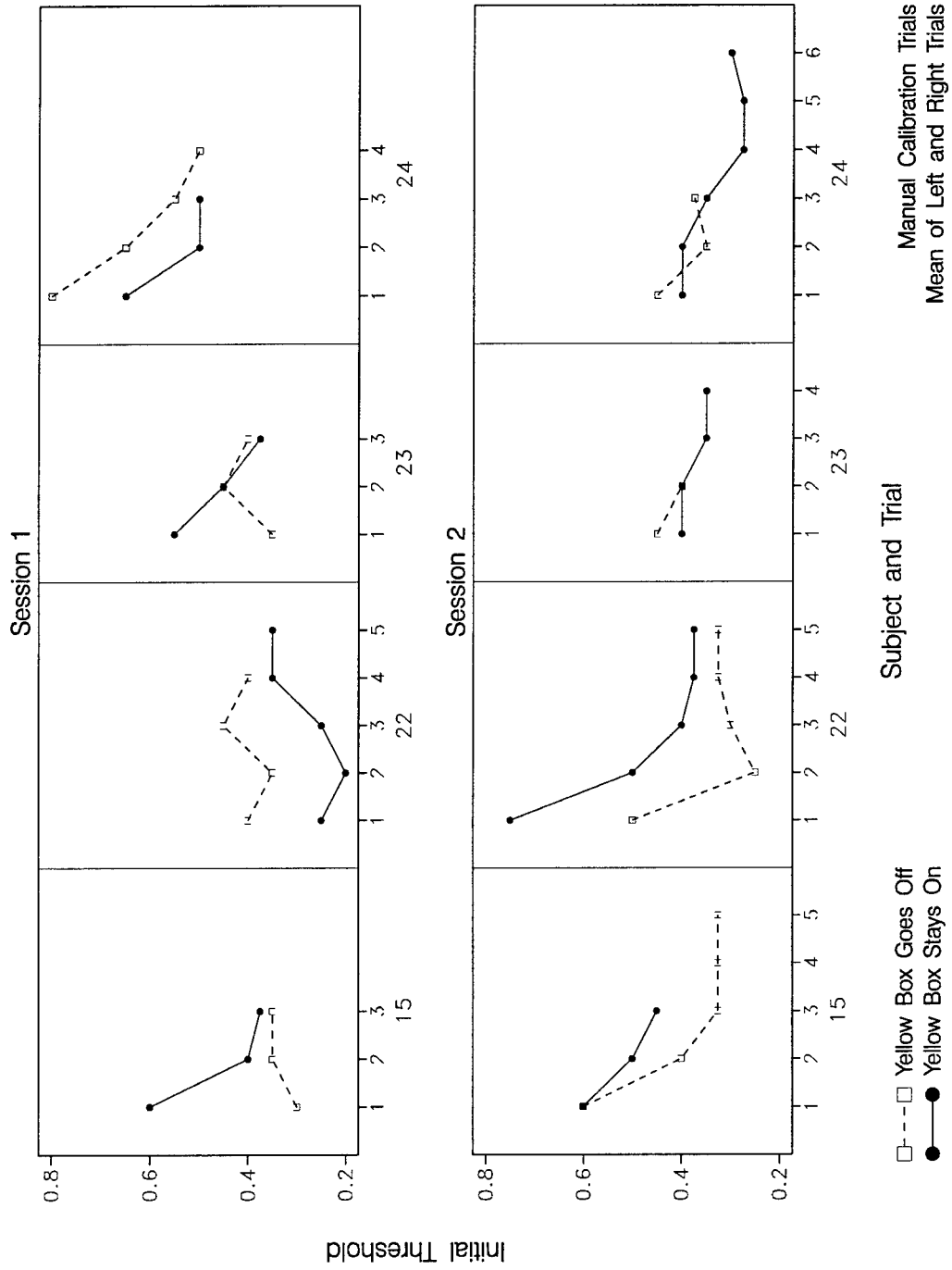


Figure 2

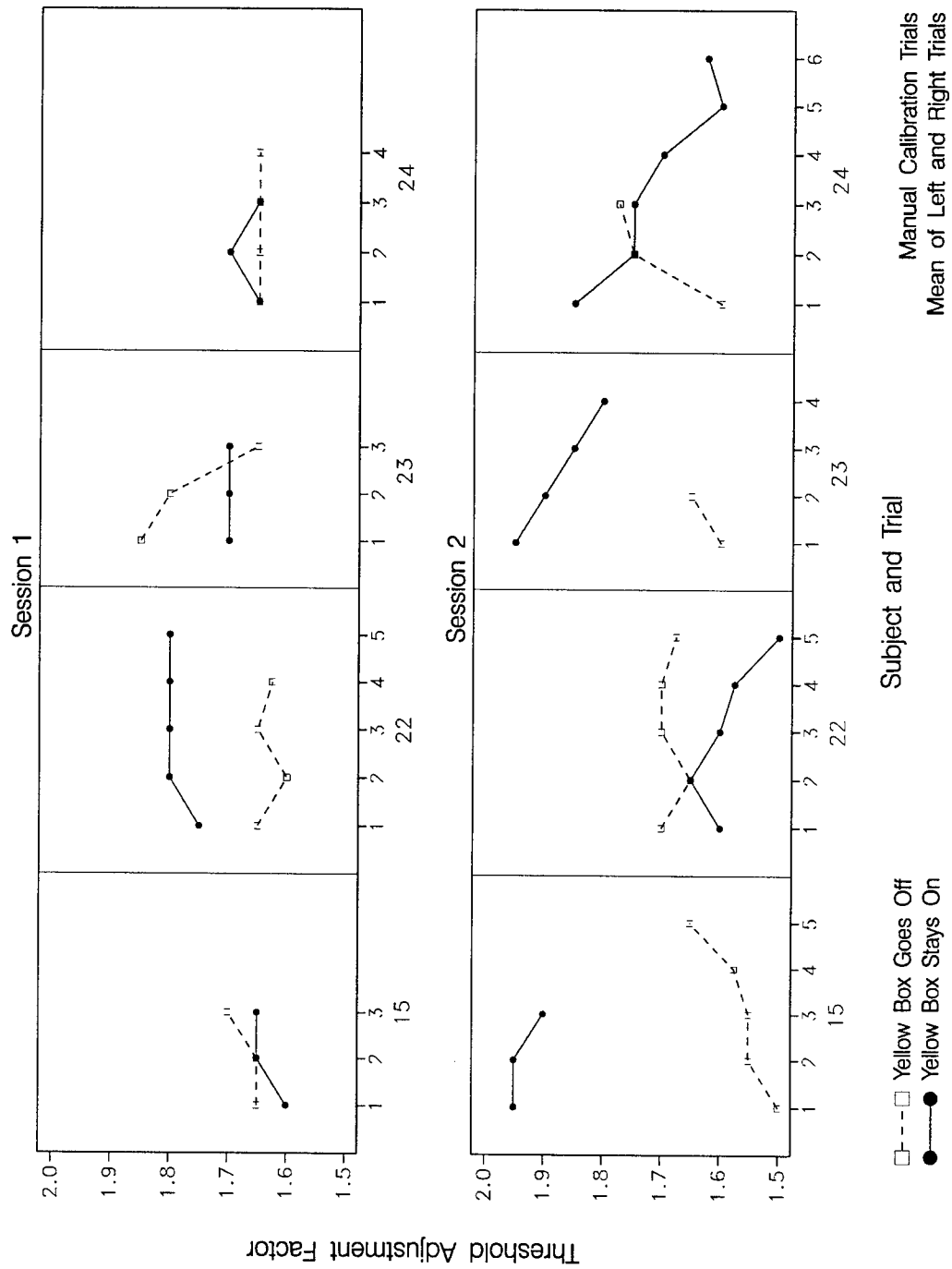


Figure 3

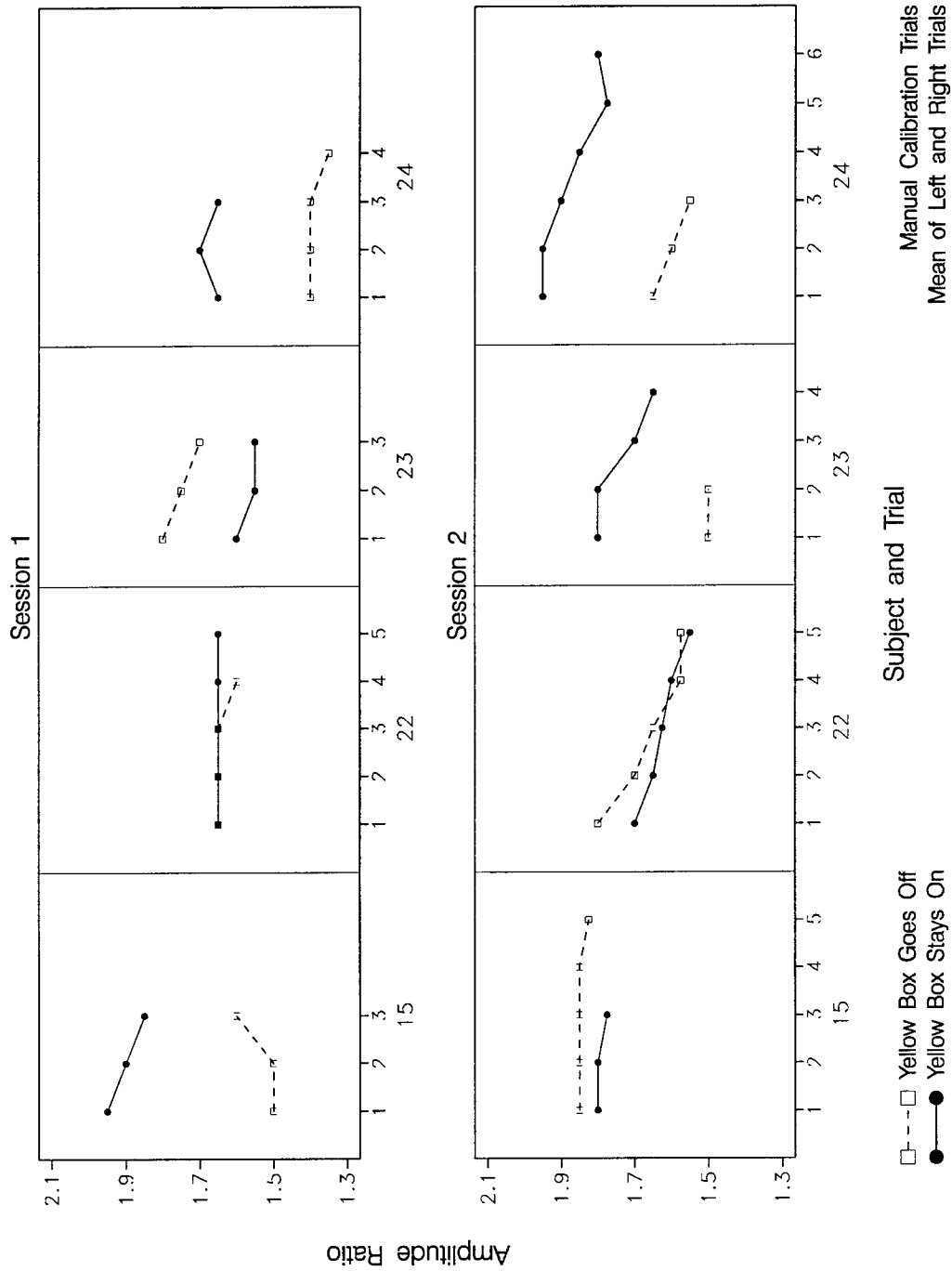


Figure 4

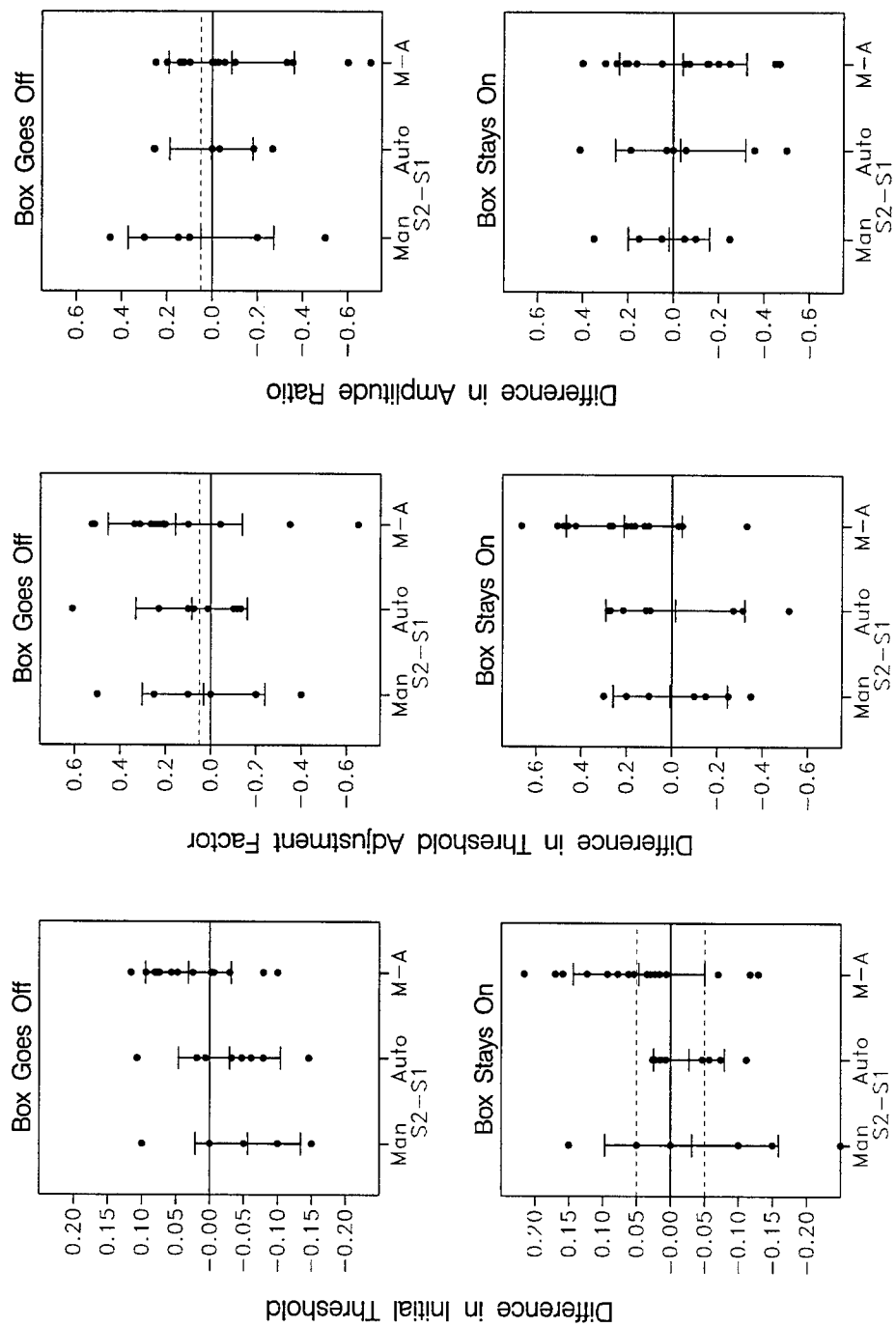
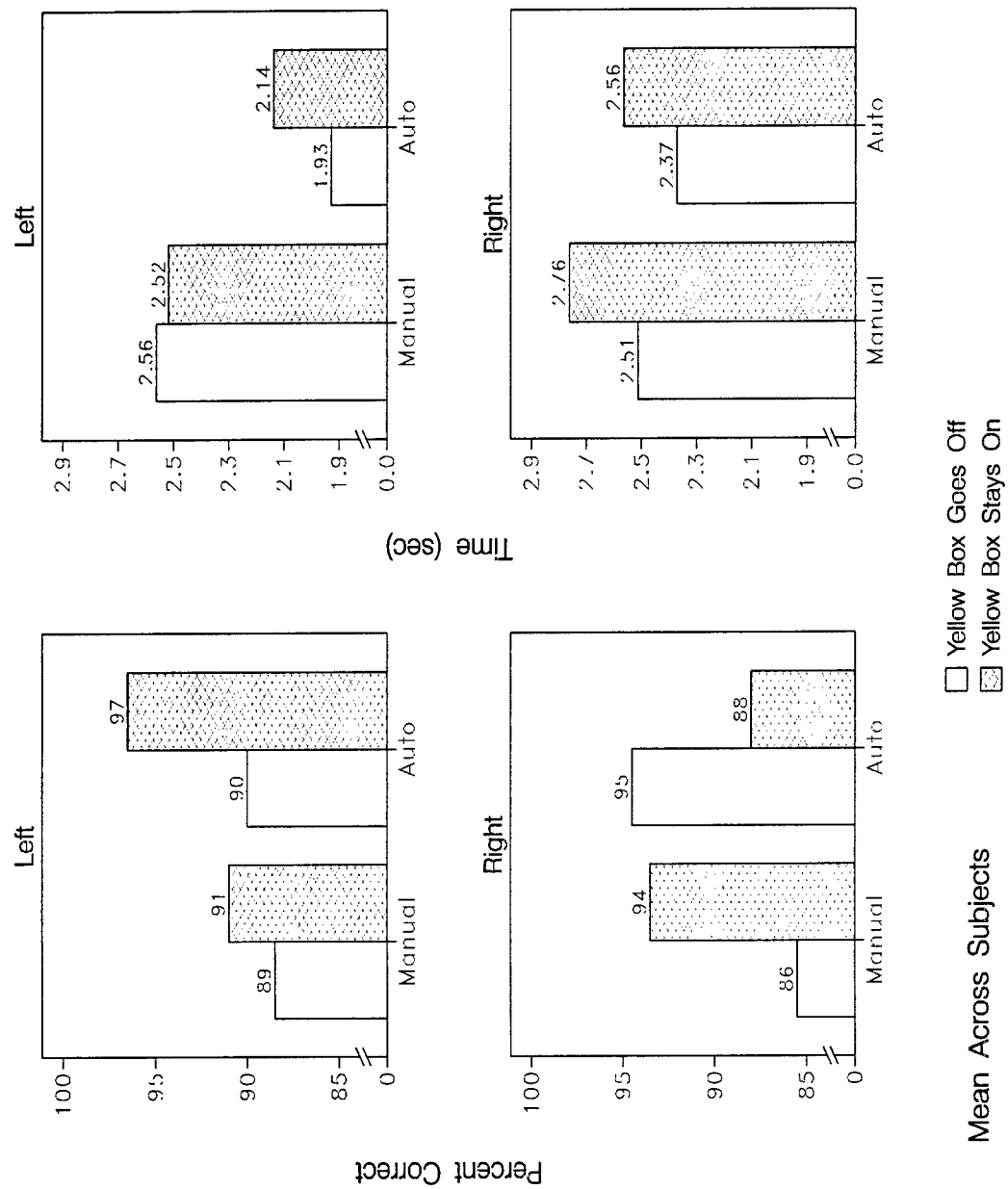
Individual differences with mean $\pm$ std

Figure 5



that investigation suggest that this SSVER-based control has merit. Participants in that study were able to perform quite accurately with low selection times. Furthermore, it was suggested that performance with this control compares favorably with that of other control modalities. Performance with this SSVER-based control averaged over 90 percent correct selections with selection times averaging slightly over 2 seconds. In addition, any performance limitations may be offset by the unique advantages of SSVER-based control.

Having established the usefulness of this control. Study II sought to evaluate an enhancement to the existing system: an automated calibration procedure. The results of this experiment suggest that this procedure consistently set the threshold and threshold adjustment factor values lower than the manual calibration settings. This fact could have resulted in: faster selection times, increased the incorrect selection rate and/or an increased the amount of time spent in the incorrect state. Analyses targeting these issues suggested that selection time was in fact decreased (a positive result). However, the levels of incorrect selections or time spent in the incorrect state were not increased. Due to the lack of negative consequences associated with the differences between the manual and automated calibration settings, the automated calibration software was deemed useful.

In sum, these two studies describe the ongoing testing and development of a novel control device. This research was able to experimentally validate the efficacy of this control. In addition, a system enhancement that has wide implications for usability was evaluated and found to be effective.

REFERENCES

- Calhoun, G.L., Arbak, C.J., and Boff, K.R. (1984). Eye-controlled switching for crew station design. Proceedings of the Human Factors Society - 28th Annual Meeting, 258-262.
- Calhoun, G.L., Janson, W.P., and Arbak, C.J. (1986). Use of eye control to select switches. Proceedings of the Human Factors Society - 30th Annual Meeting, 154-158.
- Card, S.K., English, W.K., and Burr, B.J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step-keys, and text keys for text selection on a CRT. Ergonomics, 21(8), 601-613.
- Epps, B.W., Synder, H.L., and Muto, W.H. (1986). Comparison of six cursor devices on a target acquisition task. Proceedings of the Society for Information Display, 302-304.
- Farwell, L.A. and Donchin, E. (1984). Talking off the top of your head: Toward a mental prosthesis utilizing event-related potentials, Electroencephalography and Clinical Neurophysiology, 70, 510-523.
- Glenstrup, A.J. and Engell-Nielsen, T. (1995). Eye controlled media: Present and future state. Available at: <http://www.diku.dk/~panic/eyegaze/article.html>.
- McMillan, G.R., Eggleston, R.G., and Anderson, T.R. (1997). Nonconventional controls. In G. Salvendy (Ed.), Handbook of human factors and ergonomics, 2nd edition. NY: John Wiley and Sons, pp. 729-771.
- Nasman, V.T., Calhoun, G.L., and McMillan, G.R. (1997). Brain-actuated control and HMDs. In J. Melzer and K. Moffit (Eds.), Head-mounted displays: Designing for the user. NY: McGraw-Hill, pp. 285-310.
- Regan, D. (1989). Human brain electrophysiology. NY: Elsevier.

RESEARCH TECHNIQUES:
A SEARCH FOR CREW RESOURCE MANAGEMENT DOCUMENTS

Christian A. Kijora
Arizona State University, East
Department of Aeronautical Management Technology

Final Report for:
Summer Intern Research Program
Air Force Research Labs
Mesa, Arizona

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

And

Air Force Research Lab
Mesa, Arizona

August 1998

RESEARCH TECHNIQUES:
A SEARCH FOR CREW RESOURCE MANAGEMENT DOCUMENTS

Christian A. Kijora
Arizona State University, East
Department of Aeronautical Management Technology

Abstract

Technology brings us closer to the information we seek and yet, if not used correctly, this same technology can drive a wedge between the knowledge we seek and us. Not all documents can be found via the traditional methods of academic research. Thus many excellent sources, primarily governmental, are left idle while most researchers settle for older, more easily located documents found in the typical library. The following paper describes the techniques used to uncover documents, via Internet and academic library journal search methods, pertaining to Crew / Cockpit Resource Management (CRM) and any related topics. Also included is an annotated bibliography of the more recent and relevant documents.

RESEARCH TECHNIQUES: A SEARCH FOR CREW RESOURCE MANAGEMENT DOCUMENTS

Introduction

During the renaissance, when an individual wanted to obtain information on a given subject he would have to travel about the known world in search of each and every book written on the topic. Upon completion of this task, the scholar could be fairly assured that he had all the information available, in the entire world, on this topic. This would not be possible today, while travel would no doubt be faster, there are too many possible locations and too much information available on broad topics. A new method for gathering information is required and the seeker must be able to narrowly define the scope of information desired. We are a part of a new renaissance called the Information Age. Never before has so much information been available to the population of all developed Countries. With a computer and a modem, anyone anywhere can connect to the Internet and obtain information on any number of subjects. From ancient documents such as the Dead Sea Scrolls (<http://lcweb.loc.gov/exhibits/scrolls/scr1.html>), to live data being collected on NASA's Lunar Prospector as you watch (<http://lunar.arc.nasa.gov/dataviz/locationd.html>.) The Internet is not the only electronic source of data, most modern research libraries have databases on many different subjects contained on Compact Disk (CD) or written to a hard disk. Have I forgotten the traditional visit to the library? Depending on the facilities available, you may not need to go there until it is time to receive the documents. In this paper, I intend to describe the techniques used in building a bibliography of documents relating to Crew / Cockpit Resource Management (CRM) topics.

Methodology

The best way to find academic literature is to visit your local research library. Gone are the days of card files, today libraries use computerized databases to list titles. Because this is so, it is not always necessary to actually physically go to the facility, as is the case with the Arizona State University (ASU) research libraries. It is possible to use all the library's search facilities from a computer on campus, via the Internet, or telnet. The main catalog provided the following choices for search. I used the keywords search the most, and author only if I found someone who has done much work in the field I was interested in.

- Author
- Title
- Keywords
- Author and Title
- Journal Titles (Including Newspapers)
- Library of Congress
- Subject Headings
- Call Numbers

Figure 1 is an example of a return from the library catalog. As you can see, it provides much information including additional ideas for keyword searches.

I also used specialized databases that can provide more titles and often include abstracts and other valuable information. Because they do not necessarily tell if your library possesses these articles, it is necessary to go back through your library catalog and look them up. Some examples of the databases are:

- Silver Platter. Specialized indexes including: ABI Inform, Dissertation Abstracts, GPO Monthly Catalog, Medicine, NTIS, Sociofile, etc.
- EBSCO. General Indexes - some with full text.
- First Search. Book index as well as conference papers and proceedings.

The screenshot shows a library catalog interface. At the top, there are navigation links: ASU MAIN, ASU WEST, ASU EAST, and LAW. Below this is a search bar with the text 'You searched: AUTHOR degani'. The search results show 'Record 2 of 2'. The record details include: Gov doc # NAS 1.26.177605, Author Degani, Asaf, Title On the typography of flight-deck documentation [microform] / Asaf Degani, Publisher Moffett Field, Calif. : National Aeronautics and Space Administration, Ames Research Center ; [Springfield, Va. : National Technical Information Service, distributor, 1992], Note Distributed to depository libraries in microfiche. Microfiche. [Washington, D.C. : National Aeronautics and Space Administration, 1993] 1 microfiche. Below the record details is a table with columns LOCATION, CALL #, and STATUS. The table shows HAYDEN, NAS 1.26.177605, and AVAILABLE. Below the table are links for Description, Subject, Other author, Added title, Series, and ISBN/ISSN. At the bottom, there are links for Patron information, Suggestions, Acquire books, and Interlibrary loan.

LOCATION	CALL #	STATUS
HAYDEN	NAS 1.26.177605	AVAILABLE

Figure 0 Library search engine

It has become common today to rely heavily upon the Internet. While the Internet can be an outstanding source of information, the question must be, is this piece of information scholarly or is it "some guy's opinion". Aviation is a field with many personalities and opinions, everybody's got one. In researching the papers for this work, the Internet was used, but documents were used only from sites that could be deemed scholarly or scientific. The sites from which I found the best documents are:

- Air Force Research Laboratory - Armstrong Research Site <http://www.brooks.af.mil/AFRL/master.html>
- Aviation Week Online, CRM/HF. <http://www.awgnet.com/safety/crmhfres.htm>. Mostly good links to other sites.
- FAA, Other Links - Human Factors Activities <http://www.hf.faa.gov/LINKS/L-other.HTM>
- Flight Safety Foundation Publications http://www.flightsafety.org/pub_home.html
- NASA, Aviation Operations Branch: Research Programs <http://olias.arc.nasa.gov/groups.html>. From here you can find current publications and other links. This is one of the most useful sites.
- University of Texas Aerospace Crew Research. <http://www.psy.utexas.edu/psy/helmreich/nasaut.htm>. There are many good documents here, some are by students so be careful.
- US Navy, Office of Training. <http://www.ott.navy.mil/>
- US Navy, Science and Technology Office <http://www.ntsc.navy.mil/bf/scitech/scitech.htm>

This is not to say that other sources on the Internet were not used, but that these were the most fruitful.

Search engines are important tools when working on the web. Using commercial search engines such as Excite (<http://www.excite.com/>) and Yahoo (<http://www.yahoo.com/>) can locate good information, but they can be time consuming and lead to non-scholarly information. To speed the search process I used a multi-engine search tool called Web Ferret Pro ©. Web Ferret was useful in locating documents, which I wished to have, but did not know where to find, as well as allowing keyword searches over broad topics. By using this tool, documents and web sites were located, which I might not have found otherwise. Fewer truly academic sites and documents were located by keyword searches on commercial search engines, the preferred search engines were already attached to scientific sites such as NASA and the Navy, as seen below:

- NASA Technical Report Server (NTRS) <http://techreports.larc.nasa.gov/cgi-bin/ntrs>
- NASA Scientific and Technical Information Server <http://www.sti.nasa.gov/>
- Periscope - An international, armed forces and military database, for unclassified defense and government information, equipment, specifications, and news. <http://www.periscope.usni.com/>
- FedWorld File Transfer Protocol Search and Retrieve Service <http://www.fedworld.gov/ftp.htm>
- Federal Aviation Administration, Human Factors Program Management System Search Database <http://www.hf.faa.gov/database/templates/SearchForm.cfm>

Scope of Search

As previously mentioned, the amount of information available on a technical subject may be very broad. This may necessitate limiting your search to specific concepts or facets of the whole. In the case of this research project, I did not choose to limit the scope, in fact I enlarged it whenever a new variant became evident. However, most researchers will find themselves narrowing the scope of their search in order not to be overly broad in the final product. Using CRM as an example, Using Air Force Instruction 11-290, which describes the typical facets of a CRM program is the standard that most CRM programs follow. There are six primary areas of interest and a multitude of sub-areas beyond those. An overview may contain an amount of information on the topic of CRM as a whole, but a detailed study would usually cover one or two of the sub-areas.

Ideas change and information becomes stale. A prudent researcher will need to be abreast of the current "Thinking" and have some basic idea of what is dated material. CRM, while a relatively young area of interest, is an ever changing field of study and has numerous papers written about it's evolution. The above Air Force Instruction is current as of July 1998, it would be difficult to know this without having an insiders knowledge. All the

to measure certain data has improved, providing for the possibility for changes in the data. This is an area where the researcher will need to have some knowledge of the information available.

The Documents

Using Air Force Instruction 11-290 to organize the documents plus human factors and miscellaneous sections, the following is a short list of the most recent or at least the most important documents for each area.

- 1) 6.1.1. Situational Awareness. (Situational Awareness) Includes knowledge and skill objectives for preventing the loss of situational awareness, skills for recognizing the loss of situational awareness, and techniques for recovering from the loss of situational awareness.
 - a) Gibson, J., Orasanu, J., Villeda, E., Nygren, T. E., Loss of Situation Awareness: Causes and Consequences, 1998. <http://olias.arc.nasa.gov/projects/distr-team-decisions/publications/Gibson.pdf>
Abstract: Situation awareness (SA) has been recognized as an essential aspect of safe flight. The present analysis was conducted to identify some of the features that accompany loss of SA in actual flight situations, when the loss is recognized by flight crews. A set of 159 ASRS reports involving loss of SA was analyzed to answer several specific questions: In what phase of flight was SA lost? What factors may have caused or contributed to loss of SA? What were the consequences of the loss of SA? Who lost SA? Who performed corrective actions? The results of this study bring to light a better understanding of the multiple factors that affect SA in the flight environment, when this loss of SA occurs most often, and what consequences of the loss are likely to be.
 - b) Endsley, M. R., Situation Awareness, Automation & Free Flight, 1997. Air Traffic Management R&D Seminar, Sacalay France. <http://atm-seminar-97.eurocontrol.fr/endsley.htm>
Abstract: Free flight represents a major change in the way that aircraft may be handled in the U.S. National Airspace System. It has the potential of significantly increasing airspace utilization and thus improving aircraft throughput. The degree, to which these objectives can be met, however, without compromising aircraft safety, depends on establishing appropriate changes in the air traffic control system. These changes may include significant incorporation of new air traffic management and cockpit technologies. Studies that examine the affect of automation on situation awareness will be reviewed. These studies indicate that while certain forms of automation may be advantageous, other types of automation may induce losses in situation awareness that can lead out out-of-the-loop performance errors. A recent study was also conducted that examined changes in the locus of control between ATC and the cockpit using existing technology. This study provides an objective evaluation of the effects of free flight on controllers' ability to maintain an accurate and complete picture of the traffic situation in order to provide needed monitoring and separation functions. The study revealed that aspects of free flight can significantly hinder the situation awareness and performance of controllers. The results of this study provide information for better defining how free flight

should be implemented and for determining needed design and procedural modifications to support the concept.

- 2) 6.1.2. Crew Coordination/Flight Integrity. (Leadership and Assertiveness) Knowledge and skill objectives covering the impact on aircrew performance of command authority, leadership, responsibility, assertiveness, conflict resolution, hazardous attitudes, behavioral styles, legitimate avenues of dissent, and team-building.

- a) Pettitt, M. A, Dunlap, J. H., Cockpit Conflict: An Escalation Dynamic. 1998. Aviation Leadership & Followership. http://www.wmich.edu/aviation/pdf/papers/theory_perspectives_training.pdf

Abstract: As with any other work environment, conflicts in the cockpit will occur. Workplace conflicts can be task-oriented, people-oriented, or goal-oriented. While the first two are more common, goal-oriented conflicts happen, for example, when the goals of the company (e.g., ontime departures) conflict with the goals of the crew (e.g., adding extra fuel because of weather at the destination). Conflicts can result from a lack of communication or information. Conflicts also arise because of differences in perceptions, preferred outcomes, or values. These are typical causes of CRM-related cockpit conflict and the ones we will discuss in this article.

- b) Smith, P. J., McCoy, E., Orasanu, J., et al. Cooperative Problem-Solving in the Interactions of Airline Dispatchers and Flight Crews., 1996. <http://www.hf.faa.gov/products/coopprob/2dispcrw.pdf>

Abstract: A previous study of flight preplanning decisions by air transport pilots highlighted the increasing interdependence of flight crews and ground-based operations control staff to ensure safe and efficient flight operations. Although the Pilot in Command has ultimate authority in making decisions, it is clear that Dispatchers and other ground-based personnel are an important resource to support planning and decision-making by the flight crew. To examine this interdependence further and to identify areas for improvement, a one-day focus group was held to discuss the interactions of airline Pilots and Dispatchers. Based on the discussions of participants from eight airlines, a variety of important issues were identified.

Three major classes of issues emerged, dealing with communication, training and workload.

Communication was seen as the biggest problem and included the following: Difficulty in initiating communication links, poorly formatted messages in which critical information becomes buried, inadequate procedures during the hand-over of authority, and use of confusing and opaque abbreviations and acronyms. Discussions also identified weaknesses in the training of Dispatchers, Maintenance personnel, Pilots and Air Traffic Controllers, especially training that relates to interactions among these groups and how they can support each other. A final issue was concern with the effects of Dispatcher workload, particularly during bad weather when holdings patterns or diversions become necessary. Potential areas for improving the current system and for future research were identified.

- 3) 6.1.3. Communication. (Communication) Includes knowledge of common errors, cultural influences, and barriers (rank, age, experience and position). Skills will encompass listening, feedback, precision and efficiency

of communication with all members and agencies (i.e., Crew-members, Wingmen, Weather, ATC, Intelligence, etc.).

- a) Orasanu, J., Davison, J., Fischer, U., What Did He Say? Culture and Language Barriers To Efficient Communication in Global Aviation, 1997. Presented at the Ninth Biennial International Symposium on Aviation Psychology. <http://olias.arc.nasa.gov/projects/distr-team-decisions/publications/Orasanu.pdf>
Abstract: The most recent generation of Crew Resource Management (CRM) training has been expanded to include interactions between flight crews and flight attendants, dispatchers, and maintenance personnel. But from a safety perspective, one of the most critical interactions is between flight crews and air traffic controllers (ATC). Several accidents in the US have involved inadequate air/ground communications. Effective communication in high traffic density areas can be a challenge, especially when both pilots and controllers are busy. When the culture and native language of the participants differ, the risk of communication errors escalates even further, a situation that is becoming more common with the increasing volume of international traffic. Consider several accidents in which language problems have contributed directly to the accidents.

- b) Matchette, R, Say What! Non-standard phraseology incidents., ASRS Directline. Issue Number 7 : September 1995. http://olias.arc.nasa.gov/asrs/dl7_say.htm
Abstract: Aviation has enjoyed numerous advances in aerodynamics, power plant efficiency and reliability, flight deck automation, and navigation systems. However, ATC/aircraft communications have changed little over the years, and still exhibit the age-old limitations of natural and human-made interference that can distort messages, difficulties with language barriers, and the problems of pronunciation and phraseology. At the same time, the volume of ground-to-air (ATC/aircraft) communication has increased dramatically because of the remarkable increase in air traffic. Satellite links and discrete communication technology promise communications solutions for the future--until then, aviation is forced to deal with the communications status quo. One of the greatest problems inherent in voice communications today is the use of non-standard phraseology. The ASRS database was searched for records which made reference to phraseology in their narratives, and 260 reports were reviewed. Many reported incidents resulted in little more than momentary confusion or annoyance for pilots and controllers. However, nearly half the reports involved near mid-air collisions, loss of standard ATC separation, runway transgressions, or other conflicts with potentially serious safety consequences.

4) 6.1.4. Risk Management/Decision Making. (Decision Making) Includes risk assessment, the risk management process, tools, breakdowns in judgment and discipline, problem-solving, evaluation of hazards, and control measures.

- a) Driskill, W.E., Weissmuller, J.J., Quebe, J., Hand, D.K., and Hunter, D.R., Evaluating the Decision-Making Skills of General Aviation Pilots, 1998.
<http://www.cami.jccbi.gov/publications/Abstracts/1998/FULLTXT/9807.pdf>

Abstract: An instrument consisting of 51 items was developed to assess pilot decision-making skill. Each item consisted of a stem, a short description of an aviation scenario requiring a decision on the part of the pilot. Four alternatives were provided, and subjects were instructed to rank order the alternatives from best to worst solution to the scenario presented. Rank-ordered judgments of a sample of 246 general aviation (GA) pilots (with an average of about 500 hours of total flying experience) were compared with the recommended solutions provided by an expert panel. Results indicated that, overall, GA pilots and an expert panel of pilots agreed in their judgments of the appropriate course of action in situations critical to flight safety. However, the degree of agreement of individual general aviation pilots with the recommended solutions varied widely. An index of agreement (Safety Deviation Index) was calculated that expressed the degree of agreement of individual GA pilots with the recommended solutions. Initial evaluation of this index indicates that it demonstrates adequate psychometric properties and that, as other research would suggest, it has little relationship with common demographic or flight experience measures.

Comments: The authors indicate that the high number of GA accidents and incidents are due to faulty decision making on the part of "non-expert" pilots and explore the difference in cognitive processes.

- b) Jensen, R. S., Guilkey, J. E., and Hunter, D. R., An Evaluation of Pilot Acceptance of the Personal Minimums Training Program for Risk Management. 1998.

<http://www.cami.jccbi.gov/publications/Abstracts/1998/FULLTXT/9806.pdf>

Abstract: A new general aviation training program entitled, "Personal Minimums for Aviator Risk Management in Pre-Take-off Decisions" was field tested in five diverse geographic locations around the USA (Columbus, OH; Long Beach, CA; Anchorage, AL; Baltimore, MD/Washington, DC; and Chicago, IL) to determine its acceptability to pilot audiences and to obtain feedback for further development of the intervention. In each case, following the presentation, participants were asked to evaluate the course and its acceptability to the general aviation community. Analysis of these evaluations revealed that respondents viewed the training program as helpful and intended to use personal minimums as part of their pre-flight decision making in the future.

Comments: This is one of the few academic documents discussing CRM issues at the GA level. Since a large percentage of aviation mishaps occur with GA pilots, perhaps this is an area which needs to be explored more deeply.

- 5) 6.1.5. Task Management (Adaptability/Flexibility). Includes establishing priorities, overload, underload, complacency, management of automation, available resources, checklist discipline, and standard operating procedures.

- a) Van Selst, M., Johnston, J. C., 1997. Dual-task Interference when a Response is Not Required. Dual-task interference that cannot be attributed to output Interference. Paper presented at the 19th annual conference of the Cognitive Science Society, Stanford. <http://olias.arc.nasa.gov/publications/vanselst/go-nogo/go-nogo.html>

Abstract: When subjects are required to respond to two stimuli presented in rapid succession, responses to the second stimulus are delayed. Such dual-task interference has been attributed to a fundamental processing bottleneck preventing simultaneous processing on both tasks. Two experiments show dual-task interference even when the first task does not require a response. The observed interference is caused by a bottleneck in central cognitive processing, rather than in response initiation or execution.

- b) Schutte P. C, Trujillo A. C., 1996. Flight Crew Task Management In Non-Normal Situations. Abstract: Task management (TM) is always performed on the flight deck, although not always explicitly, consistently, or rigorously. Nowhere is TM as important as it is in dealing with non-normal situations. The objective of this study was to analyze pilot TM behavior for non-normal situations. Specifically, the study observed pilots' performance in a full workload environment in order to discern their TM strategies. This study identified four different TM prioritization and allocation strategies: 'Aviate-Navigate-Communicate-Manage Systems;' 'Perceived Severity;' 'Procedure Based;' and 'Event/Interrupt Driven.' Subjects used these strategies to manage their personal workload and to schedule monitoring and assessment of the situation. The 'Perceived Severity' strategy for personal workload management combined with the 'Aviate-Navigate-Communicate- Manage Systems' strategy for monitoring and assessing appeared to be the most effective (fewest errors and fastest response times) in responding to the novel system failure used in this study.

- 6) 6.1.6. Mission Planning/Debrief (Mission Analysis). Includes pre-mission analysis and planning, briefing, ongoing mission evaluation, and post mission debrief. Also, specific tools and techniques to be used in operational and training missions.

- a) Dismukes-R.K.: Jobe-K.K.: McDonnell-L.K., LOFT Debriefings: An Analysis of Instructor Techniques and Crew Participation, 1997. Gov. Doc N19970014649XSP
http://olias.arc.nasa.gov/publications/dismukes/final_loft_tm/Final_LOFT_TM.html
Abstract: This study analyzes techniques instructors use to facilitate crew analysis and evaluation of their LOFT performance. A rating instrument called the Debriefing Assessment Battery (DAB) was developed which enables raters to reliably assess instructor facilitation techniques and characterize crew participation. Thirty-six debriefing sessions conducted at five U.S. airlines were analyzed to determine the nature of instructor facilitation and crew participation. Ratings obtained using the DAB corresponded closely with descriptive measures of instructor and crew performance. The data provide empirical evidence that facilitation can be an effective tool for increasing the depth of crew participation and self-analysis of CRM performance. Instructor facilitation skill varied dramatically, suggesting a need for more concrete hands-on training in facilitation techniques. Crews were responsive but fell short of actively leading their own debriefings. Ways to improve debriefing effectiveness are suggested.
- b) McDonnell L. K. , Jobe K. K. , Dismukes R. K., 1997. Facilitating LOS Debriefings: A Training Manual. NASA Technical Memorandum 112192 DOT/FAA/AR-97/6. <http://olias.arc.nasa.gov/projects/distr-team->

[decisions/publications/Mcdonn.pdf](#)

Abstract: This manual is based on our study of LOFT debriefings at several U.S. airlines. The suggestions in this manual are derived from the data from that study, our subjective impressions, the experiences the LOFT instructors shared, and general literature on facilitation. Data and references to relevant literature from the study are available in the published report: LOFT Debriefings: An Analysis of Instructor Techniques and Crew Participation, by R.K. Dismukes, K.K. Jobe, and L.K. McDonnell (NASA Technical Memorandum 110442; March 1997). This material is presented as suggestions rather than rules because facilitation is very much a personal skill and each instructor must develop an approach with which he or she is comfortable. These suggestions provide a tool kit of techniques instructors may draw upon to develop their own style.

7) Human Factors:

- a) Dunlap J. H., Pettitt, M. A., 1998. Training vs. Selection: The Right (CRM) Stuff. [Online]

<http://www.wmich.edu/aviation/pdf/papers/crm.pdf>

Abstract: Over the past decade considerable resources have been expended to identify the causes of airline accidents and develop training programs that provide pilots with the skills necessary to transport passengers safely. Given the high percentage of human factors underlying of air transport accidents, many of these resources have been devoted to crew resource management (CRM) and related training. This effort has been jointly supported by industry, the FAA's Office of Human Factors (AAR- 100), NASA, and the academic research community. Most recently, airlines have embraced the FAA's Advanced Qualification Program (AQP). The AQP training curricula are proficiency-based -- rather than requirement-based -- and AQP requires that CRM skills be included in the training regimen. These efforts have made a significant contribution to airline safety. However, it is worth asking: Have these measures individually been enough? Could a more systemic approach to CRM result in a complementary, even synergistic, effect."

Comments: A human factors specialist will tell you to fit the job to the person, this paper suggests that through thorough selection an airline will find people who are naturally adept at CRM and flying., rather than train people how to fit into the cockpit environment.

- b) , Rosekind, M. R., & Neri, D. F., Dinges, D. F., 1997. From Laboratory to Flightdeck: Promoting Operational Alertness. In Fatigue and Duty Limitations – An International Review. The Royal Aeronautical Society (pp. 7.1 - 7.14). London: The Royal Aeronautical Society.

<http://olias.arc.nasa.gov/zteam/fcp/pubs/raes.html>

Abstract: Human operators remain central to safe aviation operations. Fatigue, sleep loss, and circadian disruption created by flight operations can degrade performance, alertness and safety. An extensive scientific literature exists that provides important physiological information about the human operator that can be used to guide operations and policy. For example, there are human physiological requirements for sleep, predictable effects on performance and alertness with sleep loss, and patterns for recovery. The

circadian clock is a powerful modulator of human performance and alertness and it can be disrupted in aviation through night flying, time zone changes, and day/night duty shifts. Scientific examination of these physiological considerations has established a direct relationship to errors, accidents, and safety. This scientific information can be incorporated into flight/duty/rest regulatory considerations. Managing fatigue in the complex and diverse aviation environment requires an integrated and multicomponent approach. These factors preclude a simple solution and managing fatigue will benefit from addressing education, hours of service, strategies, technology, design, and research. Concept development should be initiated to move beyond current flight/duty/rest regulatory schemes and toward operational models that provide flexibility and maintain the safety margin

- c) Rosekind, M. R., Neri, D. F., Miller, D. L., Gregory, K. B., Webbon, L. L., Oyung, R. L., 1997. Crew Fatigue Research Focusing on Development and Use of Effective Countermeasures. ICAO Journal, 52. 20 - 22. <http://olias.arc.nasa.gov/zteam/fcp/pubs/ICAOjournal.html>

Abstract: AROUND-THE-CLOCK aviation operations pose unique challenges. Physiological requirements related to sleep, the internal circadian clock, and human fatigue are critical factors that are known to affect safety, performance, and productivity. Understanding the flight crews' capabilities and limitations is important to address these issues as global demand for aviation continues to increase.

The U.S. National Aeronautics and Space Administration (NASA) Ames Research Center initiated a program in 1980 to examine the role of fatigue in flight operations. The program was designed to determine the extent of fatigue, sleep loss, and circadian disruption in flight operations and how fatigue affected crew performance. It was also designed to develop strategies to maximize performance and alertness during flight operations.

8) Miscellaneous:

- a) Helmreich R. L., Merritt A. C. & Wilhelm J. A. 1998. The Evolution of Crew Resource Management Training in Commercial Aviation. http://www.psy.utexas.edu/psy/helmreich/Evolution_IJAP_for_Dist.htm. Abstract: Changes in the nature of CRM training in commercial aviation are described, including its shift from Cockpit to Crew Resource Management. Validation of the impact of CRM is discussed. Limitations of CRM, including lack of cross-cultural generality are considered. An overarching framework that stresses error management to increase acceptance of CRM concepts is presented.... The roots of Crew Resource Management training in the United States are usually traced back to a workshop, Resource Management on the Flightdeck sponsored by the National Aeronautics and Space Administration in 1979 Cooper, White, & Lauber, (1980). This conference was the outgrowth of NASA research into the causes of air transport accidents. The research presented at this meeting identified the human error aspects of the majority of air crashes as failures of interpersonal communications, decision making, and leadership. At this meeting, the label Cockpit Resource Management (CRM) was applied to the process of training crews to reduce "pilot error" by making better use of the human resources on the flightdeck. Many of the air carriers represented

at this meeting left it committed to developing new training programs to enhance the interpersonal aspects of flight operations.... CRM training in the military has followed its own path of growth and evolution and will not be addressed here (see Prince & Salas, 1993, for a discussion of military CRM programs). We use the term 'evolution' in describing the changes in CRM over the last two decades.... Similarly, the very different content and foci of programs called CRM justifies defining them in terms of generations ... Our focus is on the most recent approaches to CRM training.

Comments: This is an overview of changes in CRM philosophies. The authors express changes as generations of CRM, which can be labeled by era. The authors also believe that the 5th and most recent version leaves the "Charm School" CRM programs behind, focusing instead on error trapping techniques.

- b) Mangold, S. L. & Neumeister, D. M., 1997. AQPing CRM Proceedings of the Ninth International Symposium on Aviation Psychology. http://www.wmich.edu/aviation/pdf/papers/aqping_crm.pdf
Abstract: Crew Resource Management (CRM) training, defined broadly to include human factors issues in general, has evolved through a number of stages since it was first introduced as an industry concern in the 1970's. The original emphasis on the application of business-management-based tools for supporting teamwork activities gave way to the use of behavioral markers that reflect high-level, observable crew behaviors for specific CRM categories, such as decision making, situation awareness, and communication. The event set concept reinforced the emphasis on observable crew behaviors and provided a flight-situation-oriented framework for training and evaluating CRM skills. This shift to a greater interest in observable behaviors contributed to the belief that truly effective CRM training could only occur if it was thoroughly integrated with technical training.

Comments: AQP is the FAA's Advanced Qualification Program, a new maintaining commercial pilot proficiency.

Summary

To research a topic in depth today requires computer skills relating to search engines, both on the Internet and in the library. The researcher must be able to narrow the search or an overwhelming amount of data will be received. The researcher must also be able to discern between scholarly information and information whose pedigree is somewhat questionable.

To keep abreast of a topic such as CRM requires constant and diligent effort. New information is created daily and it is up to the researcher to ferret this information out. The documents listed above are a sampling of some of the current documents available, the documents following are a larger listing which I will maintain on a web site here at ASU. The address is <http://eastair.east.asu.edu/crm>.

References

- Dismukes-R.K.: Jobe-K.K.: McDonnell-L.K., LOFT Debriefings: An Analysis of Instructor Techniques and Crew Participation, 1997. Gov. Doc N19970014649XSP
http://olias.arc.nasa.gov/publications/dismukes/final_loft_tm/Final_LOFT_TM.html
- Driskill, W.E., Weissmuller, J.J., Quebe, J., Hand, D.K., and Hunter, D.R., Evaluating the Decision-Making Skills of General Aviation Pilots, 1998. <http://www.cami.jccbi.gov/publications/Abstracts/1998/FULLTXT/9807.pdf>
- Dunlap J. H., Pettitt, M. A., 1998. Training vs. Selection: The Right (CRM) Stuff. [Online]
<http://www.wmich.edu/aviation/pdf/papers/crm.pdf>
- Endsley, M. R., Situation Awareness, Automation & Free Flight, 1997. Air Traffic Management R&D Seminar, Sacalay France. <http://atm-seminar-97.eurocontrol.fr/endsley.htm>
- Gibson, J., Orasanu, J., Villeda, E., Nygren, T. E., Loss of Situation Awareness: Causes and Consequences, 1998.
<http://olias.arc.nasa.gov/projects/distr-team-decisions/publications/Gibson.pdf>
- Helmreich R. L., Merritt A. C. & Wilhelm J. A. 1998. The Evolution of Crew Resource Management Training in Commercial Aviation. http://www.psy.utexas.edu/psy/helmreich/Evolution_IJAP_for_Dist.htm.
- Jensen, R. S., Guilkey, J. E., and Hunter, D. R., An Evaluation of Pilot Acceptance of the Personal Minimums Training Program for Risk Management. 1998.
<http://www.cami.jccbi.gov/publications/Abstracts/1998/FULLTXT/9806.pdf>
- Mangold, S. L. & Neumeister, D. M., 1997. AQPing CRM Proceedings of the Ninth International Symposium on Aviation Psychology. http://www.wmich.edu/aviation/pdf/papers/aqping_crm.pdf
- Matchette, R., Say What! Non-standard phraseology incidents., ASRS Directline. Issue Number 7 : September 1995.
http://olias.arc.nasa.gov/asrs/dl7_say.htm
- McDonnell L. K. , Jobe K. K. , Dismukes R. K., 1997. Facilitating LOS Debriefings: A Training Manual. NASA Technical Memorandum 112192 DOT/FAA/AR-97/6. <http://olias.arc.nasa.gov/projects/distr-team-decisions/publications/Mcdonn.pdf>
- Orasanu, J., Davison, J., Fischer, U., What Did He Say? Culture and Language Barriers To Efficient Communication in Global Aviation, 1997. Presented at the Ninth Biennial International Symposium on Aviation Psychology.
<http://olias.arc.nasa.gov/projects/distr-team-decisions/publications/Orasanu.pdf>
- Pettitt, M. A., Dunlap, J. H., Cockpit Conflict: An Escalation Dynamic. 1998. Aviation Leadership & Followership.
http://www.wmich.edu/aviation/pdf/papers/theory_perspectives_training.pdf
- Rosekind, M. R., Neri, D. F., Miller, D. L., Gregory, K. B., Webbon, L. L., Oyung, R. L., 1997. Crew Fatigue Research Focusing on Development and Use of Effective Countermeasures. ICAO Journal, 52. 20 - 22.
http://olias.arc.nasa.gov/zteam/fcp/pubs/ICAO_journal.html
- Rosekind, M. R., & Neri, D. F., Dinges, D. F., 1997. From Laboratory to Flightdeck: Promoting Operational Alertness. In *Fatigue and Duty Limitations – An International Review*. The Royal Aeronautical Society (pp. 7.1 - 7.14). London: The Royal Aeronautical Society. <http://olias.arc.nasa.gov/zteam/fcp/pubs/raes.html>
- Schutte P. C., Trujillo A. C., 1996. Flight Crew Task Management In Non-Normal Situations.
- Smith, P. J., McCoy, E., Orasanu, J., et al. Cooperative Problem-Solving in the Interactions of Airline Dispatchers and Flight Crews., 1996. <http://www.hf.faa.gov/products/coopprob/2dispcrw.pdf>
- Van Selst, M., Johnston, J. C., 1997. Dual-task Interference when a Response is Not Required. Dual-task interference that cannot be attributed to output Interference. Paper presented at the 19th annual conference of the Cognitive Science Society, Stanford. <http://olias.arc.nasa.gov/publications/vanselst/go-nogo/go-nogo.html>

The Documents

Error! Not a valid filename.

PERMEABILITY CHARACTERISTICS OF AN ENDOTHELIAL
CELL MODEL OF THE BLOOD-BRAIN BARRIER

Vanessa D. Le
Graduate Student
Department of Pharmaceutics

University of Texas at Austin
College of Pharmacy
Austin, Texas 78712-1074

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1998

PERMEABILITY CHARACTERISTICS OF AN ENDOTHELIAL CELL MODEL OF THE BLOOD-BRAIN BARRIER

Vanessa D. Le
Graduate Student
Department of Pharmaceutics
College of Pharmacy
University of Texas at Austin

Abstract

The purpose of this study was to characterize the permeability characteristics of an in vitro blood-brain barrier (BBB) cell system and relate this information to available in vivo data.. The model was developed by culturing cerebromicrovascular endothelial cells on gelatin-coated polyethylene terphthalate membrane and was thoroughly characterized with regards to morphological appearance, transendothelial electrical resistance (TEER) and the permeability of fluorescein isothiocyanate-dextran (FITC-dextran). The model depicted here bear resemblances to many of the vitro models reported in literature with regards a cobblestone morphology, transendothelial resistance and permeability coefficient for FITC-dextran. However, there appears to be differences in the transendothelial permeability characteristics of this model in comparison to the in vivo BBB. The established cell model was 10-100 times more permeable than that of a continuous BBB endothelium. A possible reason for this increased permeability appeared to be small intermittent gaps between endothelial cells of this model that allowed for leakage of large molecules.

PERMEABILITY CHARACTERISTICS OF AN ENDOTHELIAL CELL MODEL OF THE BLOOD-BRAIN BARRIER

Vanessa D. Le

Introduction

The BBB consists of a continuous, single layer of closely apposed endothelial cells lining the microvessels of the brain that provides the neurons of the central nervous system (CNS) with a constant environment thereby maintaining homeostasis. The structural basis for this unique barrier appears to be related to features of the endothelium of the brain capillaries, specifically, endothelial cells exhibit tight intercellular junctions, few pinocytotic vesicles, no fenestra, a complex glycocalyx, and a significant metabolic capacity which is consistent with the 5-fold greater mitochondrial content of brain endothelial cells compared to peripheral vascular endothelia (Oldendorf et al., 1977). Together these morphologic and biochemical characteristics of endothelial cells in the brain microvasculature restrict the bidirectional passage of polar, ionized, hydrophilic and even certain lipophilic compounds from the blood and the brain. Exceptions include relatively lipophilic substances and the various CNS-required nutrients, e.g., glucose, amino acids, choline) for which specific, saturable BBB transport systems exist (Betz and Goldstein, 1978; Betz et al., 1980).

A good *in vitro* model of the BBB would have wide application in experimental, pharmaceutical and clinical studies. The focus of developing an *in vitro* model for addressing specifically transendothelial drug delivery problems has, therefore, been on systems derived from blood-brain barrier endothelium either in culture systems or as isolated suspensions of endothelial cell provides a cell model that allows for the assessment of transport profiles and the study of factors that influence drug transport across the endothelial monolayer. Culture of brain

capillary cells affords several advantages, in particular the preservation of cellular polarity and metabolism. Primary cultures of brain microvessel endothelial cells isolated either enzymatically or mechanically retain many characteristics of the parent tissue including the morphology, specific BBB enzyme markers, few pinocytotic vesicles, absence of fenestra, and tight intercellular junctions (Bowman et al., 1981; Audus et al., 1990).

Isolation and characterization of primary cultures of brain endothelial cells are, however, time consuming; the resulting cultures differ in quality, contaminating cells, and their culture requirements are more fastidious than systemic and/or macrovascular endothelial cells. While passaging has the advantage that it generates large stocks of cells for experimental study, it introduces additional complications due to dedifferentiation and selection of the most rapidly proliferating clones. There is considerable variability between properties reported for bovine cultures from different laboratories. Furthermore, it has also been reported that passaged cultures can only be maintained for limited passages before onset of senescence, which is associated with reduced proliferation rate and loss of phenotypic markers (Durieu-Trautmann et al., 1991).

To facilitate the functional *in vitro* study of brain endothelial cells, it seems necessary to produce stable immortalized cell lines capable of maintaining a fully differentiated phenotype in culture. Therefore, a readily available, homogeneous, and highly proliferative microvascular capillary endothelial cell population of human origin bearing stable phenotypic properties of the BBB endothelium would serve as an invaluable tool for conducting drug transport studies across the BBB.

Rat (Greenwood et al., 1996; Trautmann et al., 1993) and bovine (Durieu-Trautmann et al., 1991; Stins et al., 1997) immortalized microvascular capillary endothelial cell lines that

maintain some of the structural and pharmacologic characteristics of the BBB endothelium have been reported. The first successful immortalization of human cerebromicrovascular endothelial cells (HCEC) by stable transfection of SV40 large T into primary HCEC (SV-HCEC) was reported by Muruganandam et al. (1997).

Extensive morphological and functional characterization of the established SV-HCEC line provides evidence that these immortalized cells retain most of the differentiated properties/functions of primary HCEC and can therefore be used to study functional and transport properties of the human BBB *in vitro* (Table 2).

Table 2. Genotypic and phenotypic characterization of SV-HCEC

Endothelial markers:	Factor VIII-related antigen Uptake of Ac-LDL and LDL Lectin binding: <i>Ricinus Communis Agglutinin I</i> <i>Ulex Europeus</i>
BBB markers:	Transferrin receptor and receptor-mediated endocytosis of transferrin P-glycoprotein (mdr-1) HT-7 ICAM-1 VCAM-1 PECAM Glucose transporter 1 γ -glutamyl transpeptidase

Likewise, the use of SV-HCEC in *in vitro* models of the BBB designed for fast screening of drug access to the brain will furnish an important advantage to existing models that use

immortalized endothelial cells from sources other than brain tissue or primary brain endothelial cells of other species such as bovine and rat. This *in vitro* model also affords the use of an immortalized human fetal astrocyte cell line, which overcomes the problems associated with cross-species interaction.

Materials and Methods

Materials:

Immortalized human cerebromicrovascular endothelial cells and human fetal astrocyte cells were generously provided by Dr. Arumugam Muruganandam and Dr. Danica Stanimirovic from the National Research Council of Canada. All culture media and HEPES were obtained from Gibco BRL (Gaithersburg, MD). Fetal calf serum (FCS), penicillin/streptomycin solution, endothelial cell growth supplement (ECGS), heparin, insulin, poly-L-lysine, L-glutamine, gelatin, fluorescein-conjugated dextran (MW 70,000), TRITC-labeled Concanavalin A, TRITC-labeled Agglutinin RCA₁₂₀ and primary polyclonal rabbit anti-human factor VIII-related antigen antibody were obtained from Sigma Chemical Co. (St. Louis, MO). Holo-transferrin and selenium were purchased from Collaborative Biomedical Pds. Calcein-AM, BODIPY FL-conjugated AcLDL and DiOC₅ were purchased from Molecular Probe (OR). Texas red-conjugated human transferrin was obtained from Accurate Chemical and Scientific Corp. (Westbury, NY). Transport studies were conducted on Falcon 6-well plates with Falcon culture inserts (pore size 1 μm ; SA 4.2 cm^2 ; track-etched polyethylene terephthalate membrane).

Methods:

1. Cell culture conditions

Human cerebromicrovascular endothelial cells (HCECs) were grown in growth media comprising of 65% medium M199 supplemented with Earle's salts, 25 mM HEPES, 4.35 g/L sodium bicarbonate, and 3 mM L-glutamine, 10% FBS, 5 µg/ml of insulin, 5 µg/ml transferrin, 5 ng/ml selenium, and 10 µg/ml ECGS. Cells were subcultured at a 1:3 and 1:10 split ratio every 7 and 3 days, respectively, by first washing with Ca^{2+} and Mg^{2+} free-Hanks' balanced salt solution (CMF-HBSS), exposed to trypsin-EDTA solution for endothelial cells for 4 minutes and then centrifuged at 1000 rpm for 10 minutes. Cells were then counted using a trypan blue stain to determine percent cell viability (>95% viability by trypan blue exclusion). Cells were then seeded into 0.5% gelatin-coated T-flask or Falcon culture inserts at a density of 75,000 cells/cm² or 3.15×10^5 cells/insert and allowed to attach and grow to a complete monolayer in the presence of 95% relative humidity, 5% CO₂, at 37°C. Confluency was assessed based on sustained TEER measurements and phase-contrast microscopy.

To coat the tissue culture plasticwares with gelatin, plasticwares were flooded with a 0.5% gelatin solution and placed in 4°C overnight. The following day, the gelatin solution was aspirated, the surface washed twice with CMF-HBSS and filled with culture medium prior to plating the cells.

2. Characterization of SV-HCEC cell monolaye

The following characterization assays were performed after every 20th passages and were based on the methods of Muruganandam et al. (1997):

2a. To determine the expression of the endothelium-specific Factor VIII-related antigen, cells

were seeded at a density of 3.15×10^5 onto human fibronectin-coated glass coverslip and allowed to incubate for 4 days until confluency is maintained. The monolayer-coated coverslips were washed with PBS, fixed with a 3% solution of formaldehyde in PBS, washed with PBS, and permeabilized in ice-cold 0.5% Triton-X100 for 10 minutes followed by a washing step. Coverslips were then incubated in 10% FCS for 30 minutes to block for non-specific binding and incubated with rabbit anti-human factor VIII-related antigen serum (1:50 dilution in PBS) for 1 hour at room temperature. The coverslips are then washed with PBS and exposed to fluorescein-conjugated goat anti-rabbit IgG serum (1:50 dilution with PBS) for 1 hr at RT. Coverslips were then washed and inverted onto a solution of PBS containing propylgalate (5%) and glycerol (50%). Coverslips were then counterstained with cationic membrane probe DiOC₅ (10 nM) for 15 minutes and analyzed on a confocal microscope using a 510-525 nm and 610 nm band pass filter for DiOC₅ and Texas red, respectively.

2b. Receptor-mediated internalization of transferrin was assessed by determining the temperature-sensitive uptake of Texas red-conjugated transferrin. Cells, seeded at a density of 3.15×10^5 , were grown on human fibronectin-coated coverslips. Once confluency has been attained, cells were then incubated in transferrin-free medium M199 for 3 hr prior to the addition of 25 $\mu\text{g/ml}$ Texas-red-labeled transferrin. Internalization of transferrin was visualized by confocal microscopy after 4 hr of incubation at 37°C. The cells were then counterstained with the vital dye calcein-AM for 10 min and analyzed using confocal microscopy with a 515-540 nm band pass and a 610 nm long pass filter for calcein-AM and Texas red, respectively.

2c. Sigma Kit No. 545 was used to quantitate the activity of γ -GTP in SV-HCEC according to the manufacturer's instructions. Cells, grown on plastic culture dishes (without coating), were scraped with a rubber policeman in the presence of 1 ml PBS. The cell suspension was collected and the dish washed with 0.5 ml PBS. The wash was added to the previously collected cell suspension and the total 1.5 ml cell suspension was sonified with a Branson sonifier, equipped with a micro-tip, for 10 sec at 20 W peak power. Aliquots of the cell suspension were taken for measurement of γ -GTP activity and protein content. Protein quantitation was determined based on the Lowry Method. Gamma GTP content was determined by incubating the cell suspension aliquot with γ -GTP substrate (51 μ mol L- γ -glutamyl-p-nitroanilide and 1.1 mmol glycylglycine) in Trizma buffer for 20 min at 37°C. The reaction catalyzed by γ -GTP, L-glutamyl-p-nitroaniline + glycylglycine $\rightarrow$ p-nitroaniline + glutamylglycylglycine. This reaction was stopped by addition of 10% (v/v) glacial acetic acid. The liberated p-nitroaniline was diazotized by a modified Bratton-Marshall reaction and the resulting diazo compound solution was reacted with N-1-naphthylethylenediamine-HCl to form a pink azo dye. The absorbance of azo dye was read at 545 nm with a Perkin Elmers spectrophotometer and was used to calculate the activity of γ -GTP. One unit of γ -GTP is defined as that amount of activity that will liberate 1 nmol of p-nitroaniline per min at 25°C.

3. Integrity of the cell monolayer

The integrity of the cell monolayer was monitored a day after seeding and thereafter until stabilized TEER values were maintained which was indicative of a confluent monolayer.

3a. Resistance measurements

Falcon culture inserts, with or without (control) confluent monolayer, were placed in a chamber containing medium M199. Resistance was measured using an Endohm 26S assembly, containing current-passing and voltage-measuring disc electrodes (World Precision Instruments Inc., New Haven, CT), according to the manufacturer's instructions. The resistance values were normalized for surface area and reported in ohms times centimeters squared. To specifically determine the resistance associated with each cell monolayer in each culture insert, TEER values (ohms-cm²) were calculated as follows: resistance of (filter membrane + gelatin layer + cell monolayer) – resistance of (filter membrane + gelatin layer) x surface area (cm²) of the insert membrane. To both increase the accuracy and lessen the variability in TEER measurements, electrodes were incubated in medium for 30 minutes before use.

3b. Transport of non-permeable, hydrophilic marker

The flux of hydrophilic markers across BCE cell monolayers has been reported to occur mainly through intercellular regions, as BCE cells do not express receptors for dextran and display only minimal pinocytotic activity. Hence, the degree to which dextran is transported transendothelially provides a good index of the tightness of the endothelial barrier. Cells were seeded at a density of 3.15×10^5 cells/culture insert and allowed to incubate until confluency is achieved. Cell monolayers were then washed once with transport buffer to rid of FCS followed by the addition of 2.8 and 3 ml buffer to the donor and acceptor chambers, respectively. At t=0, 200 µL of fluorescein isothiocyanate-dextran (FITC-dextran) solution (MW 70,000; 1 mg/ml dissolved in transport buffer) was added to the donor chamber. At t=0,1,2 and 3 hrs, 40 and 200 µl were withdrawn from the donor and acceptor chambers, respectively. After sampling, the

same volume of pre-warmed buffer was added back to the respective chambers. Transendothelial transfer rate of FITC-dextran was determined 3 hours later using a Perkin-Elmer microplate fluorimeter. Samples were excited at wavelength of 490 and the emitted light measured at 530 nm with 3-nm slit widths. The amount of FITC-dextran transported was calculated as the percent of applied FITC-dextran that passed from the donor to the acceptor chamber at varying time intervals.

It is necessary to calculate the P_e (permeability coefficient) values to perform the *in vivo/in vitro* correlation since the surface area (SA) of the endothelial monolayer ($4.2 \text{ cm}^2/\text{well}$) differs from the SA of the BBB *in vivo* ($100 \text{ cm}^2/\text{g}$ of brain for male Sprague-Dawley rats) (Rapoport, 1976). Clearance volume (μl) will be calculated as follows: $\text{Clearance } (\mu\text{l}) = ([C]_A \times V_A) / [C]_D$, where $[C]_A$ = acceptor tracer concentration, V_A = volume of the acceptor chamber, and $[C]_D$ = the initial donor tracer concentration. The slopes of the clearance curves for membranes alone and membranes with cell monolayers, representing permeability X SA product (PS, $\mu\text{l}/\text{min}$), will be calculated using linear regression analysis. The PS values of endothelial monolayers (PS_e) will be calculated based on the following equation: $1/PS_e = 1/PS_{e+m} - 1/PS_m$, where PS_{e+m} = slope of the clearance curve for cell and membrane together, PS_m = slope of the clearance curve of membranes without cells, and P_e (cm/min, permeability coefficient) will be derived by dividing PS_e by SA (4.2 cm^2) of the Falcon membrane.

Results

Morphology of Endothelial Monolayers

When stained with hematoxylin, the endothelial cell-covered filters showed a confluent layer of cells. Cells displayed elongated shape for the first 2-3 days after seeding but then

assumed the more characteristic cobblestone configuration. After 5 days in culture when confluency has been attained, positive silver staining at cell borders was present. Electron microscopy of seeded filters revealed a continuous layer of flattened endothelial cells. While most endothelial cell junctions appeared similar to those seen in vivo, when large numbers of cell junctions were counted, approximately five to ten percent of the intercellular borders were discontinuous with gaps of approximately 0.5-2 μm between cells. By day 7, the cells showed signs of senescence, which is indicative of the decreasing TEER values (Figure 1 and 2).

Indirect immunocytochemical assays showed that the cells stained positively for the endothelial marker, Factor VIII-related antigen, and showed consistent patterns of transferrin receptor distribution and receptor-mediated transferrin endocytosis. A photometric assay also indicates that SV-HCEC consistently expressed high activity of a BBB marker enzyme, specifically, γ -GTP.

Electrical Resistance

The most sensitive measure of the barrier function of tight junctions is TEER, which in turn is directly proportional to the permeability of the tight junctions to inorganic ions or electrolytes. The electrical resistance was measured at various days after seeding. As shown in Figure 1, transendothelial electrical resistance is highly dependent on seeding density. At low seeding densities, 25,000 and 50,000 per cm^2 , confluent monolayers are not formed which is indicative of the lower TEER readings. On the contrary, at supraconfluent seeding density of 100,000 per cm^2 , the cells appear to be overcrowded which retards their proliferation rate and may result in early cell death as indicated by the rapid decline in TEER following day 5. The optimal seeding density for this cell model appears to be 75,000/ cm^2 since TEER values steadily

increase to an elevated average of 55.3 ohms-cm^2 by day 5, which strongly suggests that confluency has been achieved. As is evident from Figure 1, cells seeded at all densities appeared to undergo senescence after day 5.

Permeability Studies

The permeability decrease of SV-HCEC with increasing time in culture is shown in Figure 3. FITC-dextran moved across unseeded filters in a manner consistent with free diffusion. Figure 3A showed that with increasing incubation time in the presence of FITC-dextran, the cells become more permeable to this hydrophilic tracer. The presence of cells on the filter clearly retards the rate of dextran movement (Figure 3A). This decrease in dextran permeability is related to the length of time the cells had been cultured on the filter prior to the transport experiment.

Figure 3B represents the transport of FITC-dextran across both the monolayer and the permeable membrane. From these data, PC was calculated, then corrected for the membrane PC ($3.9 \times 10^{-5} \pm 4.2 \times 10^{-6} \text{ cm/sec}$). From the figures, it is evident that the 2-day mean PC ($1.1 \times 10^{-5} \pm 2.8 \times 10^{-6} \text{ cm/sec}$) is significantly greater than either the 3-day ($6.3 \times 10^{-6} \pm 3.4 \times 10^{-7} \text{ cm/sec}$) or 5-day ($1.15 \times 10^{-7} \pm 3.3 \times 10^{-8} \text{ cm/sec}$) mean PC.

Figure 1. The effects of varying seeding density on transendothelial electrical resistance were conducted at various time points after seeding. The cell model depicted maximal resistance on day 5; thereafter, the cells appear to undergo senescence as indicated by the decreasing TEER value by day 7.

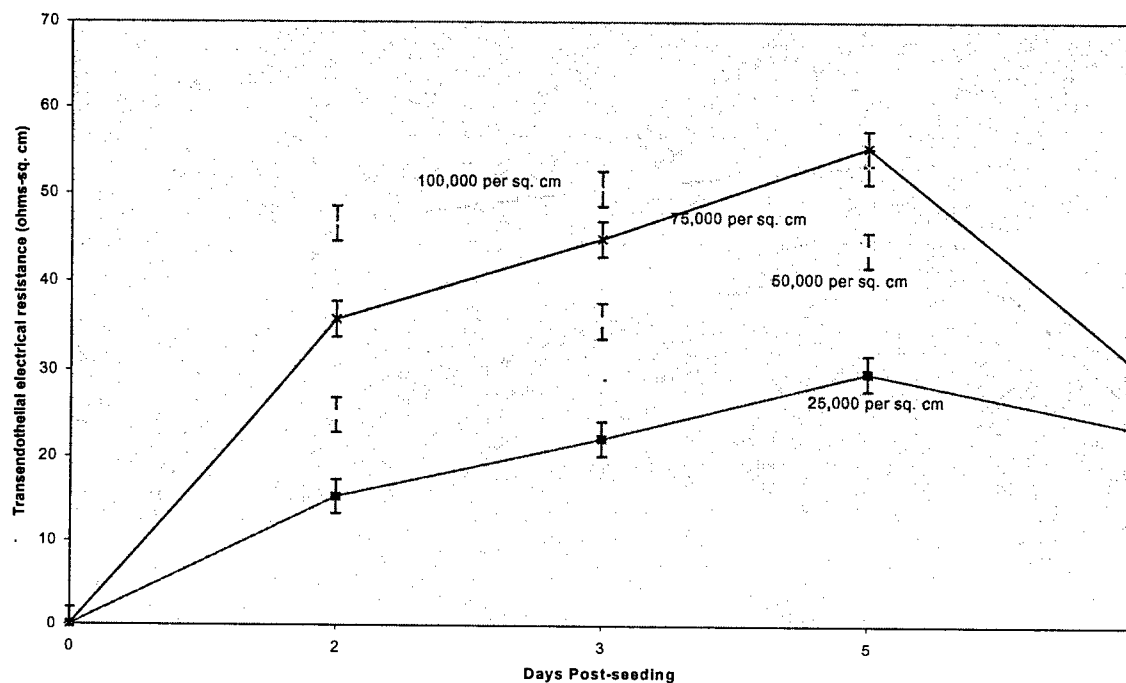


Figure 2. Cells seeded at a density of 75,000 per sq. cm (315,000 cells/culture insert) gave the optimal TEER values.

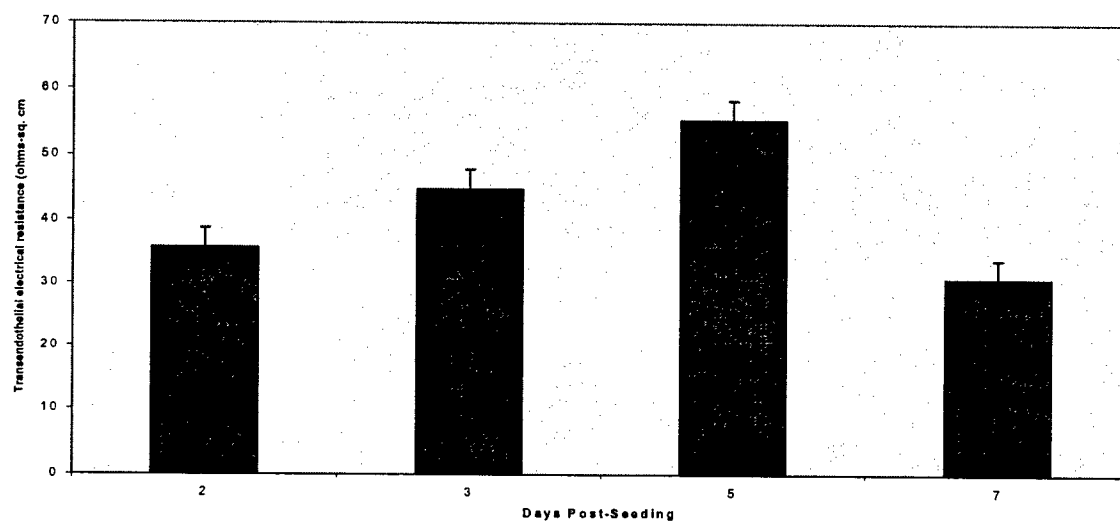
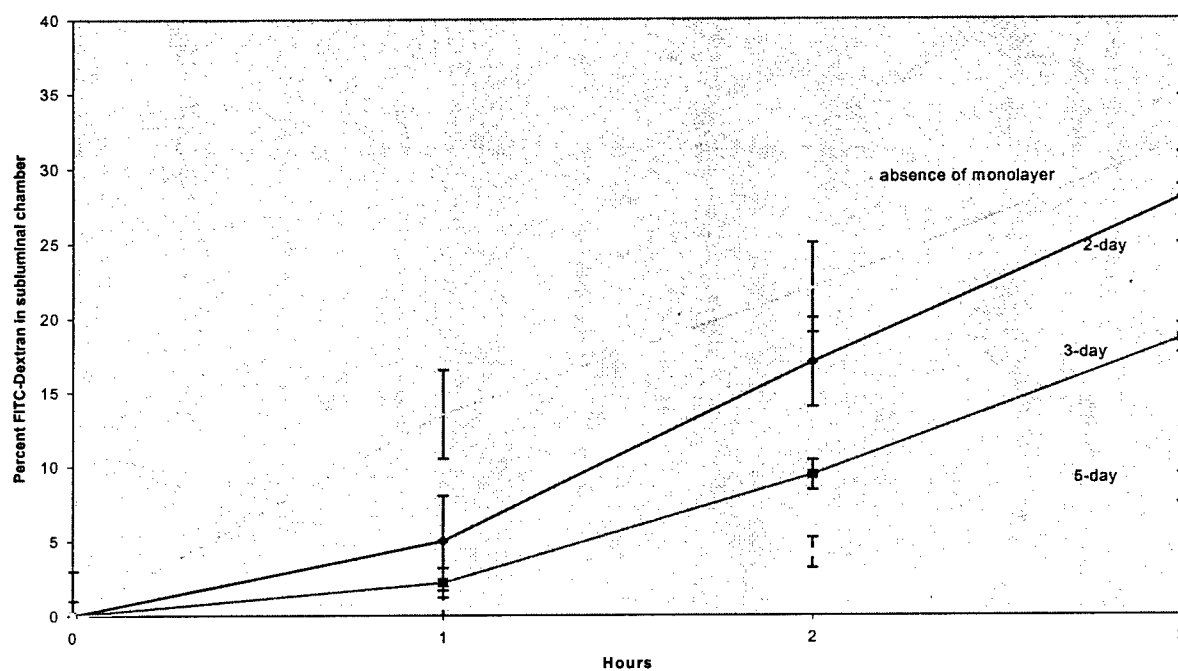
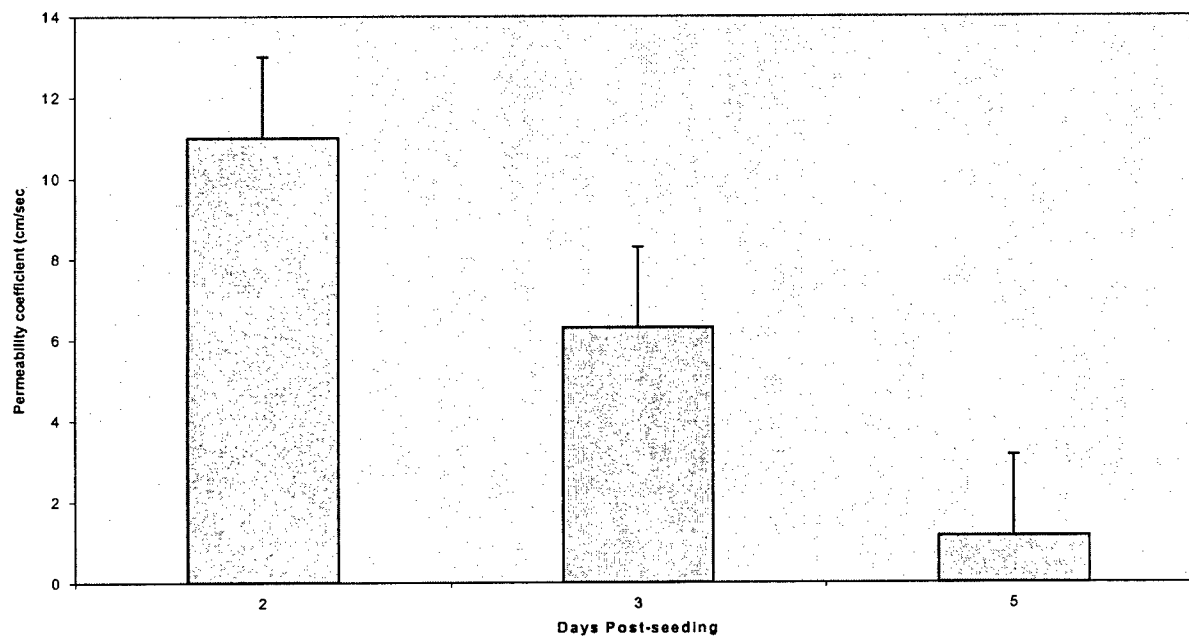


Figure 3. Effect of time in culture on permeability. A. Mass of FITC-dextran (as percent of total system dextran) transported across the monolayer vs. time. B. Permeability coefficients (cm/sec) $\times 10^6$ are shown for each postseeding time point, corrected for permeability of membrane as described in METHODS. Error bars represent SEM.

A



B



Discussion

The use of an in vitro model of permeability has a number of advantages, such as offering direct access, to luminal and subluminal fluid for analysis, being highly simplified and limited to a single cell type, and having an experimental milieu that can be defined in terms of chemical composition, pressure, and shear conditions of the incubation medium. However, the very factors that simplify the system also raise questions about its relevance to the intact in vivo situation. All models of in vitro permeability require that the endothelial cells be grown on some sort of artificial substrate, such as gelatin or fibronectin-coated polyethylene terephthalate, nitrocellulose or polycarbonate filters or treated amniotic membrane. The effect of these substrates on endothelial junctional anatomy is unclear. Additionally, the influence of the basement membrane, subcellular matrix region and interstitium on BBB permeability are altered considerably in the in vitro model.

To accurately interpret data obtained from endothelial cell monolayer systems, it is important to understand their detailed permeability characteristics. To date, these models have been analyzed by measuring transendothelial electrical resistance and/or permeability to hydrophilic, nonpolar markers, such as dextran, sucrose, inulin, horseradish peroxidase, e.g. This cell model showed TEER to be similar to that observed by other investigators (Bowman et al., 1981.; Dehouck et al., 1992); however, the permeability to FITC-dextran has been quite different from that reported in intact BBB preparations (van Bree et al., 1988).

The reasons for the discrepancy between TEER and the increased permeability to macromolecules in these models remain unexplained. A possible explanation could be due to small gaps present between cells. Careful observation of a large number of sections revealed at

least one important difference between endothelial cells seeded on the filter and those in continuous endothelium. Between 5 and 10% of the cells were not tightly joined but had small gaps (0.5-2 μm in length) between them (image not shown). A future study will need to be conducted to confirm the presence of tight junctions by means of an immunofluorescence assay employing zonula occludens-1 monoclonal antibody (R26.4C, Stevenson et al., 1986).

It also appears that the measurement of electrical resistance may not accurately reflect endothelial macromolecular permeability and thus may not be as useful as it has been in the study of epithelial barriers. Further support for this hypothesis is provided by the study of Navab et al. (1986), who showed that permeability to some macromolecules (albumin and low density lipoprotein) could be increased across an endothelial monolayer system without a concomitant increase in transendothelial resistance. This may lend support to the apparent paradox of in vitro endothelial monolayers showing "physiological" levels of electrical resistance while at the same time being much more permeable to FITC-dextran (and other macromolecules) than are in vivo endothelial barriers.

Conclusion

Morphological and physiological characterization of an in vitro cell model of immortalized human cerebromicrovascular endothelial cells was performed. The cell model depicted here resemble intact BBB endothelium with regards to cobblestone morphology and transendothelial electrical resistance; however, it appears that by physiological parameters, the system described here is 10 times "leakier" than intact BBB in vivo and therefore does not exhibit restricted diffusion as in the latter case. Other parameters, such as re-seeding experiments

to be conducted on day 3 as well as changing the volume of FHAS-conditioned medium used in the growing the cells are being considered to further develop this model so that it accurately simulates the in vivo conditions.

References

Audus KL, Bartel RL, Hidalgo JJ and Borchardt RT. The use of cultured epithelial and endothelial cells for drug transport and metabolism studies. *Pharm. Res.* (NY), 7: 435-451.

Betz AL and Goldstein GW. Polarity of the blood-brain barrier: neutral amino acid transport into isolated brain capillaries. *Science*, 202: 225-227, 1978.

Betz AL, Firth JA and Goldstein GW. Polarity of the BBB: distribution of enzymes between the luminal and antiluminal membranes of brain capillary endothelial cells. *Brain Res.*, 192: 17-28, 1980.

Bowman PD, Betz AL, Ar D, Wolinsky JS, Penney JB, Shivers RR and Goldstein GW. Primary culture of capillary endothelia from rat brain. *In Vitro Cell. Dev. Biol.*, 17: 353-362, 1981.

Dehouck MP, Jolliet-Riant P, Bree F, Fruchart JC, Cechelli R and Tillement JP. Drug transfer across the blood-brain barrier: correlation between *in vitro* and *in vivo* models. *J. Neurochem.*, 58: 1790-1797, 1992.

Durieu-Trautmann O, Foignant-Chaverot N, Perdomo J, Gounon P, Strosberg AD and Couraud PO. Immortalization of brain capillary endothelial cells with maintenance of structural characteristics of the blood-brain barrier endothelium. *In vitro Cell. Dev. Biol.*, 27A: 771-778, 1991.

Greenwood J, Pryce G, Devine L, Male DK, Dos Santos WLC, Calder VL and Adamson P. SV40 immortalized cell lines of the rat blood-brain and blood retinal barriers retain their phenotypic and immunological characteristics. *J. Neuroimmunol.*, 71: 51-63, 1996.

Muruganandam A, Herx LM, Monette R, Durkin JP and Stanimirovic DB. Development of immortalized human cerebromicrovascular endothelial cell line as an *in vitro* model of the human blood-brain barrier. *FASEB J.*, 11: 1187-1197, 1997.

Oldendorf WH, Cornford ME and Brown WJ. The large apparent work capability of the blood-brain barrier: a study of the mitochondrial content of capillary endothelial cells in brain and other tissues of the rat. *Ann. Neurol.*, 1: 409-417, 1977.

Rapoport SI. *Blood-brain barrier in physiology and medicine*. Raven Press, New York, 63-74, 1976.

Stevenson BR, Siliciano JD, Mooseker MS and Goodenough DA. Identification of ZO-1: A high molecular weight polypeptide associated with the tight junction (zonula occludens) in a variety of epithelia. *J. Cell Biol.*, 103: 755-766, 1986.

Stins MF, Prasadaraio NV, Zhou J, Arditi M, and Kim KS. Bovine brain microvascular endothelial cells transfected with SV40-large T antigen: development of an immortalized cell line to study pathophysiology of CNS disease. *In vitro Cell. Dev. Biol.*, 33: 243-247, 1997.

Trautmann OD, Bourdoulos S, Roux F, Bourre JM, Strosberg AD and Couraud PO. Immortalized rat brain microvessel endothelial cells: II. Pharmacological characterization. In (Drewes LR and Betz AL, Eds.) *Frontiers in Cerebral Vascular Biology: Transport and Its Regulation*. Plenum Press, NY, 205-209, 1993.

Van Bree JBMM, De Boer AG, Verhoef JC, Danhof M, Ginsel LA and Breimer DD. Characterization of an "in vitro" blood-brain barrier: effects of molecular size and lipophilicity on cerebrovascular endothelial transport rates of drugs. *J. Pharmacol. Exp. Ther.*, 247: 1233-1239, 1988.

**ANISOTROPIES IN VISUAL SEARCH PERFORMANCE ACROSS THE UPPER AND LOWER
VISUAL FIELDS AS A FUNCTION OF EXTRINSIC FEATURES**

**Dr. Peter D. Naegele
Graduate
Department of Psychology**

**The University of Toledo
2801 West Bancroft
Toledo, Ohio 43606**

**Final Report for:
Graduate Student Research Program
Brooks Air Force Base, TX**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC**

and

Brooks Air Force Base, TX

September, 1998

ANISOTROPIES IN VISUAL SEARCH PERFORMANCE ACROSS THE UPPER AND LOWER
VISUAL FIELDS AS A FUNCTION OF EXTRINSIC FEATURES

Dr. Peter D. Naegele
Graduate
Department of Psychology
The University of Toledo

Abstract

Previous research examining visual search performance anisotropies across the visual fields utilized stimuli which differed from one another in terms of intrinsic features (e.g. color). These studies yielded an upper and right visual field advantage in performance efficiency and accuracy. Recent research examining the same anisotropies utilizing stimuli which differed from one another based on extrinsic features (e.g. tilt) have yielded conflicting explanations. The current study attempted to resolve this conflict by controlling for a possible confound in previously used stimuli: luminance. Results indicated two major findings when luminance is controlled for: 1) reaction times are slowed drastically and 2) perceived three-dimensionality is not a useful feature during visual search. A discussion of these findings with respect to peripersonal and extrapersonal space and categorical and coordinate processing is included.

ANISOTROPIES IN VISUAL SEARCH PERFORMANCE ACROSS THE UPPER AND LOWER VISUAL FIELDS AS A FUNCTION OF INTRINSIC FEATURES

Dr. Peter D. Naegele
Graduate
Department of Psychology
The University of Toledo

Introduction

Recently, studies have examined the role of extrinsic features (e.g. tilt and light source) during visual search. Performance anisotropies for targets which differ from distractors by extrinsic features yielded a lower visual field (LVF) advantage for efficiency and accuracy (von Grunau and Dube, 1994). In particular, von Grunau and Dube (1994) found that cubes in the LVF were detected more efficiently and accurately when the target cube was tilted upward.

Their interpretation of the results was that a downward tilted cube in the LVF is the typical orientation that they are presented in the real world. An upward tilted cube in the LVF would therefore stick out (von Grunau and Dube, 1994). They further raised an issue of lighting which stated that the prevalent light source is from above and light sources originating from below would also stick out.

Previc (1990) offers a slightly different ecological explanation: tilted objects (regardless of direction of tilt) are more likely encountered in peripersonal space which is biased to the LVF. Objects in extrapersonal space (which is biased toward the upper visual field (UVF)) are rarely seen in extreme orientations (i.e. tilted) and 3D tilt is not a good cue for visual search.

Further data supporting Previc's (1990) hypothesis can be found in work by Niebauer and Christman (1998) who examined categorical (CAT) and coordinate (COOR) processing anisotropies across the visual fields. Niebauer and Christman found that CAT processing (which has shown to be related to visual search (Naegele, 1998)) is carried out more efficiently and accurately in the upper-right visual field and COOR processing (which is thought to be related to reaching and grasping behaviors (Previc, 1990)) is carried out more efficiently and accurately in the lower-left visual field.

However, Sun and Perona (1996), and Enns and Rensik (1991) have argued that perceived three-dimensionality of target items yields *faster* visual search reaction times when compared to targets perceived as less three-dimensional. They further suggest that it is the light source (i.e. above or below) that influences the efficiency of visual search of 3D items.

The main purpose of this study was to examine the effect of tilt and lighting direction of three-dimensional objects on visual search performance anisotropies across the visual fields. The secondary purpose was to examine the effect of perceived three-dimensionality of stimuli on visual search performance anisotropies across the visual fields.

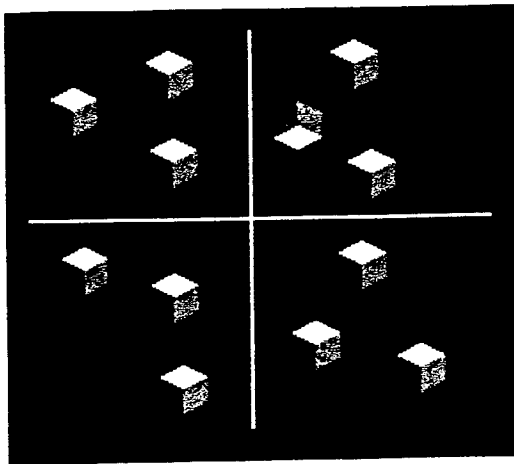
Methods

Subjects. A total of 12 subjects (10 male and 2 female) participated in the study. All but one of the subjects were either full-time or visiting military and civilian personnel at Brooks Air Force Base. None had participated in previous visual search experiments, and all were naive to the hypothesis of the study. The only requirement was that they have at least 20/30 vision with or without correction in each eye.

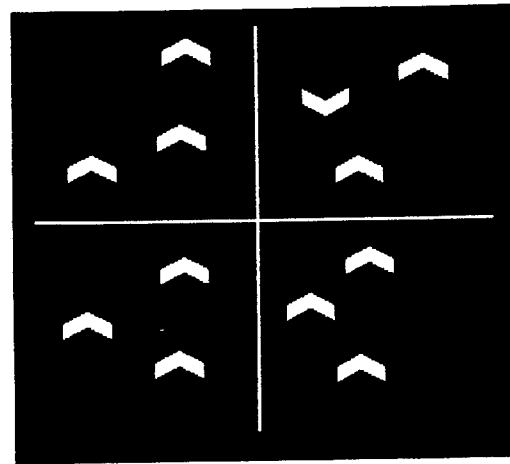
The voluntary, informed consent of subjects was obtained as required by Air Force Instructions 40-4020 and 40-403.

Stimuli and Apparatus. The stimuli were achromatic cubes generated on a Silicon Graphics Indigo 2 computer and presented on a Sony 60 Hz color monitor (model # GDM-20E21). The cubes were either large (three 1-cm² sides of varying luminance) or truncated, lit from above or below and were tilted either tilted upward or downward in relation to the subject. The truncated cubes served to control for luminance differences between targets and distractors in the large cube condition and force a judgement based on the tilt of the cubes alone.

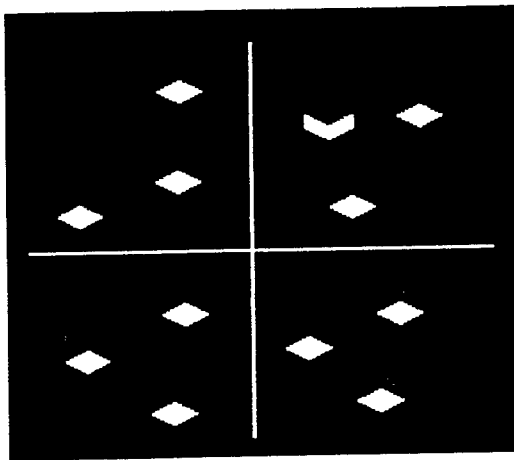
Tasks & Overall Procedures. All subjects were trained and tested in each of the four experimental conditions: 1) large cubes tilted upward or downward with a non-uniform light source, 2) truncated cubes tilted upward downward with a non-uniform light source, 3) truncated cubes tilted upward or downward with a uniform light source and 4) truncated cubes tilted upward or downward with a uniform light source and low contrast (see Figure 1). Truncated cubes were used in conditions two through four to control for luminance differences observed between targets and distractors in Condition 1.



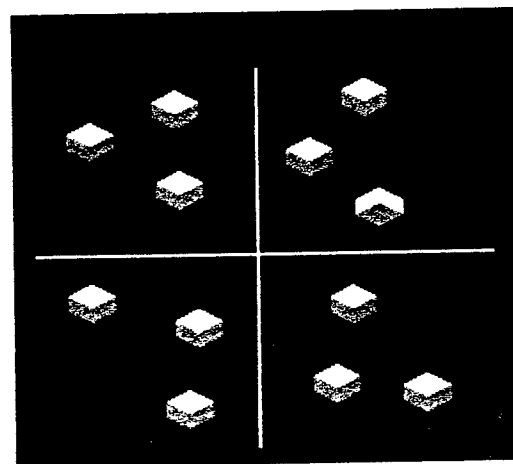
Condition 1



Condition 2



Condition 3



Condition 4

Figure 1: Sample search arrays for each of the conditions in the experiment.

The cubes in the search arrays in Condition 1 were comprised of upward tilted cubes lit from below and downward tilted cubes lit from above. The search arrays in the other three conditions were comprised of upward and downward tilted cubes lit from above *and* below.

All conditions required subjects to perform a feature search task to identify the target. In conditions one and two, the target differed from the distractor by tilt *or* light source and in conditions three and four the target differed from the distractor by tilt only.

The search field consisted of a single target and 11 fairly evenly dispersed distractors (three per quadrant with a single distractor removed in one of the four quadrants to allow for presentation of the target cube). All stimuli were located in the plane of fixation and in 3 different eccentricity rings: 2.1, 3.6, and 5.6 degrees. The overall diameter of the search field was 12.3 degrees. Subjects viewed the stimuli at a distance of 81.3 cm while their forehead and chin rested in an ophthalmologic viewing brace. During all training and testing sessions, subjects were seated in a dark room in a comfortable chair which was adjustable in height.

Subjects made a reaction time (RT) response when they found the target cube by pressing the left mouse button on a mouse (Mouse Systems Model 4Q) that was attached to the Indigo 2 computer. After making the RT response, subjects registered the target's location by moving and setting a red pointer on the video monitor using the same mouse and button. Subjects were able to pause the experiment by pressing the right mouse button when the fixation circle and cross were on the screen.

Training sessions consisted of four separate 30 minute sessions. Each session consisted of three blocks of 144 trials. Following each training block, subjects were given feedback as to their performance speed and accuracy.

Once subjects had received training for all conditions, they were given two testing sessions. Each testing session lasted 40 minutes and consisted of four blocks of 144 trials. Two of the test blocks were of one condition type and two were of another. In total, subjects received four training and two testing blocks of each condition.

Results

Preliminary analysis yielded the following means: The main effect for task showed that the fastest and most accurate responses were made under Condition 3 (mean RT = 1166 msec, accuracy = 92%) and

the slowest and least accurate responses were made under Condition 4 (mean RT = 1728 msec, accuracy = 74%). Mean RT and accuracy for Conditions 1 and 2 fell in between Conditions 3 and 4 respectively (see Figure 2). The main effect for visual field indicated that the fastest and most accurate responses were made in the UVF (mean RT = 1333 msec, accuracy = 88%) (see Figure 3). The main effect for shape showed that the fastest and most accurate responses were made for downward tilted cubes lit from below (mean RT = 1303, accuracy = 88%) (see Figure 4).

The condition by shape interaction was analyzed in the following manner: Condition 1 was analyzed separately because it utilized large versions of the upward tilted cubes lit from below and the downward tilted cubes lit from above only. Findings for Condition 1 indicated that the upward tilted cubes lit from above were responded to faster and more accurately than the downward tilted cubes lit from below (RT difference = 206 msec, accuracy difference = 3%) (see Figure 5).

For the remaining three conditions, the shapes that were responded to fastest and most accurately were the downward tilted cubes lit from below and the upward tilted cubes lit from above. However, in Condition 2, the downward tilted cubes lit from below were responded to faster than the upward tilted cubes lit from above (RT difference = 115 msec), but accuracy remained equal (see Figure 6)

The condition by visual field interaction indicated that the UVF advantage was greatest for Condition 1 (126 msec), however accuracy differences between visual fields was 3% or less across all conditions (see Figure 7).

The visual field by shape interaction indicated that the fastest and most accurate responses were made for downward tilted cubes lit from below in the UVF (mean RT = 1238, accuracy = 89%) (see Figure 8).

The condition by visual field by shape interaction was again separated into Condition 1 and Conditions 2 through 4. The visual field by shape interaction for Condition 1 indicated that the upward tilted cubes lit from below yielded the fastest and most accurate responses (RT = 1104 accuracy = 95%) (see Figure 9).

The condition by visual field by shape interaction for Conditions 2 through 4 indicated that the fastest and most accurate responses were made for downward tilted cubes lit from below presented in the UVF in Condition 3 (see Figure 10).

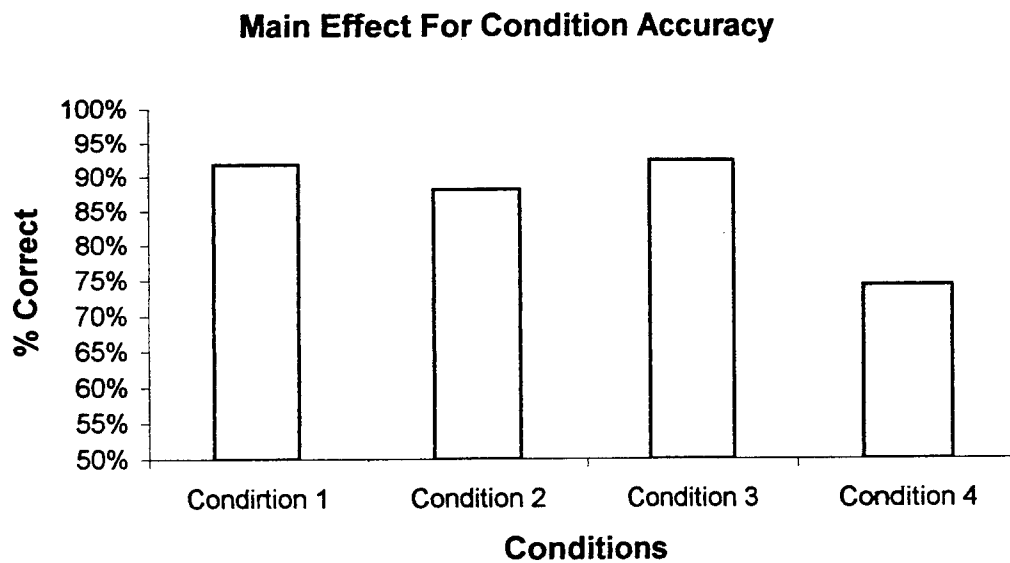
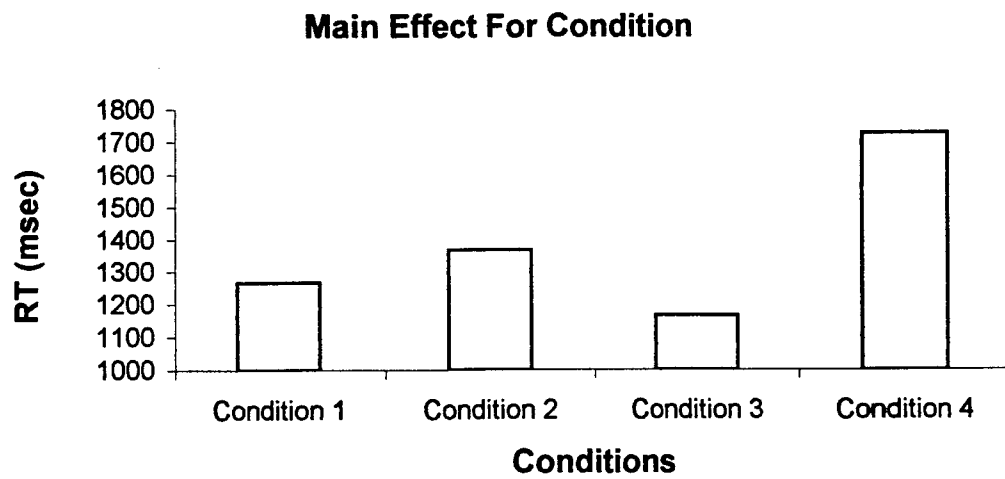


Figure 2: Main Effect for Condition RT & Accuracy

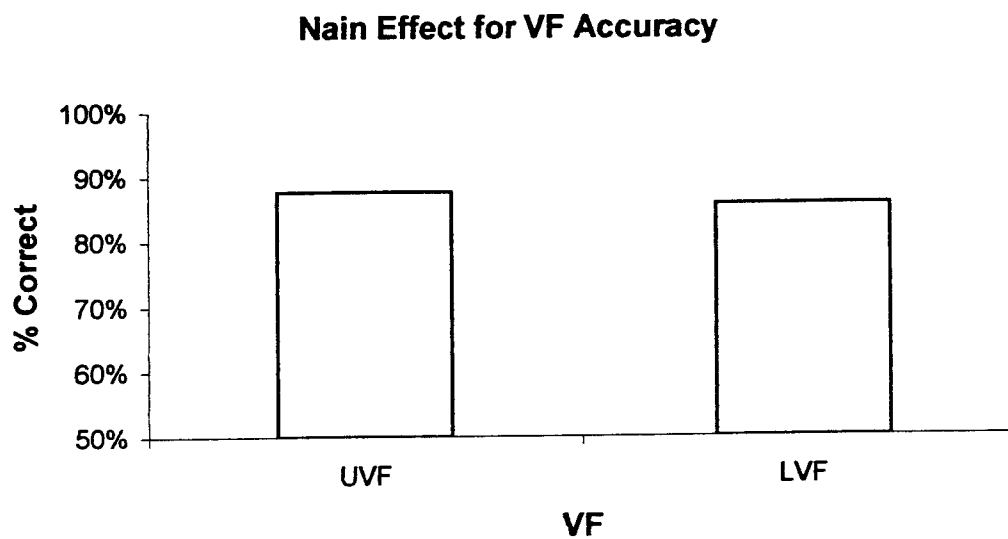
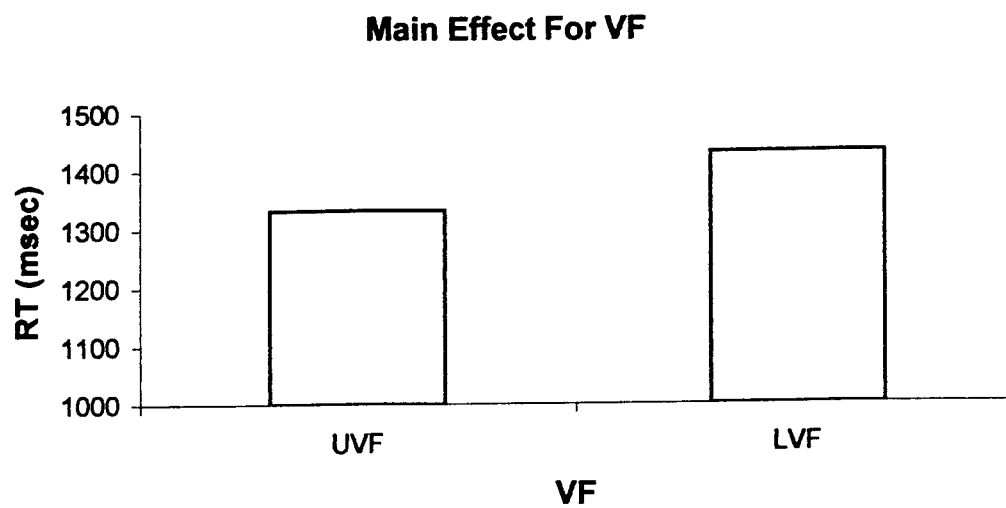


Figure 3: Main effect for VF RT & Accuracy

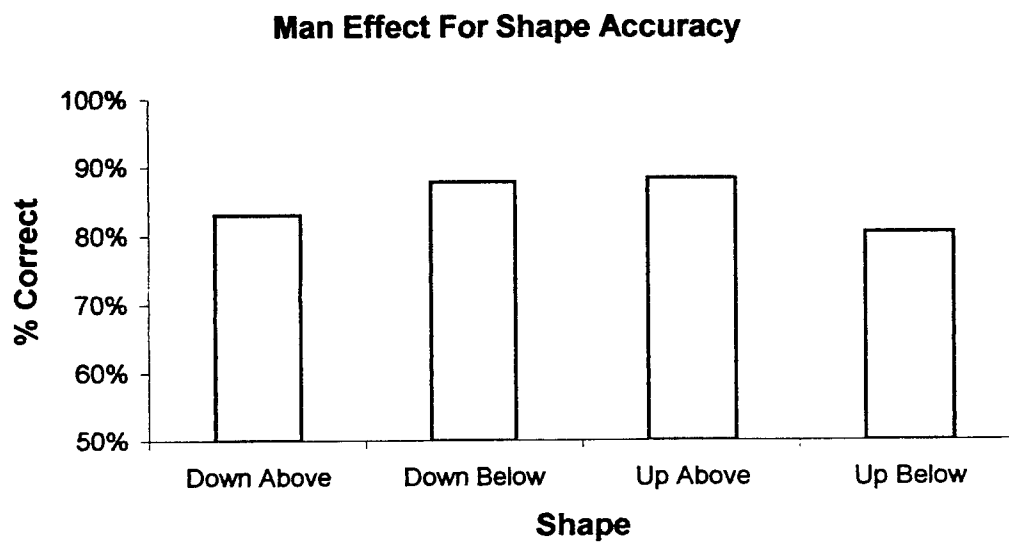
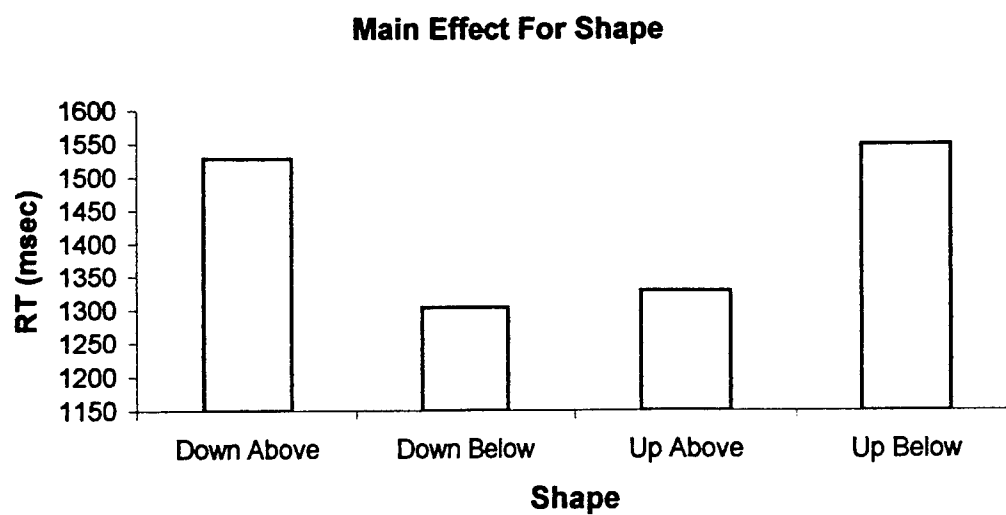


Figure 4: Main Effect for Shape RT & Accuracy

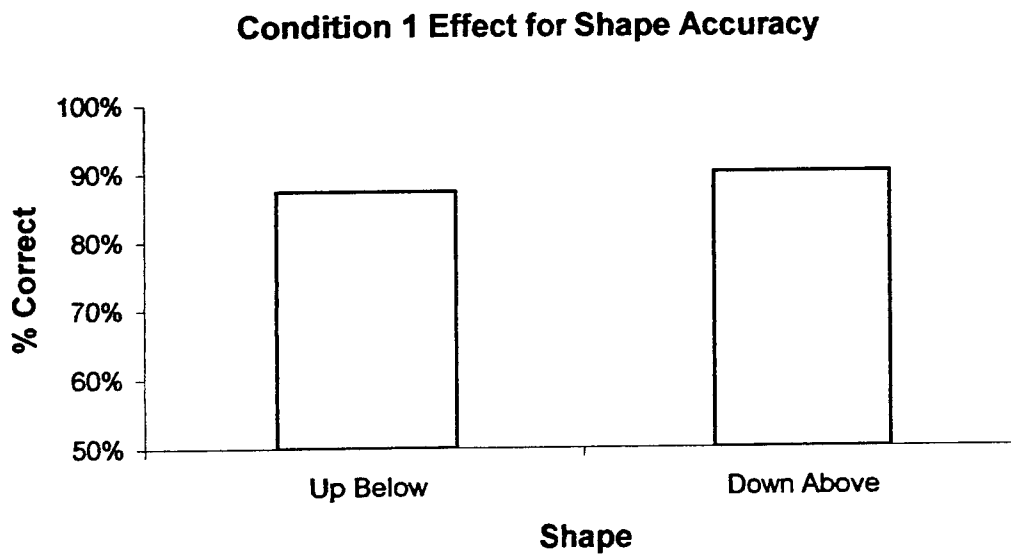
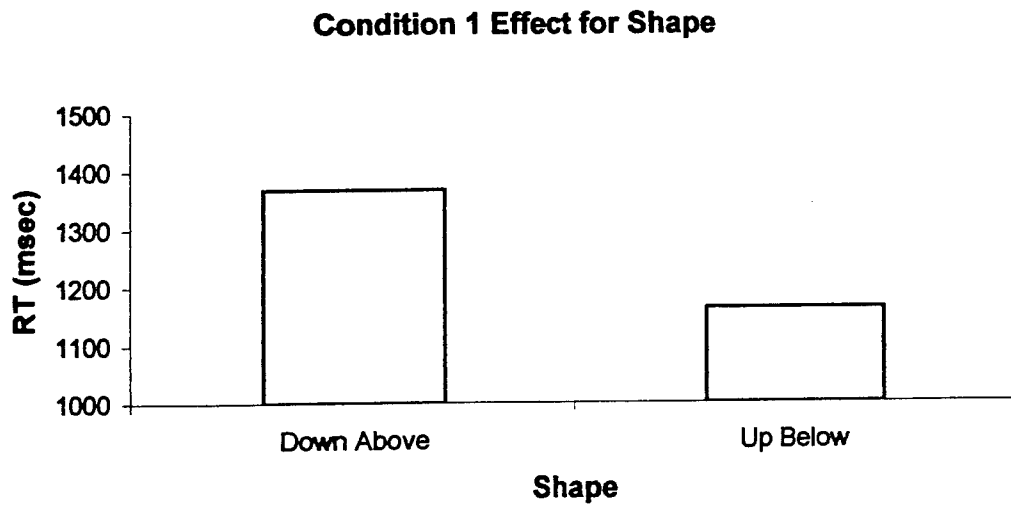


Figure 5: Condition 1 Effect for Shape RT & Accuracy

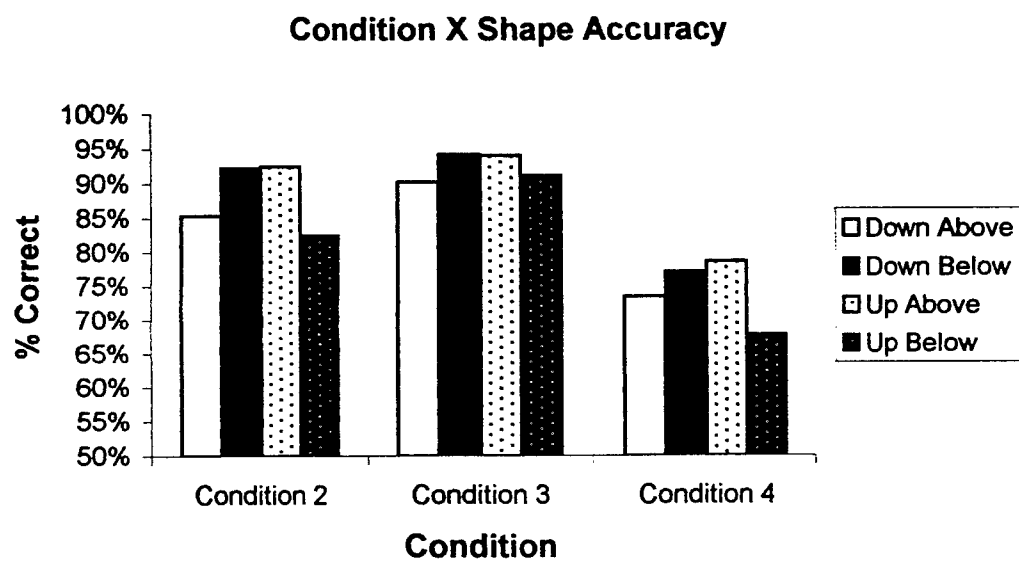
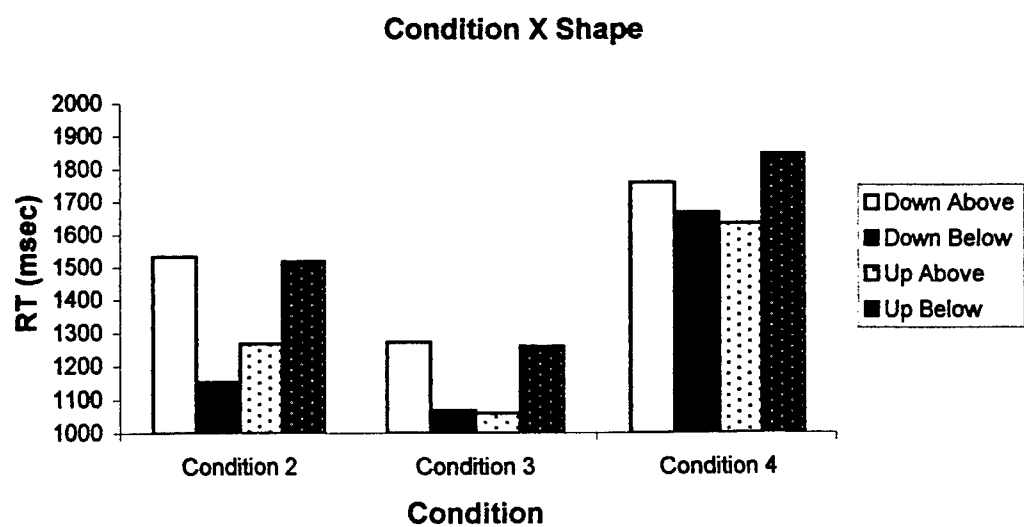


Figure 6: Condition by Shape Interaction RT & Accuracy

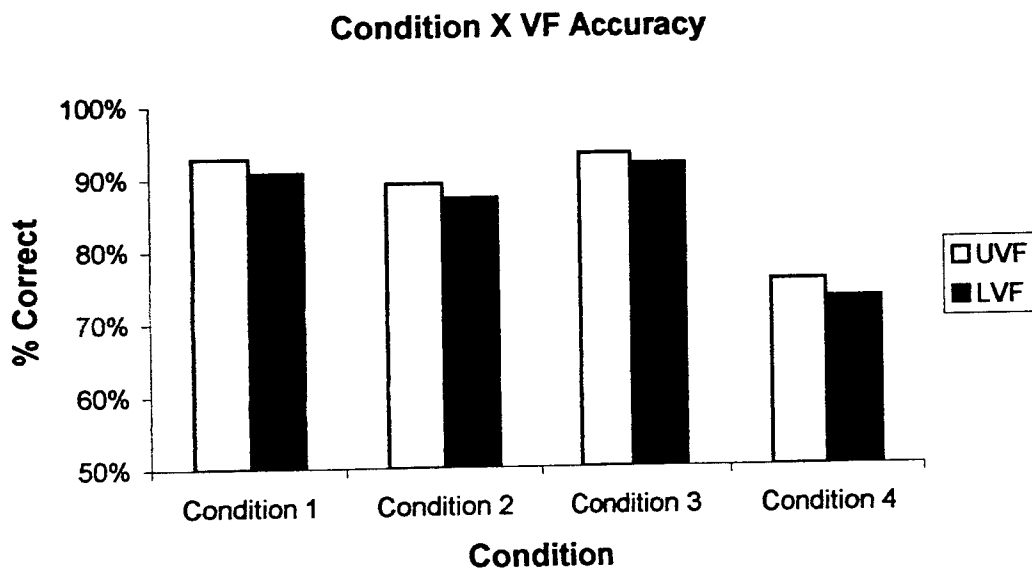
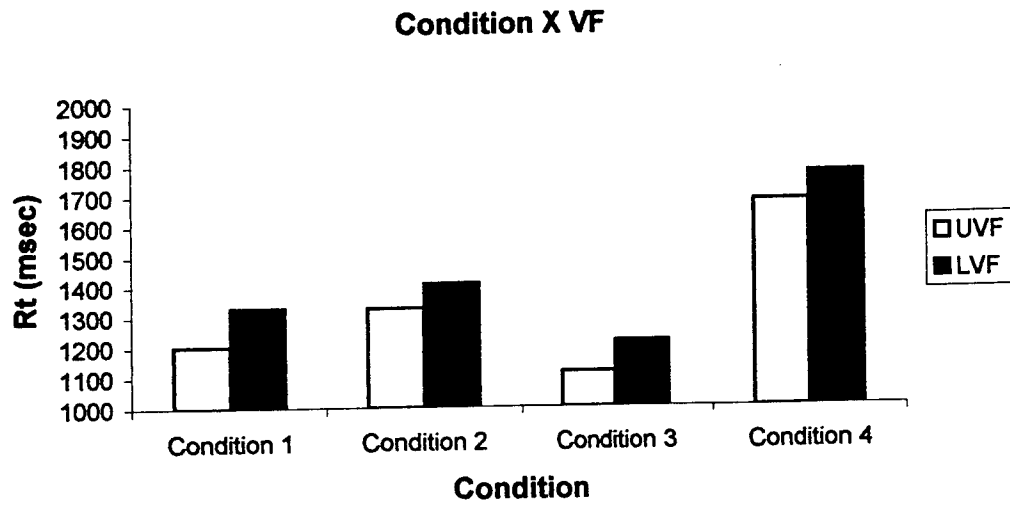


Figure 7: Condition by VF RT & Accuracy

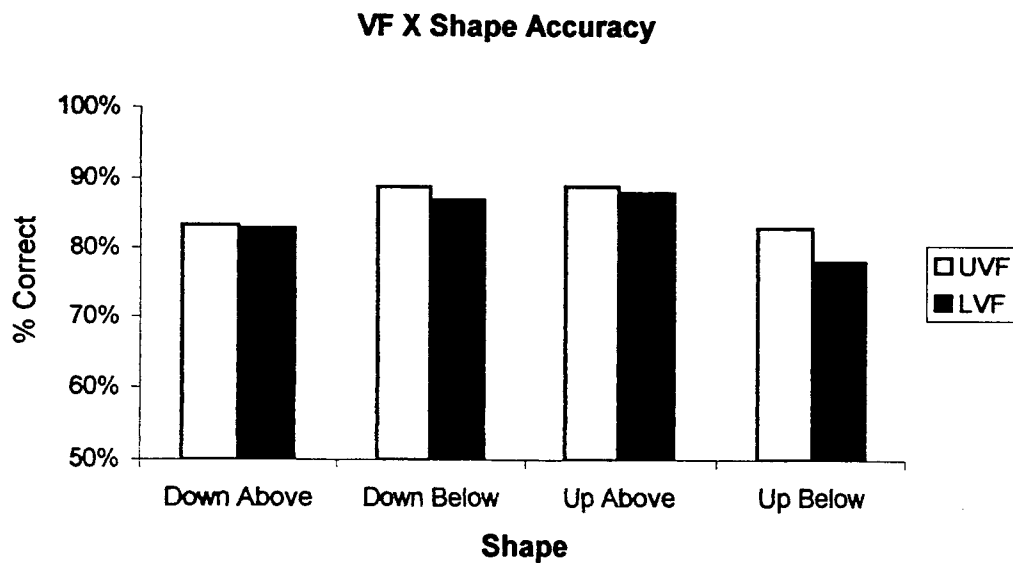
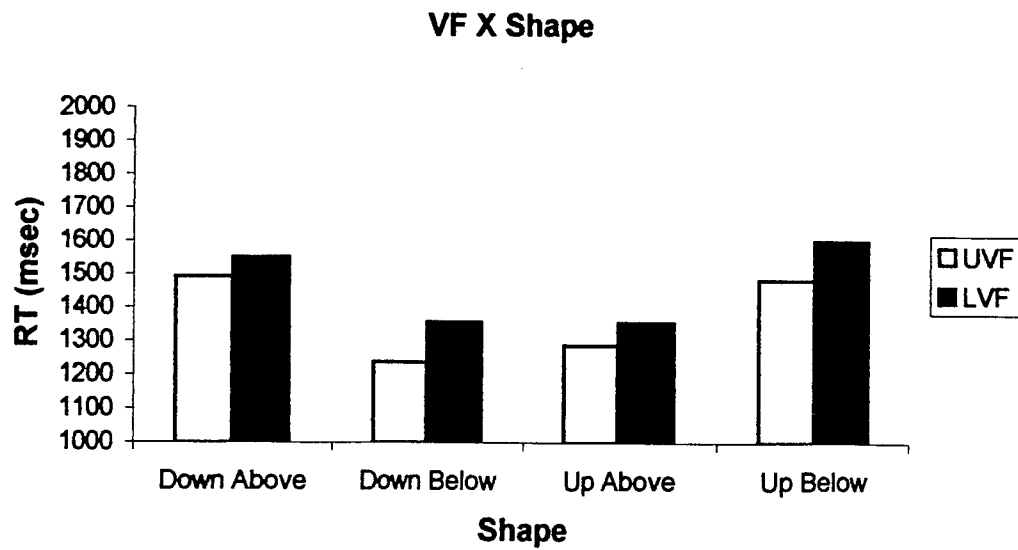


Figure 8: VF by Shape Interaction RT & Accuracy

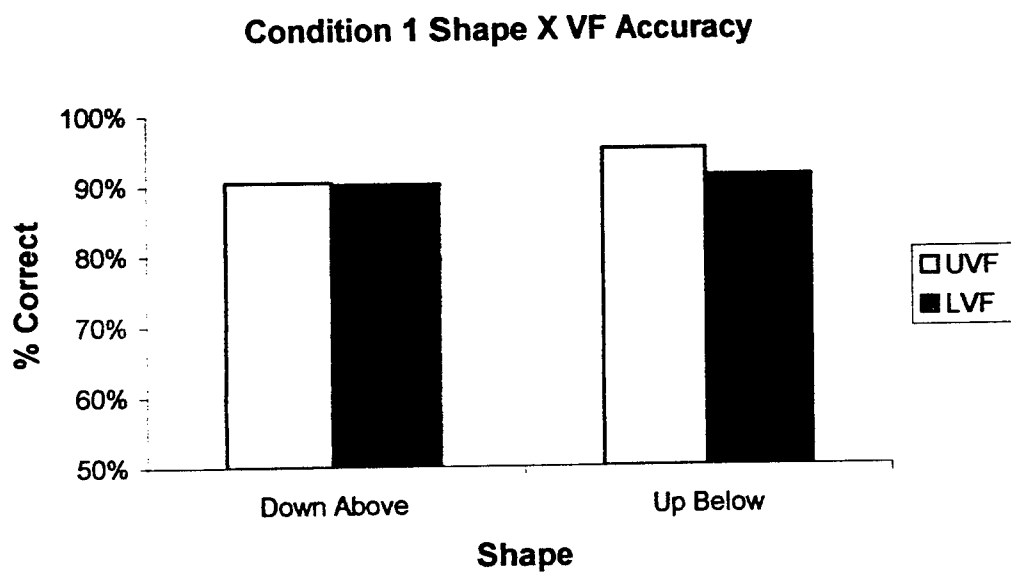
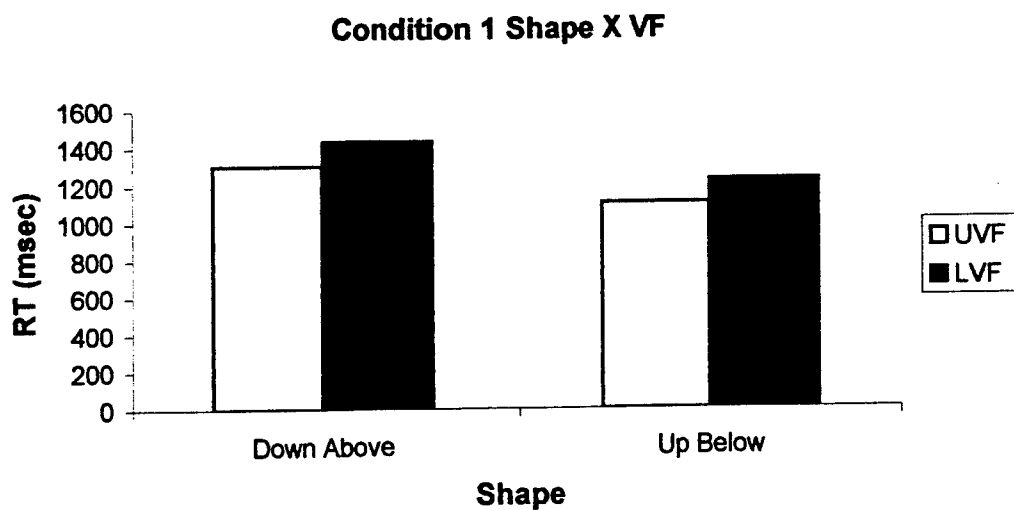


Figure 9: Condition 1 Shape by VF RT & Accuracy

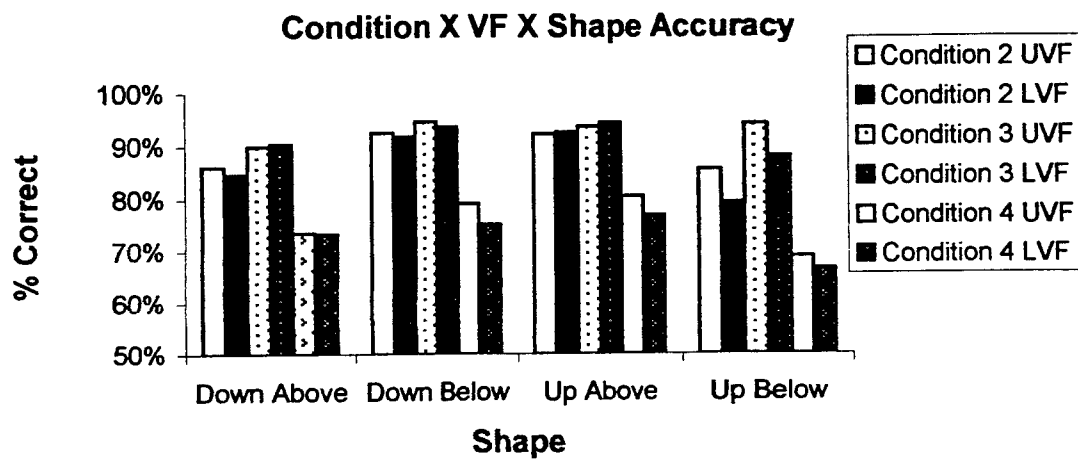
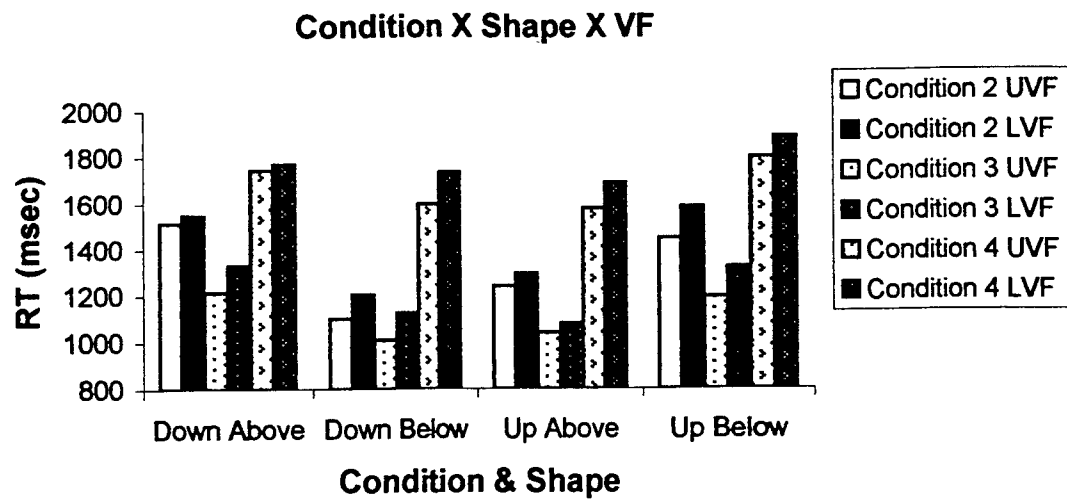


Figure 10: Condition by VF by Shape RT & Accuracy

Discussion

The most noticeable feature of these data is the very long RT's. Typically, RT's for feature search is two to three times faster than the responses observed here (Christman & Naegele, 1995; Previc, 1990; Tresiman and Gelade, 1980). This finding supports the notion put forth by Previc (1998) that three-dimensionality is not a good cue for the visual search system.

Previc's assertion is based on the notion that objects encountered during visual search are typically located in focal extrapersonal space (Previc, 1990). Objects in this location in space are rarely presented in any gross orientation; they are not tilted at extreme angles upward or downward. This being the case, attempting to conduct a search for a feature rarely encountered during visual search slows the process down considerably.

An extension of this theory is found in data collected by Niebauer and Christman (1998) examining UVF and LVF differences in categorical and coordinate processing. Their results indicate that categorical processing (used during object recognition) is performed faster in the UVF and coordinate processing (used during object localization and manipulation) is performed faster in the LVF.

The relationship to Previc's (1990) theory is that CAT processing is used during visual search for object detection (Naegele, 1998) and both search and CAT processing yield UVF performance advantages. Further, CAT processing allows for object identification regardless of an object's orientation in space: the object is identified by its parts which remain stable regardless of orientation.

Therefore, *extrinsic* properties of objects, such as those examined in the current study, are of little use to the categorical processing system. Taken to the extreme, one might assert that the categorical processing system identified only two types of stimuli in this study: large and small cubes.

However, there were processing differences observed for tilt and light source across the cubes and conditions. The cube that was responded to fastest and most efficiently in Condition 1 was the upward tilted cube lit from below presented in the UVF. An ecological explanation of this finding might be that in the real world, objects in the UVF are rarely lit by a light source from below and hence, it tends to stick out from objects lit from above (von Grunau and Dube, 1994).

However, as mentioned previously, there were differences in luminance levels between targets and distractors in this condition. It may have been the case that the targets were detected based on differences in

luminance levels and in the case of the upward tilted cube lit from below presented in the UVF, subjects could have been drawn to the darkest object.

Condition 2 controlled for this possible confound by using a truncated version of the cubes which removed luminance cues. This slowed reaction times drastically compared to Condition 1. Specifically, when direct comparisons are made to the cubes utilized in Condition 1, there is a RT difference of 259 msec. suggesting that visual search during Condition 1 was either carried out primarily by luminance cues or perceived three-dimensionality.

The cube which was responded to the fastest and most accurately in Condition 2 was the downward tilted cube lit from below presented in the UVF. As in Condition 1, the favored target cube was one which was lit from below further suggesting that the non-ecological light source may have played some role in target detection.

Conditions 3 and 4 differed from 2 based on their light source. All objects in their search arrays had a uniform light source, be it from above or below. The target cubes which were responded to the fastest and most accurately to in Conditions 3 and 4 were the downward tilted cubes lit from below and the upward tilted cubes lit from above presented in the UVF. In both Conditions, the two types of cubes yielded RT's within 50 msec. of one another.

Since all objects in the search arrays for these two conditions received light from a uniform light source, the target discrimination could have been made on luminance differences alone, as was the case in Condition 1. This explanation is further strengthened when one examines the differences in RT and accuracy between Conditions 3 and 4.

Condition 4 utilized a low contrast display for the cubes. In this condition the luminance difference between target and distractors was reduced and made the search task much more difficult to conduct.

The most interesting aspect of Conditions 3 and 4 however, is that they more closely resemble real world processing than Conditions 1 and 2. In the real world, the majority of objects in the visual scene are lighted by a uniform light source from above. Although they are not usually presented at extreme tilted orientations, when they are, a search for an object at a particular tilt may be carried out by luminance differences and *not* three-dimensionality.

This is supported by two findings in this experiment: 1) the fastest reaction times were observed when targets and distractors were lit by a uniform light source and 2) subjects reported that the truncated cubes appeared less three-dimensional than the full size cubes. In fact, during Conditions 2 through 4, most subjects reported seeing chevrons and diamonds although they were told during training that they would always be presented with cubes.

While this is contrary to previous work which has suggested the three-dimensionality plays a large role in visual search (Sun and Perona, 1996; Enns and Rensik, 1991; von Grunau and Dube, 1994), it appears that they failed to take factors such as luminance and uniform versus non-uniform light sources into account. Three-dimensionality *may* play a large role in coordinate processing, however.

When we engage in coordinate processing, we are typically interacting with objects in peripersonal space located in the LVF (Niebauer & Christman, 1998; Previc, 1990). The typical interactions include such things as object manipulation and spatial navigation which rely heavily on the extrinsic features of objects such as tilt (Previc, 1990). Hence, the greater the degree of perceived three-dimensionality, the easier it should be to identify the extrinsic features of an object. The easier it is to identify the extrinsic features of an object, the easier it should be to carry out coordinate processing.

References

Enns, J. T. and Rensik, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. Psychological Review, 98, 335-351.

Naegle, P. D. (1998). Categorical and coordinate processing within a visual search paradigm. *Unpublished Doctoral Dissertation, The University of Toledo, Toledo, Ohio.*

Niebauer, C. L. and Christman, S. D. (1998). Upper and lower visual field differences in categorical and coordinate judgements. Psychonomic Review and Bulletin, 5, 148-151.

Previc, F. H. (1990). Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. Behavioral and Brain Sciences, 13, 219-542.

Previc, F. H. (1998). Vertical field differences in feature and conjunction search for tilt and motion. *Paper in preparation.*

Sun, J. Y. and Perona, P. (1996). Preattentive perception of elementary three-dimensional shapes. Vision Research, 36, 2515-2529.

Treisman, A. and Gelade, G. (1980). A feature integration theory of attention. Cognitive Psychology, 12, 97-136

von Grunau, M and Dube, S. (1994). Visual search asymmetry for viewing direction. Perception and Psychophysics, 56, 211-220.

NEGOTIATION AT A DISTANCE: WHY YOU
MIGHT WANT TO USE THE TELEPHONE

Nicole Proulx
Graduate Assistant
Department of Psychology

University of Dayton
300 College Park
Dayton, OH 45469

Final Report for:
Summer Faculty Research Program
Wright Research Site

Sponsored by:

Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Research Site

October 1998

NEGOTIATION AT A DISTANCE: WHY YOU
MIGHT WANT TO USE THE TELEPHONE

Nicole Proulx
Graduate Assistant
Department of Psychology
University of Dayton

Abstract

Same gender dyads engaged in a 4-issue integrative bargaining task. Negotiator accountability and communication channel (face-to-face, teleconferencing, videoconferencing) were manipulated. Negotiators in the VC condition spent less time negotiating, obtained lower outcomes overall, and engaged in less logrolling than dyads in the other communication conditions. As predicted, negotiators in the TELE condition performed relatively well in the absence of visual access.

NEGOTIATION AT A DISTANCE: WHY YOU MIGHT WANT TO USE THE TELEPHONE

Nicole Proulx

In The Death of Distance, Cairnicross (1997) predicts the emergence of the networked computer as the main platform for supporting global communication and collaboration. This movement toward distributed teams, computer-mediated communication (CMC), and “groupware” is well underway in many organizations and will undoubtedly have profound organizational and psychological consequences. A growing body of theory and research emphasizes the importance of selecting the right tool for the task. For some collaborative activities, certain forms of technological support may be inappropriate and may even introduce new and substantial process losses. For example, a number of studies have demonstrated that synchronous, text-based messaging (i.e., electronic “chat”) may be a poor medium for sharing and integrating information in distributed expert groups (Strauss, 1996). Similarly, the expected benefits of electronic brainstorming tools for promoting synergy and enhancing creativity have largely failed to materialize in the laboratory. With the exception of rather large teams ($N > 12$; Gallupe, Bastianutti, & Cooper, 1991), computer-supported brainstorming groups tend to perform about as well as unsupported groups, who in turn tend to generate as many ideas as isolated individuals, if not fewer (Diehl & Strobe, 1987). Thus, the allure of inexpensive distributed collaboration should be balanced by a firm understanding of the connection between technological support and group task characteristics.

A potential task-technology mismatch involves the use of videoconferencing to support negotiation and bargaining processes. In organizational environments, the growing interest and investment in videoconferencing technology is being driven largely by a desire to “look the other person in the eye.” Presumably, this is deemed important when engaging in distributed interactions wherein trustworthiness must be either assessed or communicated. Research suggests that individuals view videoconferencing as more enjoyable, more informative, and more personal than other forms of CMC (Tang & Isaacs, 1993). However, there is also evidence that negotiators use the visual channel to dominate, deceive, threaten, or pressure their opponents and rely heavily on visual cues (e.g., staring) to assess the degree to which an opposing negotiator is trying to dominate them. In a study by Carnevale, Pruitt, and Seilheimer (1981), for example, negotiators who felt accountable to a constituent and who had direct visual access to the opposing negotiator engaged in more contentious behavior and obtained significantly lower outcomes than negotiators who could not see one another. This suggests that the strong preference for videoconferencing as a means of building trust between two distributed parties may be misguided. The current study attempts to replicate and extend the Carnevale, Pruitt, and Seilheimer findings by comparing negotiations in three different communication modalities: face-to-face (FTF), teleconferencing (TELE), and videoconferencing (VC). It was hypothesized that visual access, either FTF or VC, would lead to poor negotiation performance, relatively high levels of frustration and contentious behavior, and low levels of trust. Restricting visual access (i.e., TELE) would presumably reduce the level of contentious behavior and promote interest integration (i.e., logrolling) and the development of trust.

Method

Participants

One hundred and thirty four undergraduates, 68 females and 66 males, participated in the study in partial fulfillment of an introductory psychology research requirement.

Independent Variables

Communication. Dyads were randomly assigned to one of three communication conditions: FTF, TELE, or VC. In the FTF condition, negotiators sat across a table from one another in a private room. Negotiators used notebook computers during the negotiation that did not obstruct negotiators' visual access. In the TELE condition, negotiators were seated in private cubicles (approximately 6 sq. ft), each equipped with a small table, a headset (i.e., hands free) telephone, and a PC. In the VC condition, participants were also seated in private cubicles. The VC connection allowed negotiators to see and hear one another. The VC system provided high quality video and audio over an ISDN line. Video was close to broadcast quality with no noticeable voice lag.

As illustrated in Figure 1, the video window was located in the upper corner of the 17 in color monitor. The cameras were positioned approximately 2.5 ft from the faces of the negotiators and pointed at a slight downward angle. This provided a good view of the opponent's face and upper torso. All discussions were audio taped.

Accountability. Dyads were randomly assigned to either the accountable or not accountable condition. Using a protocol similar to that used in other studies, experimenters told accountable negotiators that their respective "supervisors" would monitor the session. Participants were told that the experimenters would be playing the supervisory roles. Experimenters instructed participants,

"Although I will not communicate with you at any point during the negotiation, I will be able to hear the discussion. As your supervisor, I will be evaluating your performance as a negotiator using a standard Negotiation Effectiveness Checklist. Based on your performance during the negotiation, you will receive a certain proportion of the total points that result from the final agreement. If I evaluate your performance positively, you will get all of the points. If I evaluate you negatively, you will only get a portion of the points. Remember, as your supervisor I want you to negotiate the best agreement for our side that you can."

In the not accountable condition, the supervisor section of the introduction protocol was omitted. Participants in this condition were simply told to, "Try to negotiate the best agreement for yourself that you can."

Negotiation Task

All dyads engaged in a 4-issue integrative bargaining task similar to those used and described extensively elsewhere (see Pruitt & Lewis, 1975; Thompson & Hastie, 1990; Thompson, 1991). As listed in Table 1, payoff schedules displayed five discrete levels per issue, with points per level listed in parentheses. The task involved the purchase of land from a state land manager by a developer interested in building an amusement park. The negotiation included one compatible issue, the earliest date for opening the park (Opening Day), and one zero-sum issue, the percentage of the park's gross that would go to the state (Gross). Negotiators' point distributions on the compatible issue were identical. For the zero-sum issue, the two negotiators' point distributions, while equal in

magnitude, reflected diametrically opposed preferences. The other two issues, the number of acres sold (Acres) and the percentage of instate park employees (Instate), combined to form a logrolling pair. Although interests on both issues were opposed, each logrolling issue was less important to one negotiator than to the other. The developer valued Acres over Instate, while the buyer preferred Instate to Acres.

As illustrated in Figure 1, all negotiators viewed a computerized version of their payoff schedule, which was located in the lower half of the computer screen. The issue levels and points per level were displayed in menus that dropped down when clicked with the mouse. Negotiators could then click on an issue level and their total points would be calculated and displayed in a box. Thus, negotiators could use this tool to compute the number of points that they would earn for any potential agreement. At no point during the session were the contents of this window transmitted to the other negotiator. Negotiators were only privy to their own payoff schedules.

Procedure

Participants arrived in same-gender pairs. During the recruitment phase, an experimenter ensured that the dyads were unfamiliar with one another. Two experimenters conducted each session. Upon arrival, participants sat at private desks to minimize pre-experimental contact and discussion. After obtaining informed consent, participants read a page of introductory material that described their role, the negotiation issues, and the method of payment. Individuals earned \$5 for their participation in the study plus a chance to win a \$100 bonus prize. The instructions informed each individual that, following completion of the entire study, the names of individuals who earned a high number of points in the negotiation would be entered into the drawing for the \$100. The instructions informed participants that the probability of a tie was quite high and that the recipient of the prize would, "...most likely be selected randomly from among the top point earners." All negotiators were told to, "try to reach a deal that results in the best possible outcome for yourself or for your company."

Experimenters then escorted participants in the FTF condition to the meeting room. Lines of communication were opened between negotiators in the TELE and VC conditions. An experimenter introduced the negotiators to one another by role and advised them as to the 30-min time limit and the consequences of not reaching an agreement. Dyads that failed to reach an agreement within 30 min received 0 points for the negotiation. Experimenters told dyads that a 2-min warning signal would be provided at the 28-min mark. Experimenters informed negotiators that they were not allowed to show their payoff schedules to their opponents during the session.

The experimenters timed the duration of the session. After reaching an agreement, participants in the FTF condition returned to their private rooms. Lines of communication were closed in the TELE and VC conditions. Participants then completed the post-negotiation questionnaire. Finally, experimenters paid and debriefed the participants.

Dependent Variables

Agreements. Analyses were conducted using time to decision and total points across all four issues as dependent variables. In addition, outcomes on the logrolling issues and the compatible issue were analyzed separately. A logrolling score was constructed where 2 = complete logrolling (each negotiator conceded all points on the lesser issue while gaining 4000 points on the more important issue), 1 = partial logrolling (each negotiator

settled for 400 points on the lesser issue while earning 3000 points on the more important issue), 0 = distributive solution (negotiators agreed to split both issues down the middle), -1 = win/lose solution (one negotiator managed to exceed his or her distributive outcome on both issues simultaneously and at his or her opponent's expense), and -2 = lose/lose solution (negotiators reached an agreement that resulted in both obtaining less than the distributive total, essentially logrolling backwards). A score for the dyad was obtained by averaging across negotiators. Performance on the compatible issue was assessed using the total points obtained on that issue.

Post negotiation questionnaire. The post-negotiation questionnaire assessed perceived issue importance and conflict of interest, frustration, trust, and various contentious behaviors. [Note. Specific descriptions of these variables were omitted for the sake of brevity].

Results

Agreements

All of the dyads reached agreement within the time allotted. Averaged across all conditions, the duration of the negotiation was 9.07 min (SD = 4.47). A univariate analysis of variance (ANOVA) using time to decision as the dependent variable¹ obtained a significant main effect for communication condition, $F(2, 53) = 4.11, p < .05$. Average time to decision (SD in parentheses) was 11.10 min (4.09), 8.98 min (4.77), and 7.13 min (3.72) for the FTF, TELE, and VC conditions, respectively. Pairwise comparisons using a Tukey HSD test revealed a significant difference between the FTF and VC conditions, with negotiators interacting significantly longer in the FTF condition.

Averaging across all conditions, dyads earned 13,400 points² (SD = 1150), a value significantly greater than the distributive total (12,800), $t(64) = 4.20, p < .001$, but significantly less than the integrative total (15,200), $t(64) = -12.61, p < .001$. An ANOVA using total points as the dependent variable revealed a significant communication main effect, $F(2, 53) = 3.53, p < .05$. The average point total (SD in parentheses) was 13,800 (693), 13,487 (1431), and 12,905 (1035) in the FTF, TELE, and VC conditions, respectively. Pairwise comparisons using a Tukey HSD test obtained a significant difference between the FTF and VC conditions only, with negotiators earning significantly more points overall in the FTF condition. The ANOVA also obtained a significant main effect for gender, $F(1, 53) = 4.62, p < .05$. Female dyads earned significantly fewer points overall than male dyads (Ms = 13,006 and 13,697, respectively; SDs = 1191 and 1074, respectively).

Analyzing the logrolling issues separately, 12% of the dyads successfully logrolled, 20% partially logrolled, 17% reached a distributive solution, 45% achieved a win/lose solution, and the remaining 6% arrived at a lose/lose solution. An ANOVA using the logrolling score as a dependent variable obtained a significant main effect for communication, $F(1, 53) = 4.06, p < .05$. The average logrolling score (SD in parentheses) was -0.33 (1.06), 0.39 (1.34), and -0.48 (0.93) in the FTF, TELE, and VC conditions, respectively. Pairwise comparisons using a Tukey HSD test obtained a significant difference between the TELE condition and both the FTF and VC conditions. Table 2 lists the percentages of logrolling, distributive, win/lose, and lose/lose solutions by communication condition. Significantly more logrolling occurred in the TELE condition than in either the FTF or VC condition.

Trust

On the post-negotiation questionnaire, participants rated the degree to which they trusted the opposing negotiator (trust-other) and the perceived level of trust afforded them by their opponent (trust-you). The overall means for trust-other and trust-you were 6.79 ($SD = 1.30$) and 6.71 ($SD = 1.13$), respectively. An ANOVA using trust-you ratings as the dependent variable revealed no significant effects. An ANOVA using trust-other ratings as the dependent variable revealed a significant three-way interaction, $F(2, 53) = 3.39, p < .05$, illustrated in Figure 2. Simple comparisons using the pooled error term revealed no significant communication differences for accountable male dyads ($M = 6.59$, averaged across communication conditions). For accountable female dyads, trust-other varied by communication condition, with significantly greater levels of trust reported in the TELE condition versus the FTF condition, $t(10) = 3.05, p < .05$.

Conversely, in the not accountable condition, female negotiators' trust-other levels did not differ significantly across communication conditions ($M = 7.17$, averaged across communication conditions), however, male negotiators' trust-other levels were significantly greater in the TELE condition versus the VC condition, $t(8) = 2.61, p < .05$. The only significant gender difference occurred in the not accountable/VC condition, $t(8) = -2.44, p < .05$, with female negotiators reporting significantly greater trust-other scores than male negotiators.

Contentious Behavior

On the post-negotiation questionnaire, participants rated the extent to which their opponent engaged in a variety of positive and negative behaviors. A series of ANOVAs using each behavior as a dependent variable revealed a number of significant effects. A significant main effect for communication condition emerged for the item "made concessions," $F(2, 53) = 3.87, p < .05$. Means for this item (SD s in parentheses) were 5.40 (1.30), 6.20 (0.75), and 5.31 (1.40) for FTF, TELE, and VC, respectively. Pairwise comparisons using a Tukey HSD test revealed that negotiators in the TELE condition perceived that their opponent made significantly more concessions than did negotiators in the VC condition.

The ANOVA also revealed significant three-way interactions for the items, "listened to my suggestions/interests," $F(2, 53) = 4.42, p < .05$, and "appeared genuinely friendly," $F(2, 53) = 3.33, p < .05$, with patterns similar to that obtained for trust-other. The means for these items are listed in Table 1.

Correlational Analyses

Table 4 lists the correlations between the post-negotiation items pertaining to opponents' behaviors and both perceived frustration and trust-other. Most of the items were significantly correlated with frustration such that, the greater perceived levels of negative behavior were associated with greater frustration and less trust. Further analyses examining the potential moderational effects revealed that communication condition significantly moderated the relationship between reported frustration and responses on the item, "appeared inflexible," $F(2, 59) = 4.55, p < .05$. Although significant correlations were observed between this item and frustration in both the VC, $r(20) = +.68, p < .01$, and FTF, $r(20) = +.61, p < .01$, conditions, the two were not correlated in the TELE condition, $r(22) = -.08$.

Discussion

Based on the Carnevale, Pruitt, and Seilheimer (1981) findings related to visual access and integrative bargaining, it was predicted that dyads in the TELE condition would have a more accurate impression of their opponent's interests, engage in less contentious behavior, and achieve better joint outcomes than dyads in the FTF

and VC conditions. Although the communication channels investigated in the current study certainly differ along multiple dimensions, the relative absence of major differences between FTF and VC negotiations is a strong indication that the two were qualitatively similar. Thus, it seems unlikely that the poor performance of VC negotiators was due solely to the quality of the visual channel.

Current results indicated that VC may not be the optimal medium for integrative bargaining. Negotiators in the VC condition spent less time negotiating, obtained lower outcomes overall, and engaged in less logrolling than dyads in the other communication conditions. As predicted, negotiators in the TELE condition performed relatively well in the absence of visual access. The study also suggested that the effects of communication condition on perceptions of trust are a complex function of negotiator gender and accountability. Future research should attempt to isolate the mechanisms that may be driving this interaction effect. Finally, the moderational effect of communication condition on the relationship between frustration and perceptions of inflexibility suggests that such perceptions may be strongly influenced derived by visual cues. Thus, one explanation for the superior performance of TELE negotiators may be that they were not exposed to visual indicators of inflexibility and positional commitment.

References

- Cairncross, F. (1997). The Death of Distance. Boston: Harvard Business School Press.
- Carnevale, Pruitt, & Seilheimer (1981). Looking and competing: Accountability and visual access in integrative bargaining. Journal of Personality and Social Psychology, 40, 111-120.
- Diehl, M., & Strobe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. Journal of Personality and Social Psychology, 53, 497-509.
- Gallupe, R.B., Bastianutti, L., & Cooper, W.H. (1991). Unblocking brainstormers. Journal of Applied Psychology, 76, 137-142.
- Pruitt, D. G., & Lewis, S. A. (1975). Development of integrative solutions in bilateral negotiation. Journal of Personality and Social Psychology, 31, 621-633.
- Straus, S.G. (1996). Getting a clue: The effects of communication media and information distribution on participation and performance in computer-mediated and face-to-face groups. Small Group Research, 27, 115-142.
- Tang, J.C., & Isaacs, E. (1993). Why do users like video: Studies of multimedia-supported communication. Computer Supported Cooperative Work, 1, 163-196.
- Thompson, L. L. (1991). Information exchange in negotiation. Journal of Experimental Social Psychology, 27, 161-179.
- Thompson, L. L., & Hastie, R. (1990). Social perception in negotiation. Organizational Behavior and Human Decision Processes, 47, 98-123.

Footnotes

¹ All statistical tests used $\alpha = .05$. For post hoc comparisons, $\alpha_{FW} = .05$. All ANOVAs included communication, accountability, and gender as independent variables

² Reported point values were rounded to the nearest whole number.

Table 1

Park Developer and Land Manager Payoff Schedules for the 4-Issue Task

Park Developer's Payoff Schedule			
Acres	Earliest Opening Day	% of Gross	% Instate Employees
110 Acres (4000)	1 Year (2400)	2% (2400)	10% (1600)
100 Acres (3000)	2 Years (1800)	4% (1800)	20% (1200)
90 Acres (2000)	3 Years (1200)	6% (1200)	30% (800)
80 Acres (1000)	4 Years (600)	8% (600)	40% (400)
70 Acres (0)	5 Years (0)	10% (0)	50% (0)

Land Manager's Payoff Schedule			
Acres	Earliest Opening Day	% of Gross	% Instate Employees
110 Acres (0)	1 Year (2400)	2% (0)	10% (0)
100 Acres (400)	2 Years (1800)	4% (600)	20% (1000)
90 Acres (800)	3 Years (1200)	6% (1200)	30% (2000)
80 Acres (1000)	4 Years (600)	8% (1800)	40% (3000)
70 Acres (1600)	5 Years (0)	10% (2400)	50% (4000)

Note. Points for each issue are listed in parentheses. All payoff schedules used in the study presented levels for both players arranged from most to least preferred.

Table 2

Logrolling Performance by Communication Condition

Outcomes	FTF	Telephone	Videoconference
Complete Logrolling	9%	26%	0%
Partial Logrolling	14%	26%	19%
Distributive Solution	9%	17%	24%
Win/Lose Solution	68%	22%	48%
Lose/Lose Solution	0%	9%	9%

Table 3

Mean Responses to Contentious Behavior Items by Communication,
Accountability, and Gender

	Not Accountable		Accountable	
	Male	Female	Male	Female
FTF				
Trust-Other	6.25	7.50	6.60	6.42
Listened-Other	6.58	7.88	7.70	6.83
Friendly-Other	6.92	7.50	7.40	7.08
Friendly-Own	7.33	7.25	7.60	7.33
Telephone				
Trust-Other	7.58	6.70	6.67	7.67
Listened-Other	8.08	7.30	7.75	8.08
Friendly-Other	7.83	6.80	7.08	8.42
Friendly-Own	7.75	7.30	6.75	8.42
Videoconference				
Trust-Other	5.75	7.33	6.50	6.33
Listened-Other	6.97	7.27	8.20	7.67
Friendly-Other	6.38	7.75	8.00	7.67
Friendly-Own	5.83	7.92	7.60	7.25

Table 4

Simple Correlations of Behavioral Items With
Frustration and Trust-Other

Items	r(frus)	r(trust)
Made concessions	-.32*	+.29*
Communicated interests	-.14	+.41*
Offered Alternatives	-.43*	+.40*
Interested in reaching agreement	-.50*	+.60*
Listened to my suggestions	-.26*	+.50*
Appeared friendly	-.35*	+.66*
Purposely misled	+.39*	-.50*
Used verbal pressure tactics	+.29*	-.28*
Used nonverbal pressure tactics	+.42*	-.50*
Appeared inflexible	+.40*	-.46*
Appeared mainly interested in self	+.23	-.27*

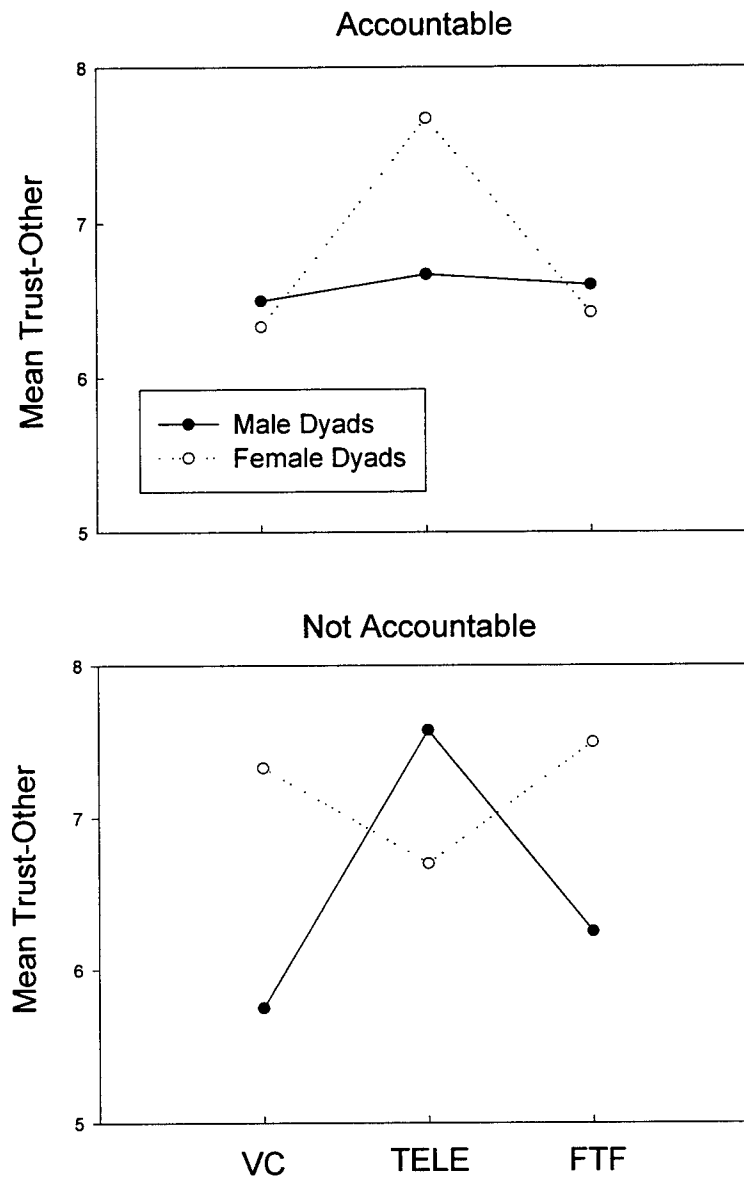


Figure 1. Mean Responses on Trust-Other by Communication, Accountability and Gender

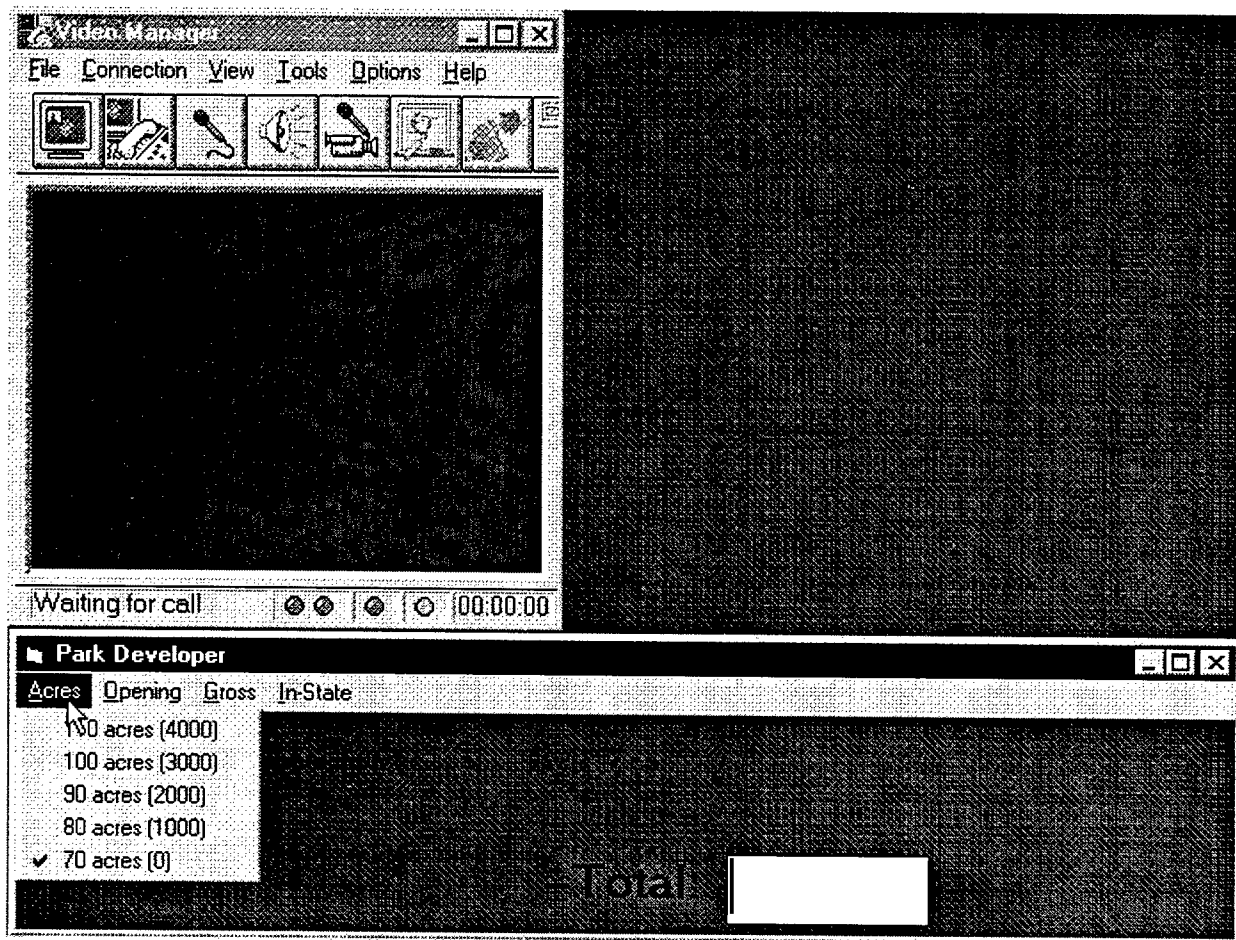


Figure 2. Screen Shot of Video Window and Computerized Payoff Schedule

UNDERSTANDING DISAGREEMENT ACROSS RATING SOURCES:
AN ASSESSMENT OF THE MEASUREMENT EQUIVALENCE OF RATERS

M. Kathleen Sheehan
Graduate Student
Department of Psychology

Texas A&M University
Department of Psychology
College Station, TX 77843-4235

Final Report for:
Graduate Student Research Program
Air Force Research Laboratory Site
Brooks Air Force Base, San Antonio, TX

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Air Force Research Laboratory Site
Brooks Air Force Base, San Antonio, TX

September 1998

—
UNDERSTANDING DISAGREEMENT ACROSS RATING SOURCES:
AN ASSESSMENT OF THE MEASUREMENT EQUIVALENCE OF RATERS

M. Kathleen Sheehan
Graduate Student
Department of Psychology
Texas A&M University

Abstract

Relations among ratings of multiple performance dimensions by multiple rating sources are examined. Confirmatory factor analysis is used to evaluate a series of models representing levels of measurement equivalence across rating sources. Results indicate the relative impact of dimension, rating source, and unique effects on ratings. Implications for practice (e.g., 360° feedback systems) are considered.

UNDERSTANDING DISAGREEMENT ACROSS RATING SOURCES:
AN ASSESSMENT OF THE MEASUREMENT EQUIVALENCE OF RATERS

M. Kathleen Sheehan

Few would argue that individual performance appraisal is a commonplace, if not vital, component of the majority of human resource management (HRM) systems in today's organizations. The frequency and/or importance of performance appraisal to HRM systems is reflected in the I/O and HRM scientific literature both in terms of the amount and the longevity of the research interest. Further, it would be difficult to argue that the vast majority of performance appraisal involves the subjective rating of an individual's performance by another individual. Within this context, traditional approaches to performance appraisal in organizations have typically emphasized a top-down approach in which supervisors provide evaluations of subordinate performance. Certainly differences exist in the extent to which subordinates are permitted to participate in the process, but this participation has traditionally been limited to the feedback and goal setting process as opposed to the actual evaluation of performance. Recent trends in performance appraisal (e.g., 360° degree feedback), however, have moved away from single sources of evaluative performance information and incorporate multiple rating sources. These sources may include one or more supervisors, peers, subordinates, and in some cases customers or clients.

An important issue for performance appraisals using multiple rating sources is the extent to which the various sources converge. Here two questions may be raised. First, to what extent do the ratings from multiple sources rating a particular individual converge or agree? Second, to what extent should the multiple rating sources be expected to agree? The first question is an empirical question while the second is a conceptual question that can only be addressed within a theoretical context.

A good deal of literature has focused on the first question (e.g., Harris & Schaubroeck, 1988; Mabe & West, 1982; Viswesvaran, Ones & Schmidt, 1996), with much of this literature indicating that individuals view their own performance quite differently than others view their performance (Ashford, 1989; Harris & Schaubroeck, 1988). A meta-analytic review of self and other ratings of performance revealed that self-ratings were only moderately correlated with supervisors ($r = 0.35$ for self-supervisor agreement) or peers ($r = 0.36$ for self-peer agreement), while correlations between peer and supervisory

ratings were substantially higher ($r = 0.62$) (Harris & Schaubroeck, 1988). Similarly, Viswesvaran et al. (1996) also demonstrated higher levels of agreement between different supervisors rating the same individual than between different peers rating the same individual.

While the literature to date generally provides ample evidence with respect to levels of agreement or disagreement, very little attention has focused on elucidating the nature of this agreement or disagreement. Traditional approaches to performance ratings, drawing largely from classical test theory, have viewed ratings as being comprised of 'true score' and either systematic or random error (e.g., Wherry, 1983). From this perspective, agreement across rating sources may be viewed somewhat analogously to reliability and thus the level of interrater agreement may serve as an indication of the quality of the performance measurement. Recently, however, it has been increasingly recognized that level of agreement (or lack thereof) may reflect not only true and error components, but multiple true score components (i.e., raters may be rating different aspects of performance).

Various explanations for the low levels of agreement across rating sources have been postulated. Consistent with classical test theory explanations, for example, it may be that there is systematic bias operating in these systems, or raters may be unreliable. It has also been suggested that the low levels of agreement may represent the different perspectives that the different raters bring to a 360-degree feedback system. Multiple raters may provide a richer and more comprehensive representation of an individual's performance (Buckner, 1959). The difference in ratings may be due to different opportunities to observe the individual's work behavior (Murphy & Cleveland, 1995), as peers tend to have more opportunities to observe the individual employee's performance than do supervisors. This difference may also be a result of the raters actually observing different behavior (London & Beatty, 1993) or the individual displaying different behaviors in different situations due to the differing dynamics of relationships with peers, subordinates, and supervisors (Funder, 1987). Along these lines, Church (1997) found that individuals high in self-monitoring had more agreement with other raters, as they were better at assessing their performance on the job.

The rating sources may also differ in how they arrive at their evaluations of the target employee.

Different rating sources may be focusing on the aspects of the individual's performance that are the most salient and relevant to them. For instance, subordinates are often more attuned to some supervisory behaviors than other supervisors due to the influence that supervisors have on their subordinates (Maurer, Raju, & Collins, in press). Also, there may be differences in processing that lead to disagreement between rating sources. Attributional tendencies, such as self-serving bias and actor-observer effects, have been shown to contribute to low agreement between rating sources (Fahr & Dobbins, 1989a). Additionally, Fahr and Dobbins (1989b) found that differences in social comparison information among raters impacted the level of agreement in performance ratings between rating sources.

Additionally, different rating sources may have different standards of performance. Recent research by Maurer et al. (in press) indicated that peer raters might have higher standards of performance, resulting in lower performance ratings compared with other rating sources. Campbell and Lee (1988) noted that rating sources might differ in defining effective performance. If the rating sources are each defining effective performance in a way that is most advantageous for them, it is possible that each rating source is measuring a different performance construct.

Lastly, the low levels of agreement in agreement between raters in a multisource feedback system may suggest that the raters are measuring the individual's performance differently. Rating may differ due to how the scale was used in different populations (Maurer et al., in press), suggesting that the rating sources are not equivalent in measuring the target employee's performance. Similarly, the performance rating instruments may differ in terms of the performance constructs being measured across different rating groups, in that items may relate to different performance constructs for each rating source (Fecteau, 1998).

The purpose of the present study is to provide an empirical examination of the nature of rating differences across multiple rating sources. Toward this end a series of latent variable models representing differing levels of measurement equivalence across rating sources were evaluated. A general model (which serves as the basis for all of the models evaluated) is presented in Figure 1. This model incorporates 24 performance ratings. These ratings correspond to ratings of eight performance dimensions, each rated by three different rating sources (self, supervisor, and peer). In addition, the model postulates 8 latent

variables corresponding to each of eight performance dimensions assessed. Each of the 24 ratings loads on the performance dimension it taps, thus each dimension has three associated ratings. Based on this general model three models representing three levels of measurement equivalence are initially evaluated and compared (Nunnally & Bernstein, 1994). The first model is a model of congeneric measurements. This model specifies that: (a) ratings of the same dimension actually measure a common performance dimension; (b) the loadings of the three ratings from the three different rating sources with respect to the performance dimension are of potentially different magnitudes (i.e., unconstrained); and, (c) the unique variances of the ratings are potentially unequal (i.e., unconstrained). The second model is a model of tau equivalent measurements. This model specifies that: (a) ratings of the same dimension from the three different rating scales measure a common performance dimension; (b) the loadings of the three ratings from the three different rating scales with respect to the performance dimension are of equal magnitudes; and, (c) the unique variances of the ratings are of potentially different magnitudes (i.e., unconstrained). The third model is a model of parallel tests. This model specifies that: (a) ratings of the same dimension from the three different rating scales measure a common performance dimension; (b) the loadings of the three ratings from the three different rating scales with respect to the performance dimension are of equal magnitudes; and, (c) the unique variances of the ratings are of equal magnitudes. Finally, a null model is specified for comparative purposes and in essence postulates that each rating measures a different latent variable (i.e., there is no commonality across the ratings sources).

Each of the three models has different implications for the equivalence of the ratings obtained across the three rating sources. Each of the models separates the variance in performance ratings into that attributable to the performance dimension being assessed and that which is unique to each rating. Support for the first (congeneric) model would indicate that ratings of a performance dimension from the three rating sources measure the same construct but not to the same degree and not with equal reliability. Support for the second (tau equivalent) model would indicate that the ratings tap the same construct to the same degree, but not with equal reliability. Support for the third (parallel) model would indicate that the ratings measure the same construct to the same degree and with equal reliability. Finally findings that none

of the three models provide a better level of fit to the data than does the null model would indicate the different rating sources are providing measurements of different constructs.

Here it must be noted that the models presented above only indicate the extent to which the three rating sources measure the same construct and not their equivalence in a strict classical test theory sense. That is, only if it can be demonstrated that the unique variance associated with each rating contains only random error variance (i.e., does not include any specific, systematic variance) do the models conform exactly to classical test theory definitions of congenericism, tau equivalence, and parallelism. This is unlikely to be the case in the present study in that the rating sources are likely to introduce specific systematic variance (i.e., source effects). Therefore, a modified version of the models described above is also evaluated. This modification involves allowing the error terms among ratings from the same rating source across dimensions to correlate. In essence this modification adds a third potential source of variance in the performance ratings. That is, shared variance attributable to the rating source (source effects) is included in addition to the variance attributable to the performance dimensions and the specific unique variance associated with each rating. Consequently the magnitude and degree of equivalence of parameter estimates associated with each of these sources provides an indication of the nature of the commonality across rating sources not attributable to the performance dimension being assessed.

Method

Participants

Data for this study were collected as part of the US Air Force Job Performance Measurement (JPM) project (Hedge and Teachout, 1986). The JPM project was a large scale concurrent validation effort focusing on the measurement of several criterion measures including work samples, interviews, and job performance ratings. Job performance ratings were obtained from airmen, immediate supervisors of the airmen, and peers/coworkers of the airmen. The sample in the present study consisted of 1,028 airmen from seven different Air Force job categories.

Rating Instruments

All participants were rated on job specific tasks as well eight general job performance dimensions.

The general job performance dimensions were assessed using a five-point adjectivally anchored rating scale-- developed to measure factors associated with overall job performance, across all Air Force job categories. Ratings were generated for each of the eight dimensions which included: technical knowledge; initiative and effort; knowledge of and adherence to regulations and orders; integrity; leadership; military appearance; self development; and self control (Augustin, Gillet, Guerrero, & Ey, 1989).

Results

The goodness-of-fit of the three original models (the 3 measurement equivalence models with uncorrelated unique variances) and the three modified models (the 3 measurement equivalence models with the unique variances within rating sources allowed to correlate) was evaluated using a confirmatory factor analytic application of LISREL 8 (Jöreskog and Sörbom, 1996). The observed covariance matrix among the 24 performance ratings was used to determine individual chi-square values for the each of the three models. The overall fit of each model was initially assessed by examining a number of criteria, including the χ^2 , χ^2/df ratio, NCP, GFI, NNFI, CFI, IFI, and RMSEA.

In addition, it is important to note that the six models represent a parameter-nested sequence of models which are hierarchically nested from most restricted (parallel model with uncorrelated uniquenesses) to the next level (tau equivalent model with uncorrelated uniquenesses) to the least restricted (congeneric model with within source correlated uniquenesses). Thus, it is possible to use a difference of chi square test to assess whether the difference in fit between each successive model is significant. In this analysis it is preferable to accept the most restrictive model that does not result in a significant reduction in fit over less restricted models.

Results of the model evaluations are presented in Table 1. These results indicate a significant and substantial difference in chi square between the original models and the modified models allowing within rating source correlated uniqueness. This suggests that the modified models provide a significantly better representation of the data than do the original models. Among the three modified models, the results indicate the lowest chi square value (and thus best fit) for the congeneric model. However, the difference in chi square values between the congeneric model and the tau-equivalent models is not significant

(difference of $\chi^2 = 201.82 - 178.09 = 23.73$; $df = 156 - 140 = 16$). The difference in chi square values between the tau equivalent and parallel models, however, is significant (difference of $\chi^2 = 346.42 - 201.82 = 144.61$; $df = 172 - 156 = 16$). This indicates that the parallel model represents a significantly lower level of fit to the data; thus, the tau-equivalent model provides the overall best representation of the current data.

Examination of the individual parameter estimates for the tau equivalent model with correlated uniquenesses reveals that factor loadings for the performance dimensions range from 0.35 to 0.50 with a mean value of 0.43. All of the t -values for the performance loadings are significant. Thus, all parameter loadings on the performance dimensions are significantly different than zero, yet are of relatively low magnitude. In addition, the squared multiple correlations of the ratings with the performance dimension latent variables are quite low, ranging from 0.14 to 0.39, suggesting that the performance ratings are relatively weak or unreliable indicators of the performance dimensions.

In addition, the parameters representing the commonality across dimensions and within rating source ranged from 0.12 to 0.24 with a mean value of 0.20. Further, all of these parameters are significantly different than zero.

Finally, parameter estimates associated with the unique variance components were quite high, ranging from 0.58 to 0.84 with a mean value of 0.73. All of these parameters are significantly different than zero and exceed factor loadings on all performance dimensions.

Discussion

In the present study confirmatory factor analysis was used to evaluate a series of models representing the relations among performance ratings from multiple sources across multiple performance dimensions. Our results suggest that a model in which performance rating are attributable to three distinct sources of variance provides the best representation of the data. The three sources of variance are: a) commonality attributable to a latent variable representing the performance dimension being assessed; b) commonality attributable to the rating source (self, supervisor, or peer); and, c) unique variance over and above both performance dimension and source effects. Our results further suggest that: a) the different sources are relatively equivalent (i.e, tau equivalence) in the extent to which they assess the performance

construct; b) systematic effects (i.e, biases) attributable to the specific rating source, while statistically significant, are smaller than effects attributable to the performance dimensions; and, c) parameters not associated with either the performance dimensions or rating source had the largest individual influence on ratings.

These results are encouraging with respect to the performance dimensions being assessed and with respect to source-specific systematic biases. Different rating sources are able to assess commonalities in performance to an equivalent extent and to an extent greater than that associated with source specific effects. The results are less clear with respect to unique effects. However, the findings with respect to these unique effects must be interpreted with caution. Two factors may contribute to these effects. First, these effects may reflect random error variance in the ratings and thus provide an indication of reliability (or lack thereof). These effects may also reflect systematic variance not attributable to either commonalities in the performance dimension or to the rating source. This systematic variance may reflect the unique aspects of performance observed and rated by each rating source. For example, all raters may have had the opportunity to observe some aspects of performance. They may also have had the opportunity to observe unique aspects of performance. This would be reflected in the pattern of results obtained. Of course it is most likely that the unique effects represent both systematic and random error components. A important issue for future research is the disentanglement of these effects. One suggestion for how this issue might be addressed is to use confirmatory factor analytic techniques to look at the relationship of these unique effects with other criterion measures (e.g., work samples, job knowledge tests, etc.). To the extent that these effects reflect systematic, performance-related variance, they should correlate with other criterion measures and to the extent that they represent random effects they should not. This approach might also be extended to the latent variables representing the performance dimensions as well. That is, correlations of these latent variables with other criterion measures might be examined along with similar correlation with the unique effects.

In summary, the present study represents an attempt to empirically model the relations among ratings of multiple performance dimensions by multiple rating sources. These results appear to shed some

light on the nature of the convergence or lack thereof of ratings across sources. Future research is needed to further elucidate this issue. A follow-up study has been planned in an attempt to address this issue. Additionally, this project has been submitted to be presented at annual conference for the Society for Industrial and Organizational Psychology's in April 1999.

References —

- Ashford, S. J. (1989). Self-assessment in organizations: A literature review and integrative model. In L.L. Cummings & B. M. Staw (Eds.), Research in Organizational Behavior, Vol.11, (pp.133-174). Greenwich CT: JAI Press.
- Augustin, J. W. ,Gillet, A. H., Guerrero, Jr., C., & Ey, A. W. (1989). Documentation of the job performance measurement system data base contents. (UES Report #788-038). Dayton, OH: Universal Energy Systems, Inc. Prepared under Contract F41689-86-D-0052, Air Force Human Resources Laboratory, Brooks AFB, TX.
- Buckner, D. N. (1959). The predictability of ratings as a function of interrater agreement. Journal of Applied Psychology, 43, 60-64.
- Campbell, D. J. & Lee, C. (1988). Self-appraisal in performance evaluation: Development versus evaluation. Academy of Management Review, 13, 302-314.
- Church, A. H. (1997). Managerial self-awareness in high-performing individuals in organizations. Journal of Applied Psychology, 82, 281-292.
- Facteau, J. D. (1998, April). Are Performance Appraisal Rating Obtained from Different Rating Sources Comparable? Paper presented at annual meeting of the Society for Industrial and Organizational Psychology, Dallas, TX.
- Fahr, J. I., & Dobbins, G. H. (1989a). Effects of self-esteem on leniency bias in self-reports of performance: A structural equation model analysis. Personnel Psychology, 42, 835-850.
- Fahr & Dobbins (1989b). Effects of comparative performance information on the accuracy of self-ratings and agreement between self- and supervisor ratings. Journal of Applied Psychology, 74, 606-610.
- Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. Psychological Bulletin, 101, 75-90.
- Harris, M. M. & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. Personnel Psychology, 41, 43-62.

Hedge, J. W., & Teachout, M. S. (1986). Job Performance Measurement: A Systematic program of Research and Development (AFHRL-TP-86-87). Brooks AFB, TX: Training Systems Division, Air Force Human Resources Laboratory.

Joreskog, K. G., & Sorbom, D. (1996). LISREL 8: User's reference guide. Chicago: Scientific Software International.

London, M., & Beatty, R. W. (1993). 360-degree feedback as a competitive advantage. Human Resource Management, 32, 353-372.

Mabe, P. A. III., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. Journal of Applied Psychology, 67, 280-296.

Maurer, T. J., Raju, N. S., & Collins, W. C. (In press). Peer and subordinate appraisal measurement equivalence. Journal of Applied Psychology.

Murphy, K. R., & Cleveland, J. N. (1995). Understanding performance appraisal: Social, organizational, and goal-based perspectives. Thousand Oaks, CA: SAGE Publications, Inc.

Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. Journal of Applied Psychology, 81, 557-574.

Wherry, R. J. (1983). Wherry's Theory of Rating. In F. J. Landy & J. L. Farr, The Measurement of Work Performance: Methods, Theory, and Application. New York: Academic Press.

Table 1. Goodness of Fit Statistics for parallel, tau equivalent, congeneric, and baseline/null models for Air Force-wide Ratings

Model	χ^2	df	χ^2/df^b	$\Delta\chi^2$	Δdf	NCP	RMSEA	GFI	NNFI	CFI	IFI
Null	10502.76	276	38.05	--	--	10227	0.19	0.33	0.0	0.0	0.0
Parallel	1026.05	256	4.01	9476.7	20	770.05	0.05	0.92	0.95	0.96	0.96
Tau Equivalent	971.39	240	4.05	54.66	16	731.39	0.05	0.92	0.95	0.96	0.96
Congeneric	952.70	224	4.25	18.69	16	728.70	0.06	0.92	0.95	0.96	0.96
Parallel*	346.42	172	2.01	606.28	52	174.42	0.03	0.97	0.98	0.99	0.99
Tau Equivalent*	201.82	156	1.29	144.61	16	45.82 ^a	0.02 ^a	0.98	0.99	0.99	0.99
Congeneric*	178.09	140	1.27	23.73	16	38.09 ^a	0.02 ^a	0.99	0.99	0.99	0.99

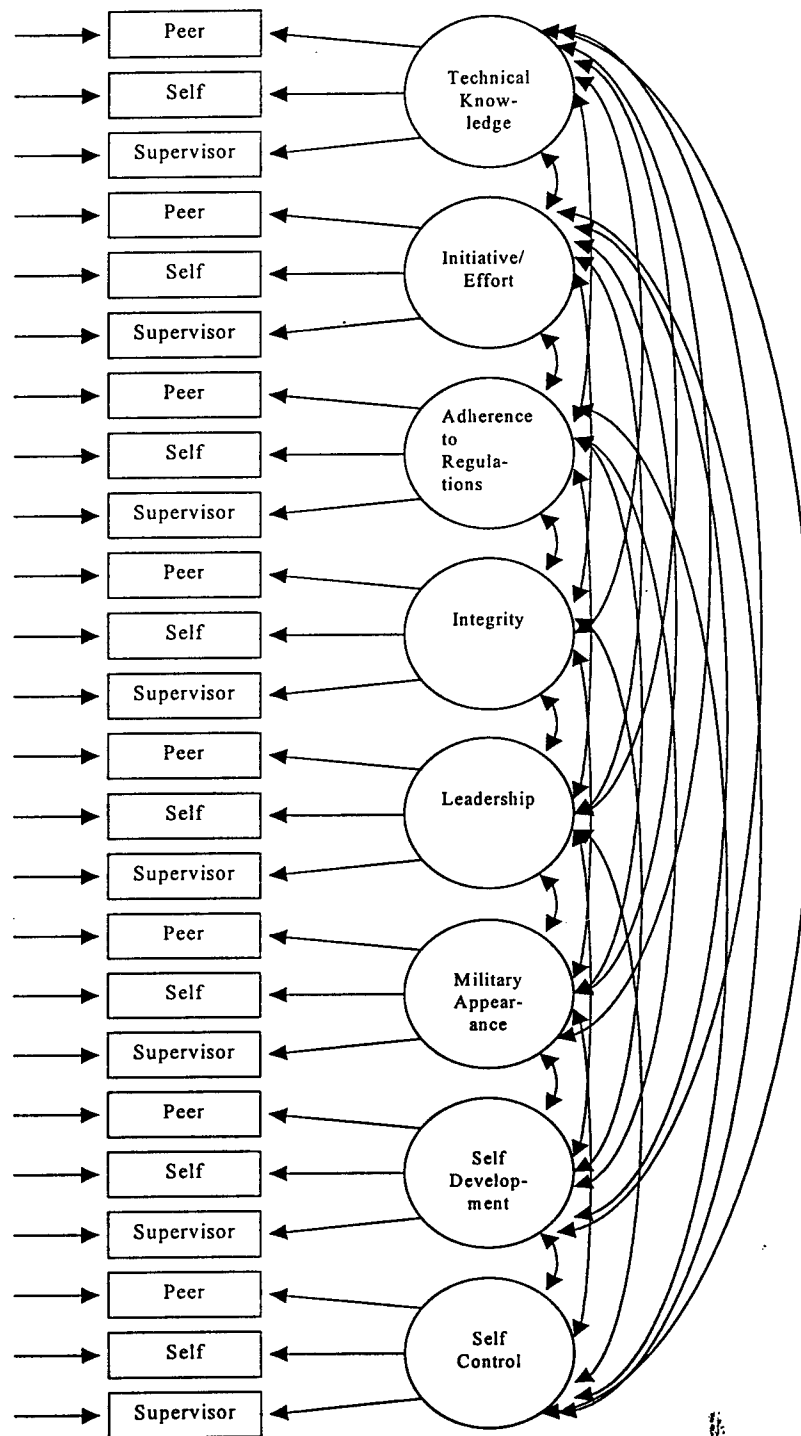
* Modified to include correlated uniqueness within rating source across dimensions.

^a Denotes inclusion within 90% confidence interval

^b The difference in χ^2 to the next more restricted model (e.g., the congeneric model is compared with the tau-equivalent model, the tau-equivalent model with the parallel, and the parallel with the null).

χ^2_{crit} for 16 df @ $\alpha = .05$ is 26.296

Figure 1. Performance Rating Factor Model



THE EFFECT OF SPATIAL SEPARTATION AND ONSET
ASYNCHRONY ON THE DETECTABILITY AND INTELLIGIBILITY OF
A CRITICAL CALL SIGN PHRASE IN A MULTI-TAKER
ENVIRONMENT

Brian D. Simpson
Graduate Research Assistant
Department of Psychology

Wright State University
3640 Colonel Glenn Highway
Dayton, OH 45435

Final Report for:
Graduate Strdent Research Program
AFRL/Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

And

Armstrong Laboratory

September 1998

THE EFFECT OF SPATIAL SEPARATION AND ONSET ASYNCHRONY ON THE DETECTABILITY AND INTELLIGIBILITY OF A CRITICAL CALL SIGN PHRASE IN A MULTI-TALKER ENVIRONMENT

Brian D. Simpson
Graduate Research Assistant
Department of Psychology
Wright State University

Abstract

Two experiments were conducted to examine the detectability and intelligibility of a target call sign and subsequent target speech phrase in the presence of competing messages, were examined. In both experiments, the speech phrases were presented virtually over headphones to a listener whose head position was tracked. In the first experiment, multiple simultaneous talkers were located at virtual positions along the median plane of a listener. The number of competing speech phrases was varied from 2-8 and the spatial separation of the simultaneous phrases was manipulated. In the second experiment, from 2-8 talkers were located at virtual positions along the horizontal plane, and the relative onset times of the individual speech phrases was varied. On half of the trials, the speech stimuli all appeared to originate from the same location, and on the other half of the trials the stimuli appeared to be spatially separated on the horizontal plane. Preliminary results indicate that both detection and intelligibility decrease as the number of speech phrases increases. Furthermore, both detection and intelligibility were improved when the speech phrases were spatially separated over the case in which all speech stimuli appeared to originate from one location. In addition, listener self-reports indicate that longer stimulus onset asynchronies improve one's ability to detect a critical call sign. The results from these experiments have relevance to the design of auditory displays for communications, warning systems, and virtual environments, the development of hearing aids for patients with sensorineural hearing impairments, and to the study of the mechanisms by which the auditory system segregates information associated with a target sound source from information produced by competing sound sources.

THE EFFECT OF SPATIAL SEPARATION AND ONSET ASYNCHRONY ON THE DETECTABILITY AND INTELLIGIBILITY OF A CRITICAL CALL SIGN PHRASE IN A MULTI-TALKER ENVIRONMENT

Brian D. Simpson

INTRODUCTION

Recent advances in technology have led to increases in the rate of information presented to human operators in complex operational settings (e.g., aircraft cockpits, command and control centers, etc.). Most current interface systems emphasize the presentation of information via visual displays, thus placing great demands on the visual information processing capacity of the operator. In an effort to lessen these demands, a great deal of research has focused on the viability of interfaces that access other sensory modalities (e.g., auditory, haptic/tactile). Spatialized auditory displays, in particular, have shown great promise (McKinley, Ericson, and D'Angelo, 1994). These displays take advantage of the fact that the auditory system is capable of monitoring the entire surrounding acoustic environment simultaneously, such that an operator may focus his/her attention on a relevant auditory stimulus presented from any direction.

Three-dimensional auditory displays may be used in many important applications. For instance, auditory stimuli generated from positions that are spatially coincident with visual targets have been shown to reduce target acquisition times by as much as a factor of 2-8 (Perrot, Cisneros, McKinley, and D'Angelo, 1995; Gilkey, Simpson, Isabelle, Anderson, and Good, 1997; McKinley and Bolia, 1998), and to reduce cognitive workload (Nelson, Hettinger, Cunningham, Brickman, Haas, and McKinley, 1997). Moreover, the auditory cue appears to be most beneficial when the visual search task is most complex (e.g., the visual target is a conjunction of two features, the target is surrounded by multiple distracters, or the field of view is narrowed as is the case in helmet-mounted displays). In addition, the position of an auditory stimulus may also be informative. For example, in a cockpit situation, a pilot may be able to keep track of his/her wingman's position by monitoring the position of his/her voice, potentially increasing a pilot's overall situation awareness.

The focus of this research has been to determine the potential benefits and limitations of 3-dimensional auditory displays for improving speech intelligibility. Specifically, the main goal has been to determine the maximum

number of talkers a listener can simultaneously monitor, and the specific stimulus characteristics that maximize this number.

Before the experiments are described, the relevant background literature must be discussed, including a brief review of the mechanisms underlying human sound localization and their contributions to 3-dimensional auditory displays.

Sound Localization. As an acoustic waveform travels from a sound source to a listener's tympanic membrane, temporal, intensive, and spectral modifications are imposed on the waveform by a listener's pinnae, head, and torso, which provide the physical cues that are the basis for sound localization. These physical modifications, collectively known as the head-related transfer function (HRTF), vary as a function of the position of a sound source relative to a listener's head. A listener can utilize temporal and intensive modifications in a stimulus by comparing the waveforms at the two ears (interaural difference cues), and can make use of spectral modifications by analyzing the characteristics of a stimulus at only one ear (monaural spectral cues).

The relative usefulness of these temporal, intensive, and spectral cues varies with the spectral content of the auditory stimulus. For low-frequency stimuli (below about 1000 Hz), ongoing interaural temporal differences (ITDs) provide a robust cue for azimuthal localization. For stimuli above approximately 2000 Hz, the wavelengths are small relative to the size of a listener's head. Thus, the head casts an acoustic shadow such that the stimulus at the far is attenuated relative to the stimulus at the near ear. This interaural level difference (ILD) provides another robust cue for azimuthal localization that operates over a frequency region complementary to the cues provided by ITDs. These two binaural cues are the basis of the Duplex Theory (Rayleigh, 1907). However, these interaural cues cannot explain a listener's ability to localize in the vertical dimension. Spectral cues, introduced by the pinna, appear to mediate performance in this dimension. For very high frequency stimuli (i.e., above 5000 Hz), the wavelength of the auditory stimulus is small relative to the cavities and folds of the pinna (outer ear) of a listener. Thus, the pinnae impose modifications on the spectrum of the stimulus that vary with the direction of the stimulus. The resulting peaks and notches provide the cues for vertical localization. Finally, overall low-pass filtering of an auditory stimulus presented from behind a listener provides the cues that help to disambiguate stimuli originating from a location in the frontal hemifield from those originating from the rear hemifield. This is particularly important

when the two locations in question fall on an iso-interaural difference cue contour, the so-called “cone of confusion” (Woodworth, 1938).

Localization Cues and Speech Intelligibility. While most current communications systems in complex environments employ only monaural presentation of auditory stimuli (i.e., presentation over a single earphone), the benefits of binaural hearing to communication effectiveness has long been known (Licklider, 1948; Hirsh, 1950; Cherry, 1953; Cherry and Taylor, 1954; Spieth, Curtis, and Webster, 1954; Webster and Thompson, 1957). For instance, researchers have found improvements for speech intelligibility in noise and for speech intelligibility in the presence of competing speech when the multiple speech phrases arrive at different ears (Hirsh, 1950; Cherry, 1953). This has been attributed to the fact that these speech phrases arrive at a listener’s ears with different interaural cues, thus providing a means of segregating these multiple speech phrases. In relation to the free-field, sounds with common interaural cues are assumed to come from a common spatial location; sounds coming from a single location most likely arise from a single sound source. Actual spatial separation in the free-field appears to provide even greater improvement in speech intelligibility. For example, Spieth, Curtis, and Webster (1954) found that speech intelligibility improved as the spatial separation between two simultaneous speech phrases increased. The fact that sounds in the real world occur external to the listener, and occupy actual locations in space appears to help a listener segregate multiple simultaneous sounds. This experience can be simulated by filtering these speech phrases through HRTFs, thus generating stimuli that, when presented over headphones, appear to exist at locations external to a listener.

Recent work at the Air Force Research Laboratory has examined the effect of the number of simultaneous talkers and spatial separation (in both real and virtual environments) on a listener’s ability to detect and identify a target call sign phrase (Nelson et al., 1998). In these experiments, from 2 to 8 simultaneous talkers were separated in azimuth on the horizontal plane of a listener. Five different spatial configurations were examined: four spatialized configurations plus one condition in which all talkers came from the same spatial location. They found that as the number of simultaneous talkers increased (from 2 to 8), detection and identification decreased. Moreover, they found that in all “number-of-talker” conditions, spatializing the speech phrases improved performance over the case in which all speech phrases originated from the same spatial location. However, when three talkers were present,

intelligibility fell below 50%, and when more talkers were present, both detection and intelligibility degraded appreciably.

Although the contribution of differences in interaural cues to speech intelligibility has been examined by spatially separating multiple talkers in azimuth, the benefits of spatial separation along the median plane, where interaural differences are identical and equal to zero, has not been explored. Recall from the section on sound localization that localization in the vertical dimension is governed largely by monaural spectral cues. Therefore, different spectral characteristics of the multiple simultaneous sound sources resulting from the filtering by the HRTFs may in fact also contribute to the segregation of concurrent speech phrases, thus leading to an improvement in intelligibility. This may lead to an improvement in performance over that found by Nelson et al. (1998). Although absolute localization is worse in the vertical dimension than in azimuth (Middlebrooks and Green, 1991; Makous and Middlebrooks, 1990; Oldfield and Parker, 1984a), it was believed that spectral cues would be sufficiently robust to provide adequate segregation and subsequent gains in speech intelligibility. Experiment 1 was a first attempt to examine the benefit of spatial separation in elevation for the case in which only spectral cues are available.

EXPERIMENT 1

Method

Subjects

Five listeners (3 males and 2 females), ranging in age from 18 to 45, were paid for their participation in the experiment. All subjects had clinically normal hearing, as determined by standard audiometric measurements. Subjects were chosen from the Veridian In-House Listening Panel. This panel of human subjects is maintained by Veridian for the Air Force Research Laboratory.

Apparatus

The experiment took place in the anechoic chamber of the Auditory Localization Facility at the Air Force Research Laboratory. This facility is housed in Building 441 in Area B of Wright-Patterson Air Force Base. The walls of the chamber are lined with 1.1-meter thick fiberglass wedges in order to reduce unwanted echoes. Within the chamber exists a 4.3-meter geodesic sphere. The surface of the sphere contains 277 loudspeakers which surround the subject in azimuth and elevation. Subjects were seated on a padded bench in the center of the sphere.

Stimuli

The stimuli utilized in this experiment were speech phrases taken from the Coordinate Response Measure (CRM) sentences (Moore, 1981). Eight talkers (four male, four female), read these sentences into a microphone, where they were digitally recorded at a 40-kHz sampling rate with 16-bit resolution using the Tucker-Davis Technologies (TDT) DD1 analog-to-digital converters. The RMS level of the individual speech phrases was adjusted such that they would all phrases would have the same average power. None of the talkers participated as subjects in the experiment.

Each sentence consisted of a call sign a color and a single-digit number. There were 8 possible call signs (Arrow, Baron, Charlie, Eagle, Hopper, Laker, Ringo, and Tiger), 4 possible colors (Red, Blue, Green, and White), and 8 possible numbers (1-8), for a total of 256 possible sentences. For example, one possible sentence was “Ready Baron go to red one now.” Each of the 256 sentences was read by each of the 8 different talkers for a total of 2048 sentences.

The stimuli were presented virtually by convolving the speech phrases with head-related transfer functions (HRTFs) in order to generate spatialized auditory images. These HRTFs were recorded via microphones placed in the ear canals of an anthropomorphically correct mannequin (Knowles Electronics Manikin for Acoustic Research, or KEMAR). HRTFs were recorded from each of the loudspeaker locations on the sphere, thus constraining the locations of the speech stimuli along the median plane to the actual positions of the loudspeakers. However, because the listeners' heads were tracked, linear interpolation techniques were employed to generate stimuli at intermediate positions between the actual loudspeaker locations. These HRTFs were downloaded into two 3-D Auditory Display Generators (McKinley, 1988), each of which was capable of presenting 4 independent channels simultaneously, for a total of up to 8 simultaneous sound sources at 8 virtual spatial locations. The stimuli were presented via open dynamic supra-aural headphones (Sennheiser model HD520 II).

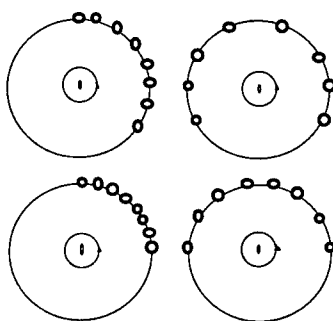
The spatial position of the target call sign (see Table 1) and sex of the target talker (male vs female) were varied, as was the number of talkers (2-8) and the spatial configuration of the multiple talkers. One of five spatialization conditions was employed, utilizing various combinations of the spatial locations shown in Table 1. These

configurations included the frontal hemifield, the upper front hemifield, the entire upper hemifield, and the entire median plane (above -45° elevation). Locations within these constraints were chosen such that the average difference in source-midline distance was maximum for the configuration (see Figure 1).

TABLE 1. Virtual Locations Along the Median Plane (degrees)

Frontal Hemifield	Rear Hemifield
0 , -44.71	180 , -44.71
0 , -31.72	180 , -31.72
0 , -10.57	180 , -10.57
0 , 0	180 , 0
0 , +10.57	180 , +10.57
0 , +31.72	180 , +31.72
0 , +44.71	180 , +44.71
0 , +57.25	180 , +57.25
0 , +69.09	180 , +69.09
0 , +82.41	180 , +82.41
0 , +90	180 , +90

Figure 1



Procedure

On each trial, the listener's task was to listen for his/her call sign in the presence of multiple competing call sign phrases. On half of the trials the target call sign was present, and on half of the trials the target call sign was absent. The subjects' first task is to press a key labeled "yes" on a standard computer keyboard if target was present, and "no" if the target was absent (the detection task). Some of the keys on this keyboard were overlaid with labels that correspond to the subjects' task. When the target was present, the subject was to perform the task instructed by the speech phrase containing the target call sign. For instance, if the target speech phrase was "Ready Baron go to blue two now" the subject was to press the blue colored key labeled "2" (the identification task). The computer waited for a response before presenting another stimulus. If the subject responded "no" to the detection task, the computer simply proceeded to the next trial in the block. In all cases, the target call sign for the listener was "Baron."

Results

Preliminary results from Experiment 1 suggest that spatial separation on the median plane does indeed lead to better performance on the detection and identification tasks. That is, spectral cues alone appear to be sufficient for the segregation of simultaneous speech phrases. Moreover, there appears to be no significant difference in performance between any of the spatialization conditions, but all spatialization conditions appear to lead to significantly better performance than the case in which all phrases came from a single virtual location. Furthermore, as expected, increasing the number of talkers led to a decrease in overall detection and identification performance. These results are consistent with the results of Nelson et al. (1998) for simultaneous speech phrases presented on the horizontal plane. [Note: Final analyses of these data could not be obtained before the end of the Summer Research Program because data collection was not completed.]

Other Cues Influencing Speech Intelligibility. In addition to spatial separation, Yost (1997) has suggested that several other physical attributes of sound also contribute to one's ability to understand speech in "cocktail party" environments. These include spectral separation, spectral profile, harmonicity, temporal separation, temporal onsets and offsets, and temporal modulation. Ericson, Bolia, Nelson, and McKinley (1998) have shown that characteristics of individual talkers related to these attributes may cause some talkers to be more or less detectable relative to the other competing talkers. These characteristics may provide a set of organizing principles whereby sounds that share common characteristics are streamed together (Bregman, 1990). Such streaming may be enhanced by an increase in the number or length of temporal gaps across the multiple talkers. One issue that has been ignored in this series of studies is the truly dynamic nature of speech. Specifically, differences in onset and offset times may provide one set of cues that can enable a listener to segregate multiple simultaneous speech phrases (Tuller and Lackner, 1976; Bregman, 1990). It is assumed that it becomes easier to detect and identify key phrases when portions of the target words occur during relatively low-amplitude periods in the competing phrases. "Listening in these gaps" may improve speech intelligibility (Ciocca and Darwin, 1993). Relative delays in onset times of the multiple speech phrases will provide these enhanced gaps (Summerfield and Assman, 1991; Culling and Summerfield, 1995a; Culling and Summerfield, 1995b). Therefore, in light of the results from these studies, it seems relevant to examine the benefit of varying the relative onset times of the multiple speech phrases in order to determine if more than 3 talkers may be simultaneously monitored successfully.

EXPERIMENT 2

Method

Subjects

In Experiment 2, four listeners (2 males and 2 females), ranging in age from 18 to 45, were paid for their participation in the experiment. The two females also participated in Experiment 1, but the two males were new to the study, although both had previously participated in auditory experiments. All subjects had clinically normal hearing, as determined by standard audiometric measurements. Subjects were chosen from the Veridian In-House Listening Panel. This panel of human subjects is maintained by Veridian for the Air Force Research Laboratory.

Apparatus

The experiment took place in the same facility described in Experiment 1.

Stimuli

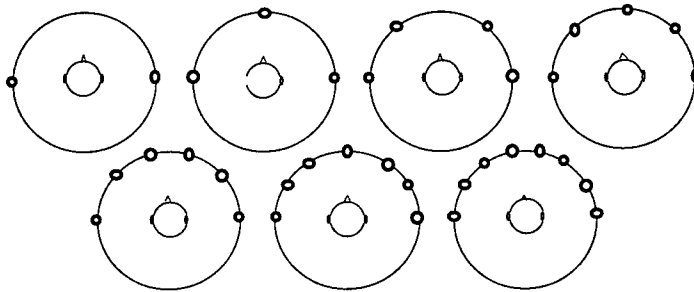
The same speech corpus used in Experiment 1 was used in Experiment 2, with one exception. The speech phrases, as they were originally recorded, began with the word "Ready." However, because the introduced onset delay between speech phrases was targeted at the critical call sign, "Ready" was removed from each of the phrases using the Cool Edit 96 digital audio editing software package. For example, the edited speech phrase could be "Baron, go to red one now." This process was time consuming, but provided another set of 2048 sentences to be added to the existing speech corpus. [Note: It is expected that this speech corpus will be made available via the World Wide Web for use in auditory experiments in other laboratories]. As in Experiment 1, the speech stimuli were presented virtually.

The spatial configuration of the stimuli on a given trial was determined by the number of speech phrases to be presented on that trial. All speech phrases originated from locations on the horizontal plane, and the locations were chosen in order to maintain maximum spacing between adjacent sound sources in virtual auditory space, given the constraints of the loudspeaker locations on the geodesic sphere. The specific locations of the stimuli in each condition are indicated in Table 2 and a schematic of these locations is shown in Figure 2.

TABLE 2. Virtual Loudspeaker Configurations for Experiment 2

# TALKERS	LOUDSPEAKER LOCATIONS (deg azimuth)
2	-90, +90
3	-90, 0, +90
4	-90, -45.3, +45.3, +90
5	-90, -45.3, 0, +45.3, +90
6	-90, -58.3, -21.9, +20.9, +58.3, +90
7	-90, -58.3, -33.0, 0, +32.0, +58.3, +90
8	-90, -58.3, -33.0, -7.6, +7.6, +32.0, +58.3, +90

Figure 2



In addition to these 7 spatialization conditions, one condition (the control) was examined in which all talkers presented on a given trial appeared to come from a single virtual location. The spatial configurations examined were identical to a subset of configurations examined by Nelson, et al. (1998) In all cases, the call sign for the listener was "Baron."

In each condition the onset time of the individual speech phrases was delayed relative to the onset times of the other speech phrases. The duration between successive onsets was fixed for a given trial. For example, on a given trial in the 3-talker condition, the difference in onset times between the first and second speech phrases was the same as the difference in onset times between the second and third speech phrases. Onset delays of 0, 40, 80, and 160 ms were examined. These delays were chosen such that, in the 8 talker condition, all speech phrases would at least minimally overlap. The target speech phrase could occur in any temporal position (i.e., 1st, 2nd, 3rd, 4th, etc.). This was varied in order to determine if the temporal position of the speech phrase contributed to its detectability and intelligibility. Furthermore, the number of speech phrases varied from 2-8, and the spatial location of the target speech phrase was randomly chosen from trial to trial. In no case were two phrases spoken by the same talker on a given trial. In addition, the target call sign only appeared in one speech phrase on any given trial

Procedure

The procedure in Experiment 2 was identical to the procedure used in Experiment 1, but the subject was only required to respond by carrying out the instruction presented in the target call sign phrase. The yes/no detection response was eliminated based on results from a brief experiment conducted concurrently with the two main experiments. This brief experiment was designed to evaluate the reliability of our response techniques with and without the yes/no portion of the task. The results showed no difference in performance between the two response techniques, and therefore the most efficient technique (that of only one response per trial) was adopted. This was additionally useful because, due to hardware constraints, the duration of Experiment 2 was significantly longer than Experiment 1.

Results

Preliminary results from Experiment 2 once again suggest that spatial separation on the horizontal plane does lead to better performance on the detection and identification tasks. That is, all spatialization conditions appear to lead to significantly better performance than the case in which all phrases came from a single virtual location. Furthermore, as expected, increasing the number of talkers led to a decrease in overall detection and identification performance. These results are consistent with the results of Nelson et al. (1998) for simultaneous speech phrases presented on the horizontal plane, and with preliminary results from Experiment 1. Finally, based on listener self-reports, increases in onset asynchrony led to better detection. Finally, it appeared that the temporal position of the target talker contributed to performance. When the target call sign was presented first, it appeared that detection increased significantly over the case when the target call sign occurred at any other temporal position. [Note: Final analyses of these data could not be obtained before the end of the Summer Research Program because data collection was not completed.]

Results from both of these experiments will be presented at the 137th Meeting of Acoustical Society of America and the 2nd Convention of the European Acoustics Association integrating the 25th German Acoustics DAGA Conference, Berlin, Germany, 14-19 March 1999

Discussion

In all cases, it appears as if detection and identification performance gets worse as the number of simultaneous talkers increases. In addition, although the final analyses have not been completed, it appears that virtual spatial separation of multiple talkers improves detection and identification of critical call sign phrases. Moreover, this improvement seems to hold for spatial separation in both azimuth and elevation. What remains to be seen is whether configuring the multiple talkers in virtual space such that they vary in both azimuth *and* elevation improves performance over spatial separation in either dimension alone. Future experiments will address these issues. In addition, delaying the relative onset times of the multiple talkers such that they are presented asynchronously appears to improve overall performance. Detection of the critical call sign may become easier as a result of the temporal gaps that occur, and streaming (Bregman, 1990) may enable a listener to determine the characteristics of the talker once the critical call sign has been detected. A listener may then more easily "follow" the subsequent critical speech phrase as it continues in the presence of multiple talkers.

Future experiments will include combining spatial separation in azimuth with spatial separation in elevation in order to determine the improvement in performance over either condition alone. Additionally, various values of onset asynchrony may be added to these conditions, and randomizing the relative onset asynchronies (as would be the case in "real-world" environments) may be examined. Improvements in communication effectiveness may also be seen when distance is added as a cue to the auditory stimuli. That is, a listener may be able to segregate the multiple talkers based on the differential cues produced by stimuli at different distances. This is particularly true if the virtual stimuli appear to be within 1 meter of the listener, where interaural cues change in ways quite different from the way they change at distances beyond 1 meter.

In addition, it may be the case that the actual position of a stimulus is informative. Another series of experiments may include examining a listener's ability to focus auditory attention to a particular location in space. A single location may initially be cued, either auditorily or visually, but later experiments may include the cueing of several locations, only one of which might contain the target stimulus. Such "functional positioning" studies may have great relevance for communications systems in which the individual talkers change regularly, but their functional relevance to the operators of complex systems does not. That is, although the voice characteristics of the target talker may change, the meaning of the direction from which the talker is heard does not.

References

- Bolia, R. S., Ericson, M. A., Nelson, W. T., McKinley, R. L., & Simpson, B. D. "A cocktail party effect in the median plane?"
- Bregman, A. S. (1990). Auditory Scene Analysis. Cambridge: MIT Press.
- Cherry, E. C. (1953). "Some experiments on the recognition of speech, with one and two ears," J. Acoust. Soc. Am. 25, 975-979.
- Cherry, E. C. and Taylor, W. K. (1954). "Some further experiments upon the recognition of speech, with one and with two ears," J. Acoust. Soc. Am. 26, 554-559.
- Ciocca, V. and Darwin, C. J. (1993). "Effects of onset asynchrony on pitch perception: Adaptation or grouping?" J. Acoust. Soc. Am. 93, 2870-2878.
- Culling, J. F. and Summerfield, Q. (1995). "Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay," J. Acoust. Soc. Am. 98, 785-797.
- Culling, J. F. and Summerfield, Q. (1995). "The role of frequency modulation in the perceptual segregation of concurrent vowels," J. Acoust. Soc. Am. 98, 837-846.
- Gilkey, R. H., Simpson, B. D., Isabelle, S. K., Anderson, T. R., and Good, M. D. (1997). "Design considerations for 3-D auditory displays in cockpits." AGARD Conference Proceedings 596: Audio Effectiveness in Aviation. Quebec: Canada Communications Group, 2-1 – 2-10.
- Hebrank, J. and Wright, D. (1974a). "Are two ears necessary for localization of sound sources on the median plane?" J. Acoust. Soc. Am. 56, 935-938
- Hebrank, J. and Wright, D. (1974b). "Spectral cues used in the localization of sound sources on the median plane," J. Acoust. Soc. Am. 56, 1829-1834.
- Hirsh, I. J. (1950). "The relation between localization and intelligibility," J. Acoust. Soc. Am., 22, 196-200.
- Licklider, J. C. R. (1948). "The influence of interaural phase relations upon the masking of speech by white noise," J. Acoust. Soc. Am. 20, 150-159.
- Makous, J. C., and Middlebrooks, J. C. (1990). "Two-dimensional sound localization by human listeners," J. Acoust. Soc. Am. 87, 2188-2200.

- McKinley, R. L. (1988). "Concept and design of an auditory localization cue synthesizer," Master's Thesis, AFIT/GE/ENG/88D-29, 1988.
- McKinley, R. L., and Bolia, R. S. (1998). "Complex field visual target acquisition with and without 3-D audio cueing.
- McKinley, R. A., Ericson, M. A., and D'Angelo, W. R. (1994). "3-dimensional auditory displays: Development, applications, and performance," *Aviation, Space, and Environmental Medicine*, 65(5, Suppl.), A31-8.
- McKinley, R. L., Ericson, M. A., Nelson, W. T., Bolia, R. S., & Simpson, B. D. "A series of studies exploring the spatial separation of multiple talkers using 3-D audio technology."
- Middlebrooks, J. C. and Green, D. M. (1991). "Sound Localization by Human Listeners," *Annu. Rev. Psychol.* 42, 135-159.
- Moore, T. J. (1981). Voice Communication Jamming Research, Advisory Group for Aerospace Research and Development (AGARD), conference proceedings No. 311, *Aural Communication in Aviation*, 2:1-2:6.
- Nelson, W. T., Bolia, R. S., Ericson, M. A., and McKinley, R. L. (1998). Monitoring the simultaneous presentation of multiple spatialized speech signals in the free field. In P. K. Kuhl and L. A. Crum (Eds.) Proceedings of the 16th International Congress on Acoustics and the 135th Meeting of the Acoustical Society of America, 4, 2341-2342.
- Oldfield, S. R., and Parker, S. P. A. (1984a). "Acuity of sound localization: A topography of auditory space. I. Normal hearing conditions," *Perception* 13, 581-600.
- Perrott, D. R., Cisneros, J., McKinley, R. L., and D'Angelo, W. R. (1996). "Aurally aided visual search under free field listening conditions," *Hum. Fact.* 38, 702-715.
- Rayleigh, Lord (J. W. Strutt, 3rd Baron of Rayleigh) (1907). "On our perception of sound direction," *Philos. Mag.* 13, 214-232 (NOT VIEWED).
- Saberi, K., Dostal, L., Sadralodabai, T., Bull, V., and Perrott, D. R. (1991). "Free field release from masking," *J. Acoust. Soc. Am.* 90, 1355-1370.
- Simpson, B. D., Bolia, R. S., Ericson, M. A., & McKinley, R. L. "The effect of sentence onset asynchrony on call sign detection and message intelligibility in a simulated cocktail party."
- Spieth, W., Curtis, J. F., and Webster, J. C. (1954). "Responding to one of two simultaneous messages," *J. Acoust. Soc. Am.* 26, 391-396.

- Summerfield, Q. and Assman, P. F. (1991). Perception of concurrent vowels: Effects of harmonic misalignment and pitch-period asynchrony," J. Acoust. Soc. Am. 89, 1364-1377.
- Tuller, B. and Lackner, J. R. (1976). "Primary auditory stream segregation of repeated word sequences," Percept. Mot. Skills 42, 1071-1074.
- Webster, J. C. and Thompson, P. O. (1954), "Responding to both of two overlapping messages," J. Acoust. Soc. Am. 26, 396-401.
- Woodworth, R.S. (1938). Experimental Psychology, Henry Holt and Company, New York.
- and McKinley (1997).
- Yost, W. A. (1997). "The cocktail party problem: Forty years later." In R. H. Gilkey and T. R. Anderson (Eds.) Binaural and Spatial Hearing in Real and Virtual Environments. New Jersey: Erlbaum (pp. 329-347).

THE RATE OF SKILL ACQUISITION FOR MALES
AND FEMALES ON SPACE FORTRESS

Julie A. Stiles-Shipley
Graduate Student
Department of Psychology

Bowling Green State University
Department of Psychology
Bowling Green, OH 43403-0001

Final Report for:
Graduate Student Research Program
Armstrong Research Site

Sponsored by:
Air Force Office of Scientific Research
Brooks Air Force Base, TX

and

Armstrong Research Site

August 1998

THE RATE OF SKILL ACQUISITION FOR MALES
AND FEMALES ON SPACE FORTRESS

Julie A. Stiles-Shipley
Graduate Student
Department of Psychology
Bowling Green State University

Abstract

The rate of skill acquisition for male and female participants will be examined using the complex computer game known as Space Fortress. Participants will work in dyads to learn how to play Space Fortress. It is expected that the rate of skill acquisition will be affected by previous experiences with video games. Participants who have more experience with video games should have a faster rate of learning than participants who have less experience with video games. Slopes for individual learning curves will be compared for male and female participants to determine the rate of skill acquisition.

THE RATE OF SKILL ACQUISITION FOR MALES
AND FEMALES ON SPACE FORTRESS

Julie A. Stiles-Shipley

The objective of the following study is to examine differences in skill acquisition for males and females on a complex computer-based task. The task is a computer game called Space Fortress, a complex psychomotor task that was specifically designed to test learning strategies. Research has shown that the components of Space Fortress are similar to other complex real-world tasks such as flying an aircraft, and that the skills used to play Space Fortress are transferable to these tasks. High fidelity aircraft, tank, and naval simulators have long been used to increase training efficiency and reduce costs for various types of military training; the possibility that simpler systems, like Space Fortress, may enhance training has fostered much recent research.

Although the best methods for computer-based training are still being investigated, observational learning has been shown to be an effective method for increasing training efficiency. Unfortunately, not everyone benefits equally from observation, and organizations that are interested in utilizing observational learning as a training option are concerned about its effectiveness. The goal of this project is to compare the rate of skill acquisition for males and females using a complex computer-based task. The introduction will first present the benefits of using computer games to train other skills and abilities, followed by a description of the research and development of Space Fortress. Finally, a review of observational learning will be presented along with factors that may affect its benefits.

Computers can be used to train specific tasks and procedures as well as general skills and strategies. For example, Computer Aided Instruction (CAI) and Intelligent Tutoring Systems (ITS) allow people to learn tasks though individualized, self-paced instruction (Shute & Psotka, 1996). Computers may also serve as a medium for training secondary skills that are not part of a specific task. For example, computer games may allow people to improve cognitive strategies, motor control, and spatial cognition.

Because videogames have become a popular pass-time for many children, adolescents, and adults, it is important to examine the benefits that videogames may have for improving mental skills. Videogames have been shown to increase cognitive abilities such as spatial visualization (Dorval & Pepin, 1986), spatial reasoning (Okagaki & Frensch, 1994; Subrahmanyam & Greenfield, 1994), and visual attention (Greenfield, DeWinstanley, Kilpatrick, & Kaye, 1996). In particular, Dorval and Pepin (1986) investigated whether spatial visualization test scores could be improved by playing video games. Undergraduate students, who had no previous video game experience, were trained to play the video game Zaxxon. Dorval and Pepin (1986) found that both males and females improved on measures of the Spatial Relations Test of the Differential Aptitude Test after learning to play Zaxxon, and they concluded that Zaxxon was effective for improving spatial visualization test scores. If video games can be used to increase test scores for psychological constructs, such as spatial visualization, then they may also be effective for equating individual differences in spatial skill performance to a certain level.

Although the benefits of videogames for improving spatial abilities have not been thoroughly substantiated, the relationship between videogames and other skills is strong enough to prompt organizations to incorporate videogames into training. For example, the military has adopted a variety of videogames and computer tasks into their training programs. The U.S. Army adapted a version of the Atari game "Battlezone" for training soldiers to maneuver tanks and the game "Star Wars" to improve shooting skills in tank gunners (Gagnon, 1985). The U.S. Navy developed their own videogame, called NAVTAG, to improve tactile training for junior officers (Jones, 1984), whereas the British Navy designed their own computer-based antisubmarine training game for officers (Kiddoo, 1982). "During the Gulf War, Commodore Tom Corcoran called the complex process of monitoring electronic screens to distinguish friends from foes in the air an enormous video game with life or death consequences" (Greenfield et al., 1996, p. 189). Finally, the military has acknowledged that the skills required to fly an aircraft are very similar to those required to play videogames (Greenfield et al., 1996). Playing videogames and piloting an aircraft require the controller to manage a lot of

different information at one time. Therefore, incorporating videogames into military training may provide programs with an opportunity to simultaneously teach and refine a variety of skills, which can reduce costs and increase efficiency without life and death risks.

In the early eighties, the Department of Defense organized a research team to develop a task that could be used to systematically investigate different learning strategies in a variety of situations. The result of this project was a video game called Space Fortress, designed to allow researchers to use one task to test a variety of strategies and still make comparisons across experiments. The completed version of Space Fortress resembled a two-dimensional overhead view of a space flight game. Space Fortress requires simultaneous coordination of perceptual and motor skills, conceptual and procedural knowledge, and high level strategies (Gopher, Weil, & Bareket, 1994). Furthermore, Space Fortress has characteristics similar to those present in aircraft piloting, such as attention allocation, decision making, prioritization, resource management, continuous motor control (e.g. moving the joystick) and discrete motor responses (e.g. pushing the bonus buttons) (Gopher, Weil, & Siegel, 1989). Furthermore,

"Space Fortress represents some of the information processing demands that are present in aviation such as, short and long-term memory loading, (e.g. remembering letters that signify foe mines) and high workload demands (e.g. simultaneous controlling the ship, identifying mines, battling the fortress, and collecting bonuses)" (Shebilske, Goettl, & Regian, 1996, p. 5).

The main objective of the Space Fortress game is to destroy the fortress, positioned in the center of the screen, without damaging the space ship. In order to do this two other primary goals must be met. First, it is necessary to identify and destroy mines, which risk destroying the space ship, that appear randomly on the screen as quickly as possible. Second, players must learn to manage game resources so as to maximize the total point score. To accomplish these goals, a player must understand the game's rules and properly integrate the appropriate strategies for optimal play. Specific skills include developing proper motor coordination and control of the joystick and mouse buttons, responding quickly and efficiently to new events (i.e., the appearance of a mine on the screen), and making appropriate decisions concerning the

ship's resources (i.e., when to select bonus missiles or bonus points). Players accumulate points when the desired strategies are implemented. Despite the game-like qualities of the task, Space Fortress is a sophisticated research tool that provides information about complex skill acquisition (Mané & Donchin, 1989). As a result researchers have been able to study complex skill acquisition for a psychomotor task in experimental environments.

Some Space Fortress research has focused on training skills that increase an individual's ability to fly an aircraft. Both time and cost of training of military and civilian aircraft pilots has increased dramatically (Hart & Battiste, 1992). High fidelity flight simulators provide a less expensive and more efficient alternative to traditional pilot training methods. The trend in simulators has led to the belief that the more closely the simulator resembles the real situation (high fidelity), the better the simulator will be for training. However, the relationship between physical fidelity and skill transfer has not yet been established. Space Fortress is a low fidelity game that is structurally rather than physically designed to simulate flying. The complex psychomotor components of the task are designed to train attentional strategies that are transferable to flight.

At first, there was only anecdotal evidence to suggest that videogames could be used to train pilot skills. As a result, Hart and Battiste (1992) used Space Fortress to systematically investigate the relationship between videogames and skill transfer in aircraft pilots. They tested whether workload-coping and attentional skills could be developed through Space Fortress experience and whether it would generalize to flight training. A second goal was to determine if a commercial videogame with similar properties would provide similar training benefits. One hundred nine students enrolled at the US Army Aviation Center participated in the experiment. Students either received training on Space Fortress or a popular videogame, *Apache Strike*, for a total of 10 hours. *Apache Strike* is a commercial videogame that is similar to Space Fortress in that it requires skills such as visual scanning, performing multiple tasks simultaneously, situation awareness and a relative high workload. A third group of students did not receive any

videogame practice. Point scores measured Videogame performances and flight school performance was measured by evaluation pilot ratings on student "check flights." Hart and Battiste (1992) found that students who received training on Space Fortress improved their piloting abilities more than students who were trained on the commercial videogame, and more than students in the control condition. Furthermore, there was evidence to suggest that the *Apache Strike* game was detrimental to flight performance. Students in the *Apache Strike* group performed worse than the students in the control group. Therefore, Hart and Battiste (1992) concluded that computer game training systems might be used to improve other skills if they are designed to train specific behaviors.

One skill that Space Fortress seems to target in training is the ability to efficiently manage attention. Similarly, important skills required for aircraft pilots include effective workload-coping strategies and attention management skills. Successful pilots are able to effectively manage information from competing visual and auditory stimuli while performing a variety of cognitive activities such as information processing, spatial transformations, and memorization (Hart & Battiste, 1992). After reviewing performances of cadets enrolled in an Israeli military flight school, Gopher et al. (1994) found that the most common cause of "washout" was cadets' inability to deal with the attentional demands associated with flight training. Although there is an attempt to select individuals who demonstrate high attentional capabilities, it does not prevent cadets from washing out (Gopher, 1982). Continuous selection of cadets who later failed is by no means a solution to this ongoing problem, and proves to be time consuming as well as costly. An alternative solution would be to offer specialized training that could increase attentional abilities so that more cadets could successfully complete the program.

Gopher et al. (1994) examined whether skills learned from playing Space Fortress would transfer to flight performance. They argued that attentional strategies and control could improve and transfer to new situations. Space Fortress provided a means for training these skills. Gopher et al. (1994) tested two groups of cadets from the Israeli Flight School who were trained for 10 one-hour sessions on Space Fortress using

two different training schedules. One schedule emphasized the individual components of the game, cadets practicing each component separately before they played the whole game. In the second schedule, cadets played the game in its entirety throughout training. A third group of cadets who not receive any training on Space Fortress. All groups were matched on ability and the groups did not differ in initial flight performance. Practice on Space Fortress was incorporated into the regular flight school curriculum. Participants completed one hour of training per day, with no more than 3 days between training sessions. Space Fortress performance was measured by the total game score and flight performance was determined by 6 standardized scores from test flights and 27 ratings for individual flight maneuvers as rated by the flight instructors.

Gopher et al. (1994) found that performance on Space Fortress in the component training condition was significantly better than Space Fortress performance in the integrated training condition. However, actual flight performance did not differ between these two groups, although cadets from the two game groups showed higher overall flight performance than cadets in the control group. Furthermore, the percentage of graduates from the game groups was twice that of the control group. These results are consistent with Hart and Battiste's (1992) findings that skills learned from Space Fortress transferred to flight performance in Army helicopter pilots. Transfer from the computer game to flight performance may have occurred because "the functional and demand characteristics of the two situations were similar and...attention strategies that had been developed while practicing the game were also efficient in actual flight" (Gopher et al, 1994, p. 403). An alternative explanation suggested that Space Fortress allowed players to explore different attentional strategies in an environment that was relevant to flight situations. Gopher et al. (1994) suggested that the Space Fortress experience motivated cadets to use similar models of behavior when dealing with flight demands and concluded that attentional strategies could be trained and transferred to new skills provided that the two tasks share similar attentional demands.

Gopher et al's (1994) interpretation is supported by Bandura's (1986) social cognitive theory of skill acquisition which states that previous experience can affect motivation for learning new skills.

According to Bandura, learning a new skill can be facilitated or debilitated by previous experiences as well as vicarious experiences. Both of these factors can affect an individual's self-efficacy which in turn may affect a person's success or failure on a task. Self-efficacy is defined as the belief about one's ability to perform a particular behavior (Bandura, 1986). Efficacy expectations are related to people's choices about what tasks to attempt, the amount of effort and persistence expended to accomplish a task, and ultimately the mastery of the task (Compeau & Higgins, 1995). Because Space Fortress is a complex task that requires high workload demands, cadets who learned to play the game may have higher self-efficacy as a result of their experience. This increased self-efficacy could transfer to other skills that required similar learning strategies, such as flying an aircraft. If this is the case, then a more detailed investigation needs to be conducted into the factors that may affect the learning of Space Fortress.

Perceptions of self-efficacy can influence a person's ability to successfully perform a behavior. Previous performance which resulted in successful outcomes, is more likely to increase self-efficacy, whereas prior difficulties with a task can lead to a decrease in self-efficacy. Even people who have no prior experience on a task may develop expectations about their performance based on past experiences with a similar task. Research (Bandura & Schunk, 1981; Compeau & Higgins, 1995, Locke, Frederick, Lee & Bobko, 1984) has shown that successful performance increases self-efficacy and can predict self-efficacy on subsequent tasks. Compeau and Higgins (1995) attempted to identify specific factors that might influence the effectiveness of training on computer tasks. They found that self-efficacy did influence computer performance and that those individuals who had higher self-efficacy performed better than individuals who doubted their ability. According to Bandura's (1986) Social Cognitive Theory, watching others perform a skill can influence an individual's perceptions of his/her own abilities. Observers often compare themselves to the model performing the task.

Observation has been used to increase training efficiency on cognitive (Berry, 1991; Berry & Broadbent, 1984) as well as motor tasks (Blandin, Proteau, & Alain, 1994; Carroll & Bandura, 1982; 1985; McCullagh

& Caird, 1990). Because learning to play Space Fortress increases an individual's ability to perform related tasks (i.e. piloting an aircraft), some researchers (Goettl & Connolly-Gomez, 1995, Shebilske, Regian, Arthur, & Jordan, 1992, Stiles-Shipley, Goettl, Tweney & Derek, 1998) began investigating issues of training method effectiveness for Space Fortress. Shebilske et al. (1992) examined the role of observational learning in a dyadic Active Interlocked Modeling (AIM) protocol. Trainees in the dyadic protocol took turns controlling different aspects of Space Fortress, which allowed each to learn the critical components by modeling the actions of their partner. Shebilske et al. (1992) found that participants who were trained in the AIM-dyad performed as effectively as the individual trainees. Therefore, observation in this dyadic protocol allowed dyads to reach the same skill level as individually trained participants despite less overall practice time with the task. However, it was unclear whether the results were due to the benefits of observation, sharing the workload with a partner, or the variable practice schedule. In a later study, Paulus & Shebilske (1998) demonstrated that passive observation was useful for complex skill acquisition.

Stiles-Shipley et al. (1998) investigated different training methods which effectively incorporated passive observation into the training of Space Fortress. They (1998.) examined the effects of observation, model skill level and practice schedule on the acquisition of a complex computer task. Participants were tested on their skill acquisition of Space Fortress in either a massed or spaced training schedule in which they observed either novice or experienced models. Both males and females served as observers to male models. Stiles-Shipley et al. (1998.) found that observation benefited Space Fortress performance when either novice or experienced models were observed but that spaced practice schedules were superior to massed practice schedules. However, these effects were found only for male participants. Females showed no effect of observation, model skill level, or training schedule on the acquisition of Space Fortress. The females in the observation conditions were no different from female controls, and the average total scores never reached positive values. The overwhelming poor performance of females on Space Fortress in

this experiment raises questions about what factors may have contributed to the gender differences in performance.

There is a large body of literature that addresses gender differences on such factors as cognitive abilities (Cochran & Wheatley, 1989; Gordon & Kravetz, 1991), spatial abilities (Law, Pellegrino, & Hunt, 1993; Maccoby & Jacklin, 1974), learning styles (Philbin, Meier, Huffman, & Boverie, 1995), and brain activity (Corsi-Cabrera, Ramos, Guevara, Arce, & Gutierrez, 1993), to name only a few. Whether gender differences are due to biology or socialization processes has not yet been established, since these are not easily manipulated (Jacklin, 1989). Furthermore, most reported gender differences are relatively small when measured by the size of the effect (Hyde, 1981). Therefore, it may be beneficial to investigate which factors may have affected differences between males and females' performance on Space Fortress.

One possible explanation offered by Stiles-Shipley et al. (1998) that might explain gender differences on Space Fortress performance, was the idea that males and females differed in terms of computer game experience. The only criterion related to computer game experience for participation in the experiment was that no individual could play more than 20 hours of videogames a week. Stiles-Shipley et al. (1998) found that males spent more time a week playing videogames than females. Previous research (Barnett, Vitaglione, Harper, Quackenbush, Steadmand, & Valdez, 1997; Braun & Giroux, 1989) has shown that males are more likely to play videogames than females and that males are more highly skilled at videogames than females (Brown, Hall, Holtzer, Brown, & Brown, 1997). Therefore, females may not have the previous experience on which to base their abilities to play or to learn new computer games and may just assume that they are less skilled than their male counterparts. It seems therefore that experience may play an important role in observational learning.

Robinson-Staveley and Cooper (1990) found that experience could moderate the effects of gender and social context for learning a computer game. Specifically, individuals, regardless of gender, were less likely to be affected by social context when they had high levels of experience as opposed to low levels of experience. Furthermore, high-experienced

individuals were more likely to have positive outcome expectancies regarding their performance whereas low-experienced individuals had more negative outcome expectancies. It is reasonable to assume that previous experience with videogames seems to be related to individuals' self-efficacy and outcome expectancies for novel computer tasks. Females and males who have less experience with videogames would be expected to exhibit lower ability for Space Fortress and would demonstrate a slower rate of skill acquisition for the game. Thus, the present study will examine gender differences for Space Fortress in relation to videogame experience and spatial ability.

The purpose of this study is to examine the rate of skill acquisition between males and females on Space Fortress. A significant number of females in Stiles-Shipley et al.'s (1998) experiment had difficulty learning to play Space Fortress and the majority of the scores never reached positive values. Even though acquisition curves between males and females were similar, the speed of acquisition was much faster for males than for females. Not only did males achieve higher skill levels than females with the same amount of practice, but also many of the males reached the criterion score that was used to classify experienced models. Males and females will be trained on Space Fortress to a criterion score that is indicative of skilled performance and their acquisition curves will be compared. Participants will continue to be trained until they have reached the criterion score. Information regarding videogame experience, spatial abilities, and hand dominance will be collected from each participant and used for post-hoc analyses. It is predicted that males will reach criterion more quickly than females and in fewer training sessions. Furthermore, it is expected that the rate of skill acquisition on Space Fortress will be related to the amount of videogame experience and spatial abilities but not handedness.

Experiment 1

Method

Participants

Males and females will be recruited from local temporary agencies in San Antonio, Texas to serve as trainees in the following study. They will be paid approximately five dollars per hour for their participation. Ages

will range from 18-35 and all trainees will have a minimum education requirement of a high school diploma or a Graduate Equivalency Degree (GED).

Equipment

Twenty-eight workstations will be used to administer questionnaires and to play Space Fortress. Each station will consist of a Gateway P5-100 computer with a Gateway Vivitron 15 monitor and PCI video card running SVGA Graphics, a standard 101-keyboard, CH Products FlightStick joystick, and a Logitech 3-button mouse.

Assessment Measures

Edinburgh Inventory. The Edinburgh Inventory (Oldfield, 1971) is a 10-item questionnaire that assesses hand dominance. The inventory will be used as a measure of handedness for trainees. The questionnaire calls for trainees to answer questions about which hand they use to perform a variety of everyday activities. The inventory will be presented on the computer using Micro Experimental Laboratory (MEL).

Videogame Survey. A list of the popular videogames was compiled to survey which videogames trainees play the most. The list contains games for four different types of systems: (a) home videogames, (b) computer games, (c) arcade videogames, and (d) hand-held videogames. Games were chosen for the list based on ratings by two popular computer magazines, *Computer Gaming World* (1998) and *GamePro* (1998). Trainees will rate their skill level only on the games which they have played using a 5 point Likert scale ranging from 1 (Beginner) to 5 (Advanced). Games are listed in alphabetical order and are grouped by system category.

Videogame Experience Questionnaire. A "Videogame Questionnaire" that was developed by Barnett et al. (1997) will be used to determine trainees' experience with videogames and their attitudes towards videogames. The questionnaire is divided into six sections. The six sections measure videogame experience, free-time activities, motivation for playing videogames, favorite characteristics of videogames, general attitudes toward videogames, and favorite videogames.

Cognitive Ability Measures (CAM). A test of cognitive ability measures will be used to assess trainees' cognitive and spatial abilities (Kyllonen, Woltz, Christal, Tirre, Shute, & Chaiken, 1990). Five tests

will measure verbal and spatial abilities for processing speed, working memory capacity, declarative (fact) learning, procedure (skill) learning and procedural knowledge (induction).

Space Fortress Rules Test. The Space Fortress Rules Test is a paper and pencil test that is composed of 30 multiple-choice questions that assess an individual's knowledge of Space Fortress instructions, procedures, and information. Trainees will have 7 minutes to complete the test.

Criterion Score. The skill level of participants on Space Fortress will be assessed using a criterion score. The criterion will be defined as a total game score of 2500 points on at least 60% of games played in one 10-game session. Gopher et al. (1994) found that the average total game score on Space Fortress after 10 hours of training, was 2251 points. Therefore, it was determined that a total score of 2500 points would be representative of experienced performance.

Computer Tasks

Aiming Task. The aiming task is a qualifying task which will be used to test trainees' ability to control the joystick. The task requires players to destroy mines that sequentially appear on the computer screen as quickly as possible by controlling a space ship with the joystick. Players control the rotation of the ship to aim, shoot and destroy the mines as they appear on the screen. The ship always stays in the center of the screen and can be rotated 360 degrees around its axis to shoot the mines. Point score on the aiming task is a combination of scores from the number of mines destroyed and the speed at which mines are destroyed.

Space Fortress. The 4.5 version of Space Fortress developed by Gopher et al. (1994) will be used in this study. In the game, there is a Space Fortress centered in the middle of the screen surrounded by two concentric hexagons. The player controls the space ship with a joystick and controls the weapon system with a 3-button mouse. For right-handed players, the joystick is positioned on the right side and the 3-button mouse on the left side and vice versa for left-handed players. The 3-button mouse is used for collecting bonuses and identifying foe mines. See Gopher et al. (1994) for a description of the game.

The main goal of Space Fortress is to obtain the highest Total Score. The Total Score is calculated by summing the four component scores. High scores are obtained when the rules of the game are followed and optimal strategies are used. The optimal strategies are: 1) to circle the Space Fortress slowly while staying within the region enclosed by the two hexagons, 2) to shoot the mines as quickly as possible while staying within the hexagons, 3) to stay on the hexagonal course if a friend mine is incorrectly identified, and let the mine destroy the ship, because it will not be able to be destroyed or energized; and 4) to take bonus points instead of missiles unless the number of shots remaining is less than 50.

Procedure

Screening. Before trainees will be allowed to participate in the study they must pass the Space Fortress Aiming task. Each trainee will play three 3-minute games and they must score a minimum of 780 points on at least one of the 3 games to qualify for the study.

Preliminary Assessment. On day 1 of training, trainees will complete a series of videogame questionnaires, cognitive assessment tests, and a handedness inventory. First, trainees will be given the paper and pencil videogame survey. Following the survey, the Edinburgh Inventory will be presented to the trainees using Micro Experimental Laboratory (MEL). Next, trainees will complete the videogame experience questionnaire, which will also be presented on the computer using MEL. After finishing the videogame experience questionnaire, trainees will receive a 5-minute break followed by the Cognitive Ability Measures (CAM) tests.

Training Week 1. Training will take place over two and a half days. The morning of day 1 will be reserved for administering preliminary assessment measures. The afternoon session will begin by giving trainees the written rules of Space Fortress and by showing them a Space Fortress instruction video. Any questions about Space Fortress, the rules or the strategies will be answered at this time.

Each trainee will play a four-game baseline session of Space Fortress at his or her own computer station. Trainees will then be paired according to their station location. A large divider will separate every other computer so that stations are divided into groups of two. Males and females will be paired separately so that only females will work with

females and only males will work with males. This will help to eliminate gender interactions during training.

Training will consist of a group of training sessions. Each training session includes eight 3-minute practice games of Space Fortress and two 3-minute test games. The test games are identical to the practice games but trainees will be told that those two games are test games. For the eight practice games, dyads will take turns playing Space Fortress and observing their partner play Space Fortress. Observers and performers will switch roles after every game. Therefore, trainees will play a total of four games and will observe a total of four games. Each trainee in the dyad will be tested individually without his or her partner observing. Trainees will complete three training sessions on day 1. After each training session, trainees will be given a 10-minute break.

On day 2, trainees will begin by watching a 10-minute summary video of the Space Fortress instructions and any questions about the game, rules, or strategies will be answered. The same procedure used for training on day 1 will be followed on day 2. Trainees will complete a total of eight sessions on day 2. Only a half-day of testing will take place on day 3. After the retention test all trainees will complete eight practice games using the same dyadic procedure as on days 1 and 2.

Training Week 2. Male and female trainees who do not reach the criterion of 2500 points on Space Fortress will be asked to return for a second week of additional training. Additional training will take place over two and a half days. Week 2 will be similar to the first week of training, but trainees will be allowed to work in teams that consist of no more than four individuals. Initially, all members of a team must be of the same gender.

At the end of each session, trainees will respond to a set of questions that ask them to describe their performance in the previous training session. Responses to these questions will be hand written in a journal, which will be collected at the end of the day. Questions will ask trainees to describe different aspects of their performance as well as strategies that they are using. Trainees who do not reach criterion will continue training on day 2 and day 3 if necessary. The procedure on day 2 will be similar to day 1. Trainees will continue to write in their

journals after each session and journals will be collected at the end of the day.

A half-day of training is scheduled for day 3 for any trainee who has not yet reached criterion. At the end of training on day 3, trainees' performance will be reviewed to determine if additional training is necessary. Any trainee who does not reach criterion will be asked to return for another week of additional training.

Training Week 3. The third week of training will take place over two and a half days. During this week however, trainees will not be required to keep journals or observe other players. Trainees will work individually to receive as much hands-on practice as possible on Space Fortress. Any trainee who does not reach criterion will be allowed to return the following week for additional training. The same procedure will continue for subsequent training sessions.

References

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. Journal of Personality and Social Psychology, 41 (3), 586-598.

Barnett, M. A., Vitaglione, G. D., Harper, K. K. G., Quackenbush, S. W., Steadman, L. A., & Valdez, B. S. (1997). Late adolescents' experiences with and attitudes toward videogames. Journal of Applied Social Psychology, 27 (15), 1316-1334.

Berry, D. C. (1991). The role of action in implicit learning. Quarterly Journal of Experimental Psychology, 43, 881-906.

Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. Quarterly Journal of Experimental Psychology, 36A, 209-231.

Blandin, Y., Proteau, L., & Alain, C. (1994). On the cognitive processes underlying contextual interference and observational learning. Journal of Motor Behavior, 26, 18-26.

Braun, C. M., & Giroux, J. (1989). Arcade video games: Proxemic, cognitive and content analyses. Journal of Leisure Research, 21 (2), 92-105.

- Brown, R. M., Hall, L. R., Holtzer, R., Brown, S. L., & Brown, N.L. (1997). Gender and video game performance. Sex Roles, 36 (11-12), 793-812.
- Carroll, W. R., & Bandura, A. (1982). The role of visual monitoring in observational learning of action patterns: Making the unobservable observable. Journal of Motor Behavior, 14, 153-167.
- Carroll, W. R., & Bandura, A. (1985). Role of timing of visual monitoring and rehearsal in observational learning of action patterns. Journal of Motor Behavior, 17, 269-281.
- Cochran, K. F., & Wheatley, G. H. (1989). Ability and sex-related differences in cognitive strategies on spatial tasks. Journal of General Psychology, 116 (1), 43-55.
- Compeau, D. R. & Higgins, C. A. (1995). Application of social cognitive theory to training for computer skills. Information Systems Research: A Journal of the Institute of Management Sciences, 6 (2), 118-143.
- Computer Gaming World (1998). July, 168.
- Corsi-Cabrera, M., Ramos, J., Guevara, M. A., Arce, C., & Gutierrez, S. (1993). Gender differences in the EEG during cognitive activity. International Journal of Neuroscience, 72, 257-264.
- Dorval, M., & Pepin, M. (1986). Effects of playing a video game on a measure of spatial visualization. Perceptual and Motor Skills, 67, 159-162.
- Gagnon, D. (1985). Video games and spatial skills: An exploratory study. Educational Communication and Technology Journal, 33, 263-275.
- GamePro (1998) July, 118.
- Goettl, B. P., & Connolly-Gomez, C. (1995). The role of observational learning in automated instruction of complex tasks. Proceedings of the 39th Annual Meeting of the Human Factors and Ergonomics Society, 2, 1335-1339.
- Gopher, D. (1982). A selective attention test as a predictor of success in flight training. Human Factors, 24, 173-183.
- Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. Human Factors, 36 (3), 387-405.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. Acta Psychologica, 71, 147-177.

Gordon, H. W., & Kravetz, S. (1991). The influence of gender, handedness, and performance level on specialized cognitive functioning. Brain and Cognition, 15, 37-61.

Greenfield, P. M., deWinstanley, P., Kilpatrick, H., & Kaye, D. (1996). Action video games and informal education: Effects on strategies for dividing visual attention. In P. M. Greenfield & R. R. Cocking (Eds.), Interacting with video (pp. 187-205). Norwood, NJ: Ablex Publishing Corporation.

Hart, S. G., & Battiste, V. (1992). Field test of video game trainer. Proceedings of the Human Factors Society 36th Annual Meeting

Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analysis using w^2 and d . American Psychologist, 36, 892-901.

Jacklin, C. N. (1989). Female and male: Issues of gender. American Psychologist, 44 (2), 127-133.

Jones, M. B. (1984). Video games as psychological tests. Simulation and Games, 15, 131-157.

Kiddoo, T. (1982). Pacman meets GI Joe? Soldiers, 37 (9), 20-23.

Kyllonen, P. C., Woltz, D. J., Christal, R. E., Tirre, W. C., Shute, V. K., & Chaiken, S. (1990). CAM-4: Computerized battery of cognitive ability tests. Unpublished computer program, Brooks Air Force Base, TX.

Law, D. L., Pellegrino, J. W., & Hunt, E. B. (1993). Comparing the tortoise and the hare: Gender differences and experience in dynamic spatial reasoning tasks. Psychological Science, 4, 35-40.

Locke, E. A., Frederick E., Lee, C., & Bobko, P. (1984). Effect of self-efficacy, goals and task strategies on task performance. Journal of Applied Psychology, 69 (2), 241-251.

Maccoby, E. E., & Jacklin, C. N. (1974). The psychology of sex differences. Stanford, CA: Stanford University Press.

McCullagh, P., & Caird, J. K. (1990). Correct and learning models and the use of the model knowledge of results in the acquisition and retention of a motor skill. Journal of Human Movement Studies, 18, 107-116.

- Okagaki, L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. Special issue: Effects of interactive entertainment technologies on development. Journal of Applied Developmental Psychology, 15 (1), 33-58.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologica, 9, 97-113.
- Paulus, L., & Shebilske, W. L. (1998). Dyadic protocols and the acquisition of complex skills. Unpublished master's thesis, Texas A & M University, College Station, Texas.
- Philbin, M., Meier, E., Huffman, S., & Boverie, P. (1995). A survey of gender and learning styles. Sex Roles, 32 (7-8), 485-494.
- Robinson-Staveley, K. R., & Cooper, J. (1990). Mere presence, gender and reactions to computers: Studying human-computer interaction in the social context. Journal of Experimental Social Psychology, 26, 168-183.
- Shebilske, W. L., Goettl, B. P., & Regian, J. W. (1996). Executive control and automatic processes as complex skills develop in laboratory and applied settings. Paper presented at Attention and Performance XVII: Cognitive regulation of performance: Interaction of theory and application. Belf Oren, Israel.
- Shebilske, W. L., Regian, J. W., Arthur, W. Jr., & Jordan, J. A. (1992). A dyadic protocol for training complex skills. Human Factors, 34, 369-374.
- Shute, V. J., & Psotka, J. (1996). Intelligent tutoring systems: Past, present, and future. In D. Jonassen (Ed.), Handbook of research on educational communications and technology: A project of the Association for Educational Communication (pp. 570-600). New York, NY: McMillan.
- Stiles-Shipley, J. A., Goettl, B. P., Tweney, R. D., & Derek, K. (1998). The effects of training schedule, observation, and model skill level on the acquisition of a complex computer-based task. Manuscript in preparation. Bowling Green State University, Ohio.
- Subrahmanyam, K., & Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. Special issue: Effects of interactive entertainment technologies on development. Journal of Applied Developmental Psychology, 15 (1), 13-32.

**WEIGHING THE IMPORTANCE OF SPATIAL ORGANIZATION
AND PRIORITY OF TARGETS ON UAV MISSION PLANNING:
WHICH FACTORS CONTRIBUTE MOST TO A TARGET'S
ATTRACTIVENESS?**

**Michael Stiso
Department of Psychology**

**Texas A&M University
College Station, TX 77843**

**Final Report for:
Summer Graduate Student Research Program
Armstrong Research Site**

**Sponsored by:
Air Force Office of Scientific Research
Brooks AFB, TX**

and

Armstrong Research Site

August 1998

WEIGHING THE IMPORTANCE OF SPATIAL ORGANIZATION
AND PRIORITY OF TARGETS ON UAV MISSION PLANNING:
WHICH FACTORS CONTRIBUTE MOST TO A TARGET'S
ATTRACTIVENESS?

Michael E. Stiso
Department of Psychology
Texas A&M University

Abstract

A relatively unstudied member of the UAV crew is the DEMPC, the person in charge of reconnaissance mission planning and preplanning. Part of the reason for that neglect is the difficulty of teasing apart all of the different factors influencing the DEMPC's decisions. The task described here is an initial attempt at such a study that simplifies the process by focusing on just the two most salient factors of those a DEMPC must consider: spatial organization and priority of targets in a ROZ box. The theory is that, when faced with several targets and a paucity of information, DEMPCs will mentally separate those targets into clusters based on Gestalt perceptual and organizational rules. Those clusters will be more or less attractive depending on the conditions of several factors surrounding them. Specifically, I propose that five features will be most salient at the start of a mission and so will also prove most influential in determining a cluster's initial attractiveness: (in order from strongest to weakest) proximity of the cluster to the ROZ box entrance, proximity to the exit, the density of the cluster, the size of the cluster, and the percentage of high-priority targets in the cluster. Preliminary results from a run of a previous version of the task show that at least some of these factors—particularly proximity to entrance and the number of priority targets—do indeed make a cluster more attractive. If a full run of the current version also proves successful, it should be a large step forward in the eventual development of decision-making support and automation systems for the DEMPC position.

WEIGHING THE IMPORTANCE OF SPATIAL ORGANIZATION
AND PRIORITY OF TARGETS ON UAV MISSION PLANNING:
WHICH FACTORS CONTRIBUTE MOST TO A TARGET'S
ATTRACTIVENESS?

EXCERPT FROM AN INTERVIEW WITH AN EXPERIENCED DEMPC/PLO:

Interviewer: *(Shows DEMPC a map.) Here are some [numbered] targets in a ROZ box. Let's assume that this is your aircraft here. You've just entered this ROZ box, and you're the DEMPC. ...What would be your next action?*

DEMPC: *Right. ...Tell the payload, let's see if we can hit 5. OK, boom, we got 5, 5 is satisfied. OK, now let's say we are right here, let's take a look back at 4. And then if we can't get 4, or 5, then let's say we take a gander at 7, see if we can get a good shot at 7. And then if not, maybe we steer this way, and kind of flatten out a little bit, then try to get these two on the back side. But now we've set ourselves up for 6. ...*

The DEMPC here makes the task he's given look deceptively simple: With little more than a cartoon map of some targets and his current position, he is able to make a snap assessment of the situation and to design a quick flight plan accordingly. However, his decisions are actually the result of a complex, perhaps intuitive risk-benefit assessment of several factors acting in combination. In fact, the plan he describes is not unlike the planning of expert billiards players: Just as a good pool player will choose her shot not just to sink a ball but also to leave the cue ball in a good location for a subsequent shot (possibly even a few moves ahead), so to is the DEMPC planning his target acquisition in such a way that future targets are more easily obtained because of the positioning of the UAV during earlier targets. Moreover, his plan provides flexibility for getting around unknown but potential obstacles along the way, such as cloud cover or high target altitude.

Such efficient decision making requires a "feel" for the situation and is a matter of expertise. Unfortunately, such expertise is difficult to study because of the large number of factors, both explicit and implicit, that the DEMPCs must consider. Some of the more salient factors in mission design include cloud cover, wind speed, and intertarget distance. Other factors, however, are somewhat more hidden but just as

important. For example, the UAV can climb only so far and so fast; even if two targets are in close proximity to each other on the map, getting from one to the other may be impossible or ridiculously inefficient if one is at sea level and the other's on a mountain. Another factor that is not readily apparent is the type of target that must be collected. Aircraft hangars, for instance, can only be collected from the side on which the hangar door is located; by itself, that bit of information would seem simple enough to incorporate into mission designs. However, DEMPCs know that the UAV also has to be at the right angle in order to photograph the inside of the hangar, and that that angle in turn depends on the UAV's distance from the hangar. What's more, the conditions around the hangar can further complicate the matter: What if the sun happens to be on the opposite side of the hangar door, preventing a good picture, or if there is cloud cover over the section of sky in which you need to be? Add wind speed to the mix, and the simultaneous estimation of the impact of all the different factors seems difficult at best. Yet, expert DEMPCs are able to do it with little more than a glance.

In fact, such an assessment of multiple factors may be where expertise becomes most important and apparent: Whereas novices will likely be able to consider only one factor (usually the most salient) at a time when designing a mission plan, DEMPCs are able to consider the impact of several at once and to design an efficient plan accordingly. And importantly, because of their experience, DEMPCs are able to provide enough flexibility in their plans to deal with unknown but *potential* obstacles, thus demonstrating the predictive ability provided by expertise.

But teasing apart the impact of several different variables alone and in combination on DEMPC decision making is difficult; there is simply too much information to analyze. We know, for example, that cloud cover is an important consideration for UAV crews, but at what level of cover does photographing that target become impossible? When does the amount of cover warrant loitering, hoping a gap comes along? And how does the current wind speed, the UAV's altitude, the target's priority, or even this target's place in line influence that decision? Also, what's more influential on the DEMPC's decision to collect a certain target: low cloud cover or its proximity to the UAV? Is that always the case? When is it not?

Why study mission planning?

Such questions are just the tip of the iceberg when dealing with the interaction of so many variables. Yet, they are important questions to answer if UAV technology is to progress beyond its current state. For example, the automation of mission planning and preplanning is a logical next step for UAVs. But to design such systems, we need to know which factors the DEMPC takes into consideration and what role each plays alone and in concert; in essence, we need a model of DEMPC expertise. And such a model is not likely to happen unless we can predict and evaluate DEMPC decision making in mission planning scenarios.

That model would have other uses, as well, of course. An application related to the automated systems would be the design of a computerized decision-making aid to support DEMPC performance and to free mental resources. The model would also indicate appropriate areas to emphasize in training: Those factors that produce bigger differences between novice and DEMPC performance are likely to be those that would benefit most from training.

To that end, two tasks have been designed. The first is an attempt by Josh Hurwitz and Robin Thomas (in development) of the AFOSR to study how the interactions of a vast number of variables—including cloud cover, wind speed, target and UAV altitude, and UAV rate of climb—influence mission design. The second is a complement to the first, sacrificing the range of variables in order to focus more closely on just a couple: spatial organization and priority of targets in a ROZ box.

Because the Hurwitz and Thomas study provides performance feedback as well as target information, it is likely to be most similar and applicable to what DEMPCs experience during mission replanning, when the UAV is already underway. On the other hand, by dealing with only spatial organization and priority, which comprises most of the information the DEMPCs have at the start, I believe my study is more applicable to the preplanning stage than it is to replanning.

The salience of spatial organization

Aside from a list of the types and priorities of the targets that they need to collect, the DEMPCs have primarily visual information at the start of a mission: a map of the target locations. Given the salience of those targets' proximal relationship to each other, as well as the absence of other information, it's likely that Gestalt perceptual-organizational processes will exert a strong influence on the direction of decision making. In particular, especially when dealing with several targets, DEMPCs are likely to mentally separate the targets into groups or clusters based on such factors as their location, their proximity to other targets, and the number of other targets nearby.

Certain characteristics of these clusters are likely to make one more attractive than another. For example, a cluster containing several more targets than another might be considered more attractive because of the potential for capturing more targets with less effort; hence, the DEMPC would be likely to head to that cluster first because it would allow a more efficient collection of targets, particularly in the event of time pressures. Other factors likely to influence a cluster's attractiveness include its density, its relation to the entrance and exit of the ROZ box, and the number of high-priority targets it contains.

As mentioned above, the salience of these organizational factors are likely to be most influential on novices; moreover, the single most salient or attractive factor should determine the route the novice chooses. DEMPCs, on the other hand, should have the expertise to choose the most efficient route, shrugging off the lure of the salient variables to consider them all in combination.

To test the effects of spatial organization and priority on mission preplanning, I presented subjects with two clusters of targets (X and Y) and asked them to design a mission plan. With each cluster having a specific manipulation, the goal was to see which manipulation made the clusters more attractive— attractiveness being operationalized as the cluster the subjects chose to head for first. There were five manipulations, two of which dealt specifically with the layout of the clusters: Size (number of targets) and Density (intertarget distance). The other three included Priority, which refers to the number of high-priority targets in a cluster; Entrance location relative to cluster X; and Exit location relative to cluster X. In

addition to those five independent variables, there was one dependent variable, which was the frequency that subjects clicked on targets in Cluster X first (see Table 1).

I hope to examine seven hypotheses with such manipulations:

1. *Entrance Effect* — All else being equal, the cluster closest to the entrance will be chosen first.
2. *Exit Effect* — All else being equal, the cluster closest to the exit will be chosen first. (In the early stages of this experiment, I had thought that the cluster farthest from the exit would be chosen first: Heading to that cluster, the subject might be able to pick up a few targets from the closer cluster as he or she flew toward the exit. However, I later realized that that feature would not be as salient as the time-saving proximity of a close exit.)
3. *Big-Cluster Effect* — All else being equal, the bigger cluster (number of targets) will be chosen first.
4. *Dense-Cluster Effect* — All else being equal, the denser cluster will be chosen first.
5. *Priority Effect* — All else being equal, the cluster having a higher proportion of high-priority targets will be chosen first.
6. All else being equal, the *Entrance Effect* will be stronger than the *Exit Effect*, which will be stronger than the *Big-Cluster Effect*, which will be stronger than the *Dense-Cluster Effect*, which will be stronger than the *Priority Effect*.
7. All else not being equal, the simultaneous presence of several constraining variables should have the greatest effect on experts' performance—in other words, novices should still follow the ranking in hypothesis 6, whereas experts' choices should more often differ from that ranking.

Those hypotheses are mostly speculation based on an educated guess concerning the order of importance of the different variables. My task, being exploratory in nature, should either confirm that ranking or provide an alternate one.

Methods

Subjects (mostly novices because of the difficulty of obtaining DEMPCs to study) are presented with a square ROZ box about 4.5" on a side that contains two groups of targets (Clusters X and Y). The number, layout, and priority of those targets is determined by the type of cluster to which they belong (see Table 2). The targets are either red dots (normal priority) or yellow (high priority). In addition, the locations of the Entrance and Exit vary from ROZ box to ROZ box. The Entrance is a green arrow that can appear on any of the four sides of the box and in one of three locations (Table 3): near to Cluster X (Near-X), in between Clusters X and Y (Neutral), and far from Cluster X (Far-X). The Exit is a red arrow that appears on the side opposite the Entrance and in one of three locations, similar to those of the Entrance.

There are five constants for each trial in this task: The Entrance always appears in one of three fixed locations particular to that side of the box; the Exit always appears in one of three fixed locations particular to the side of the box opposite the Entrance; the X and Y clusters are always side by side; the two clusters are always lined up perpendicularly to the Entrance and Exit; and unless the cluster condition specifies otherwise, the number of normal and priority targets, the average intertarget distance, and the average target distance from the Entrance and Exit is the same for each cluster. The task as a whole has one constant: The condition of every variable is matched with the condition of every other variable exactly once. For example, subjects will encounter only one trial in which a Near-X Entrance and Neutral Exit are paired with a Dense Cluster X and a Big Cluster Y.

Subjects have one hour (not real time) to collect as many targets as they can. The ROZ box is described as being 45 nautical miles on a side, and the UAV speed is 60 mph. Given those parameters, one hour is not enough time to collect all of the targets, so subjects must choose a path that will let them collect as many as they can in the time allowed. They do that by using the mouse to click on one or more targets at a time. The mouse gives them control of a circular cursor inside the box that is about half-an-inch in diameter and which represents the range of the UAV's collection capability. Each click produces a colored circle in that location; the color fades as the distance between clicks adds up, indicating that time is running out. After each click, and starting from the Entrance arrow, an animated plane travels from the previous

collection point to the current one. The plane flies at a fixed rate and is intended to give subjects a "feel" for how long it takes to travel across the box.

Subjects get one point for each red target they collect, and three points for each yellow. However, if they run out of time, they lose all points for that box. After every four missions they complete, they are told how many total points they scored for those four boxes and the total time it took them to do so. The feedback is provided after every four trials rather than after every trial in an attempt to minimize learning effects.

This task has gone through a couple of revisions to get to its present state. In the first version, subjects were required simply indicate which of two targets—X and Y, corresponding to the cluster to which they belonged—they would head to first if they were to design a mission plan. They would then just press "X" or "Y" on the keyboard. The appearance of the task was for the most part the same as in the current version; however, it was later suggested that the simplicity of the task and the lack of any feedback would leave subjects with little if any motivation to complete or even to pay attention to it.

To combat potential subject boredom, then, I added interactive elements such as target-clicking and feedback. However, when the experiment was actually run, the subjects—contrary to the instructions they were given—reportedly clicked on every target rather than just a few. The likely reason for that behavior is that they simply didn't read the instructions, which, in retrospect, were somewhat complex. In addition, though the fading color of the cursor after each click gave an estimate of the time remaining, it apparently didn't give subjects a good enough feel for the time required to get from point A to point B. As a fix, I added an animated plane that travels at a fixed rate between clicks (a suggestion by Patrick Kyllonen), which allowed me to shorten the instructions and which should also serve as a good indicator of travel time between points.

Results

Data have not yet been collected using the newest version of the task. However, the results from the run of the previous version do provide some indication of spatial organization and priority effects. No statistical analyses have been run on that data, but looking over it, it is apparent that, summing across cluster types,

Cluster X was consistently chosen more often when it was near the Entrance, less often when paired with a Neutral Entrance, and least often when the Entrance was far (Table 4). That supports the idea that having the Entrance in close proximity to a cluster is a big boost to that cluster's attractiveness. The location of the Exit, on the other hand, apparently had little effect (Table 5).

Overall, Cluster X was chosen more often when it was a Priority cluster, and somewhat less often when it was a Big cluster (Table 6). Surprisingly, the Dense and Normal clusters didn't differ much, which may indicate that cluster density isn't a very salient feature—at least, not for novices. That pattern was consistent across both the Entrance and the Exit conditions, but appeared to be strongest when summing across the Exits. In addition, the interactions of the different cluster types with each other showed the same pattern (Table 7): Cluster X was chosen most often when it was a Priority cluster, regardless of the Cluster Y condition; X was also chosen almost as often when it was a Big cluster, except, notably, when Cluster Y was a Priority cluster. Dense clusters seemed preferable only to Normal clusters.

Conclusion

In sum, then, it would seem from these preliminary findings that having the entrance close by and having a large number of high-priority targets are both important factors in the determination of a cluster's attractiveness. The size of the cluster is somewhat less salient but still has a noticeable effect on decision making. However, at least for novices, the density of the cluster and its proximity to the exit have little effect on that cluster's pull.

It would be interesting to test expert DEMPCs to see whether density and exit effects increase with expertise. It may be that the novice subjects are approaching the problem piecemeal rather than as a whole; in other words, they are starting at the entrance, looking at the clusters, finding one more attractive because of its obvious advantages (i.e. it has a bunch of targets or a high percentage of high-priority targets), and then heading there. In such start-to-finish planning, the exit—being at the end—wouldn't be much of a consideration; the entrance, however, would be the primary influence. DEMPCs, on the other hand, may assess the entire situation at once rather than just one piece at a time; instead of trying to cross

bridges when they come to them, they may have an overhead view that lets them consider the bridges ahead of time.

This is all speculation, of course: The results are unreliable, given that the subjects didn't follow the instructions and that the data were not analyzed statistically. However, the fact that there were noticeable and consistent patterns of variation makes me hopeful that the newest version, when run, will be successful in describing a ranking of the various factors involved. And if the task actually does produce some good results, it would be a small matter to add other factors into the mix, such as cloud cover, terrain, and target type, in order to broaden the range of that ranking. The end result of such information would be the assignment of some sort of attractiveness score to a given target, based on the conditions of the factors surrounding that target and the cluster in which it resides. That scoring system would allow the prediction of DEMPC flight plans as well as an evaluation of DEMPC performance. Perhaps more importantly, it would make possible the design of decision-making systems in order to automate or aid the task of the DEMPC in UAV missions.

Tables

Table 1: Variables

Independent	Dependent
Cluster's proximity to Entrance	Frequency Cluster X is chosen
Cluster's proximity to Exit	
Density of cluster	
Size of cluster (number of targets)	
Priority of cluster	

Table 2: Cluster Types

Normal	Your basic cluster
Big	More targets than in Normal clusters
Dense	Targets closer together than in Normals
Priority	More high-priority targets than in Normals

Table 3: Entrance/Exit Locations

Near X	Closer to Cluster X
Neutral	Between Clusters X and Y
Far X	Closer to Cluster Y

Note: For Tables 4-7, the numbers in the cells represent the percentage of subjects choosing to go to Cluster X first in their flight plan for that particular condition.

Table 4: Entrance Effects

Far-X	Near-X	Neutral
88.63%	95.26%	93.46%

Table 5: Exit Effects

Far-X	Near-X	Neutral
92.65%	92.28%	92.42%

Table 6: Cluster Effects

Cluster X			
Big	Dense	Normal	Priority
92.86%	91.83%	91.13%	93.98%

Table 7: Cluster Interactions

Cluster Y	Cluster X			
	Big	Dense	Normal	Priority
Big	93.00%	91.07%	91.73%	94.65%
Dense	93.05%	91.77%	90.66%	93.09%
Normal	94.32%	93.95%	91.32%	94.90%
Priority	91.07%	90.53%	90.82%	93.29%