AFWAL-TR-83-3021

(2)

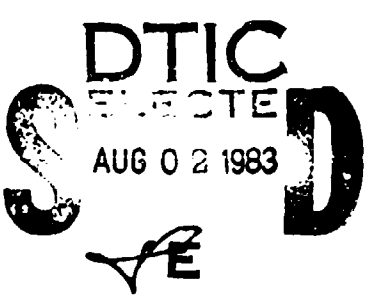
AD A131256

PROCEEDINGS OF THE EIGHTEENTH ANNUAL CONFERENCE ON MANUAL CONTROL

DAYTON, OHIO
JUNE 8 - 10, 1982

Editor: Frank L. George

JANUARY 1983



INTERIM REPORT - JUNE 1981 TO JUNE 1982

DTIC FILE COPY

Approved for public release; distribution unlimited

FLIGHT DYNAMICS LABORATORY
AIR FORCE WRIGHT AERONAUTICAL LABORATORIES
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

83 08 2 . 056

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture use, or sell any patented invention that may in any way be related thereto.

This report has been reviewed by the Office of Public Affairs (ASD/PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

Frank L. George

FRANK L. GEORGE
Project Engineer,
Control Dynamics Branch

R.O. Anderson

RONALD O. ANDERSON, Chief
Control Dynamics Branch
Flight Control Division

FOR THE COMMANDER

Charles R Goss Jr

CHARLES R. GOSS, JR, LT COL, USAF
Chief, Flight Control Division
Flight Dynamics Laboratory

"If your address has changed, if you wish to be removed from our mailing list, or if the addressee is no longer employed by your organization please notify AFWAL/FIGC, W-PAFB, OH 45433 to help us maintain a current mailing list".

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: AFWAL-TR-83-3021	2. GOVT ACCESSION NO.: AD-A131 256
	3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): Proceedings of the Eighteenth Annual Conference on Manual Control	5. TYPE OF REPORT & PERIOD COVERED: Interim, June 81 - June 82
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Frank L. George, Editor	8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Flight Dynamics Laboratory (FIGC), AF Wright Aeronautical Laboratories, W-P AFB, Ohio 45433	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62201F, 999135RX
11. CONTROLLING OFFICE NAME AND ADDRESS: Flight Dynamics Laboratory (FIGC), AF Wright Aeronautical Laboratories, W-P AFB OHIO 45433	12. REPORT DATE: January 1983
	13. NUMBER OF PAGES: 556
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approved for Public Release; Distribution Unlimited	
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)	
18. SUPPLEMENTARY NOTES: → These were presented at the 45	
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Human Dynamics; Human Modeling; Human-machine System; Manual Control; Decision Making; Attention Allocation; Simulation; Control Manipulators; Displays; Workload	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This volume contains proceedings of the Eighteenth Annual Conference on Manual Control, held at Dayton, Ohio, June 8-10, 1982. Forty-three of the forty-five Conference papers, are represented either as abstracts or complete manuscripts. Topics covered, human operator modeling from both performance and physiological viewpoints, measures of operator workload and performance, and applications of models for analysis alone or in conjunction with simulation. Applications areas ~~included~~ control

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

and controller design, analysis and definition of display requirements, and integration of control and display considerations.

EIGHTEENTH

ANNUAL

conference on

MANUAL

control

DTIC COPY INSPECTED 2

Accession For
NTIS GRA&I
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

Dayton, Ohio

June 8-10, 1982

CONFERENCE CHAIRMAN:

Frank L. George
Air Force Wright Aeronautical Labs

AIR FORCE HOST:

Ronald O. Anderson
Air Force Wright Aeronautical Labs

FOREWORD

This volume, published with the support of the Air Force Wright Aeronautical Laboratories, contains the proceedings of the *Eighteenth Annual Conference on Manual Control* held at the Daytonian Hotel, Dayton, Ohio from June eighth through tenth, 1982. All papers accepted for the Meeting are represented in this volume. In a few cases authors were unable to attend and present accepted papers, or authors who presented papers failed to provide manuscripts for the proceedings. Those cases are noted in the Table of Contents. Both formal papers, generally representing completed work, and informal papers which might represent work in progress were presented.

This was the eighteenth in a series of Conferences dating back to December 1964. These earlier meetings and their proceedings are listed below:

First Annual NASA-University Conference on Manual Control, The University of Michigan, December 1964. (Proceedings not printed.)

Second Annual NASA-University Conference on Manual Control, Massachusetts Institute of Technology, February 28 to March 2, 1966, NASA-SP-128.

Third Annual NASA-University Conference on Manual Control, University of Southern California, March 1 through 3, 1967, NASA-SP-144.

Fourth Annual NASA-University Conference on Manual Control, The University of Michigan, March 21 through 23, 1968, NASA-SP-192.

Fifth Annual NASA-University Conference on Manual Control, Massachusetts Institute of Technology, March 27 through 29, 1969, NASA-SP-215.

Sixth Annual Conference on Manual Control, Wright-Patterson AFB, Ohio, April 7 through 9, 1970, proceedings published as AFIT/AFFDL Report, no number.

Seventh Annual Conference on Manual Control, University of Southern California, June 2 through 4, 1971, NASA-SP-281.

Eighth Annual Conference on Manual Control, The University of Michigan, May 17 through 19, 1972, AFFDL-TR-72-92.

Ninth Annual Conference on Manual Control, Massachusetts Institute of Technology, May 23 through 25, 1973, proceedings published by MIT, no number.

Tenth Annual Conference on Manual Control, Wright-Patterson AFB, Ohio, April 9 through 11, 1974, AFIT/AFFDL Report, no number.

Eleventh Annual Conference on Manual Control, NASA Ames Research Center, May 21 through 23, 1975, NASA TM X-62,464.

Twelfth Annual Conference on Manual Control, University of Illinois, May 25 through 27, 1976, NASA TM X-73,170.

Thirteenth Annual Conference on Manual Control, Massachusetts Institute of Technology, June 15 through 17, 1977, proceedings published by MIT, no number.

Fourteenth Annual Conference on Manual Control, University of Southern California, April 25 through 27, 1978, NASA CP-2060.

Fifteenth Annual Conference on Manual Control, Wright State University, March 20 through 22, 1979, AFFDL-TR-79-3134.

Sixteenth Annual Conference on Manual Control, Massachusetts Institute of Technology, May 5 through 7, 1980, proceedings published by MIT, no number.

Seventeenth Annual Conference on Manual Control, University of California at Los Angeles, June 16 through 18, 1981, JPL Publ. 81-95.

The topic "human-machine system design methodology", both development and applications, received special emphasis at the Eighteenth Conference meeting.

Frank L. George,
Air Force Wright Aeronautical Labs

CONTENTS

PAGE

FOREWORD iii

SESSION 1. HUMAN OPERATOR MODELS
Moderator: David L. Kleinman, Univ of Connecticut . . . 1

1. *AN INFORMATION THEORETIC MODEL OF THE HUMAN OPERATOR (FORMAL)* Charles P. Hatsell, Patricia R. Mineer 3

2. *RISK AND DECISION PROCESSES IN MANUAL CONTROL BEHAVIOR (FORMAL)* Tom Bösser 5

3. *TIME DOMAIN IDENTIFICATION OF PILOT DYNAMICS AND CONTROL STRATEGY (FORMAL)* David K. Schmidt 19

4. *A TEST OF FITTS' LAW IN TWO DIMENSIONS WITH HAND AND HEAD MOVEMENTS (INFORMAL)* Richard J. Jagacinski, Don L. Monk (Abstract Only) 41

5. *A NONLINEAR INTERNAL FEEDBACK MODEL FOR THE PUPIL CONTROL SYSTEM (FORMAL)* Fuchuan Sun, William C. Kreuz, Lawrence W. Stark (Paper Accepted but Not Presented) 42

6. *COMPUTATIONAL PROBLEMS IN HUMAN OPERATOR ARMA MODELS (INFORMAL)* Gyan C. Agarwal, Hitoshi Miura, Gerald L. Gottlieb 58

SESSION 2. PHYSIOLOGICAL MEASURES; WORKLOAD AND PERFORMANCE MEASURES Moderator: Ronald O. Anderson, Air Force Wright Aeronautical Laboratories 75

1. *CHANGES IN HUMAN JOINT COMPLIANCE DURING PERIPHERAL ISCHEMIA (FORMAL)* Robert J. Jaeger, Gyan C. Agarwal, Gerald L. Gottlieb 77

2. *DECISION MAKING AND THE STEADY STATE VISUALLY EVOKED EEG (INFORMAL)* Andrew M. Junker, Glen F. Wilson 99

3. *OPERATOR WORKLOAD AS A FUNCTION OF THE SYSTEM STATE: AN ANALYSIS BASED UPON THE EVENT-RELATED BRAIN POTENTIAL (INFORMAL)* Richard T. Gill, Christopher Wickens (Paper Accepted but Not Presented) 100

4. *ASSESSMENT OF MANUAL, PROCESS CONTROL STRATEGIES (INFORMAL)* Joseph C. DeMaio 108

5. *DEVELOPMENT OF PERFORMANCE MEASURES FOR SELECTION OF CREWS FOR FLIGHT TRAINING (FORMAL)* Edward M. Connelly (Paper Not Available) 119

CONTENTS (Cont'd)

PAGE

6. *A FUZZY MODEL OF RATHER HEAVY WORKLOAD (FORMAL)* Neville Moray, K. Waterton 120

7. *RATING CONSISTENCY AND COMPONENT SALIENCE IN SUBJECTIVE WORKLOAD ESTIMATION (FORMAL)* Jan R. Hauser, Mary E. Childress, Sandra G. Hart 127

8. *THE SENSITIVITY OF TWENTY MEASURES OF PILOT MENTAL WORKLOAD IN A SIMULATED ILS TASK (FORMAL)* W. W. Wierwille, Sidney A. Connor (Paper Accepted But Not Presented) 150

SESSION 3. SIMULATION AND MODEL BASED ANALYSIS Moderator: Andrew M. Junker, Air Force Aerospace Medical Research Lab . 163

1. *DEVELOPMENT OF A G-SEAT ROLL-AXIS DRIVE ALGORITHM (INFORMAL)* Edward A. Martin, Grant R. McMillan (Abstract Only) . . . 165

2. *COMPENSATION FOR TIME DELAYS IN FLIGHT SIMULATOR VISUAL DISPLAY SYSTEMS (INFORMAL)* D. Francis Crane (Abstract Only) . 166

3. *AN AUTOMOBILE AND SIMULATOR COMPARISON OF DRIVER BEHAVIOR IN A COLLISION AVOIDANCE MANOEUVRE (INFORMAL)* Eric N. Solowka, Lloyd D. Reid 168

4. *AN OPTIMAL CONTROL MODEL ANALYSIS OF DATA FROM A SIMULATED HOVER TASK (FORMAL)* Sheldon Baron 186

5. *ANALYSIS OF A AAA SUPERVISORY CONTROL TASK USING SAINT (INFORMAL)* Kuang C. Wei, Blaza Toman, Joseph Whitmeyer, Sharon L. Ward 207

6. *HUMAN RESPONSE MODELING AND FIELD TEST VALIDATION OF AN AIR DESENSE GUN SYSTEM SIMULATION (FORMAL)* Y. K. Yin, K. Sahara, V. Dea, A. Netch 225

7. *EFFECTS OF PRACTICE ON PILOT RESPONSE BEHAVIOR (INFORMAL)* William H. Levison 239

8. *A MODERN APPROACH TO PILOT-VEHICLE ANALYSIS AND THE NEAL-SMITH CRITERIA (FORMAL)* B. J. Bacon, David K. Schmidt . . . 256

9. *A STRUCTURED METHODOLOGY FOR ANALYZING HUMAN INFORMATION PROCESSING IN COMMAND SYSTEMS (FORMAL)* Joseph G. Wohl, Krishna R. Pattipati, David L. Kleinman, Nils R. Sandell Jr., Elliot E. Entin (Abstract Only) 263

CONTENTS (Cont'd)

PAGE

SESSION 4. MODEL BASED DESIGN AND CONTROL INTEGRATION
Moderator: Sheldon Baron, Bolt, Beranek and Newman, Inc. 265

1. *VISUAL CUE REQUIREMENTS FOR FLIGHT CONTROL TASKS (FORMAL)* Ronald A. Hess, Arthur Beckman (Abstract only) 267

2. *VALIDATION OF AN ADVANCED COCKPIT DISPLAY DESIGN METHODOLOGY VIA WORKLOAD/MONITORING TRADEOFF ANALYSIS (INFORMAL)* Johnathan Korn, Sol W. Gully, David L. Kleinman 268

3. *HUMAN-MACHINE INTERFACE ISSUES IN THE DESIGN OF INCREASINGLY AUTOMATED NASA CONTROL ROOMS (INFORMAL)* Christine M. Mitchell 293

4. *KEYBOARD DESIGN VARIABLES IN DUAL-TASK MODE SELECTION (FORMAL)* Mark D. Hansen (Paper accepted, not presented) 320

5. *AN ANALYSIS OF CONTROL RESPONSES AS A FUNCTION OF DISPLAY DYNAMICS IN PERSPECTIVE AIRCRAFT DISPLAYS WITH PREDICTION (INFORMAL)* Jeffrey L. Maresh, Richard S. Jensen 327

6. *INSTRUMENT POSITION-CONTROL MANIPULATION (IPCOM II): HELICOPTER PILOT RESPONSES AS A FUNCTION OF CONTROL-DISPLAY COMPATIBILITY (INFORMAL)* Kathleen M. Craig 336

7. *HELICOPTER PILOT RESPONSE LATENCY AS A FUNCTION OF THE SPATIAL ARRANGEMENT OF INSTRUMENTS AND CONTROLS (FORMAL)* E. James Hartzell, S. Dunbar, R. Beveridge, R. Cortilla 345

8. *CONTROLS, DISPLAYS AND SITUATION AWARENESS ON THE FINAL APPROACH (FORMAL)* Lee H. Person, Jr. (Abstract only) 365

9. *CHARATERISTICS OF A COMPENSATORY MANUAL CONTROL SYSTEM WITH ELECTROCUTANEOUS FEEDBACK (FORMAL)* Kazuo Tanie, Susumu Tachi, Kiyoshi Komoriya, Hideo Iguchi, Masao Takabe, Minoru Abe, Yoshihiro Sakai 366

SESSION 5: HUMAN-MACHINE SYSTEMS AND CONTROLS Moderator: Ronald A. Hess, NASA Ames Research Center 387

1. *A TRANSFORM APPROACH TO COLLISION FREE PROGRAMMING OF MULTIDEGREE-OF-FREEDOM STRUCTURES (FORMAL)* John G. Kreifeldt, Stephen H. Levine 389

2. *CROSS MODALITY SCALING: MASS AND MASS MOMENT OF INERTIA (INFORMAL)* Margaret M. Clark, Stephen Handel, John G. Kreifeldt 413

3. *PROMAN: PROGRAM FOR MANAGEMENT AND ANALYSIS OF MANUAL CONTROL EXPERIMENTAL DATA (FORMAL)* Ramal Muralidaren, William H. Levison (Abstract only) 419

CONTENTS (Concluded)

PAGE

4. *DEXTEROUS MANIPULATOR LABORATORY PROGRAM (INFORMAL)* Roy E. Olsen 420

5. *STICK CONTROLLER DESIGN FOR LATERAL ACCELERATION ENVIRONMENTS (FORMAL)* D. W. Repperger, W. H. Levison, V. Skowronski, B. O'Lear, J. W. Frazier, K. E. Hudson 427

6. *FORCE-TORQUE CONTROL EXPERIMENTS WITH THE SIMULATED SPACE SHUTTLE MANIPULATOR IN MANUAL CONTROL MODE (INFORMAL)* R. S. Dotson, A. K. Bejczy, J. W. Brown, J. L. Lewis 440

7. *HELICOPTER INTEGRATED CONTROLLER RESEARCH (INFORMAL)* Sherry L. Dunbar 466

8. *TENTATIVE CRITERIA FOR CONTROLLERS OF UNCOUPLED AIRCRAFT MOTION (INFORMAL)* John M. Evans, Kevin Citurs (Paper not available) 474

9. *APPLICATION OF SPEECH SYNTHESIS TECHNOLOGY TO STINGER MISSILE SYSTEM (FORMAL)* M. J. Crisp, Y. K. Yin, L. R. Perkin . . . 475

SESSION 6: DISPLAY FACTORS Moderator: Terry J. Emerson, Air Force Wright Aeronautical Laboratories 483

1. *FACTORS AFFECTING IN-TRAIL FOLLOWING USING CDTI (FORMAL)* James R. Kelly, David H. Williams 485

2. *A PERSPECTIVE DIMENSIONAL DISPLAY OF AIR TRAFFIC FOR THE COCKPIT (INFORMAL)* Michael W. McGreevy 514

3. *EFFECT OF VFR AIRCRAFT ON APPROACH TRAFFIC WITH AND WITHOUT COCKPIT DISPLAYS OF TRAFFIC INFORMATION (FORMAL)* Sandra G. Hart 522

4. *AIR TRAFFIC CONTROL OF SIMULATED AIRCRAFT WITH AND WITHOUT COCKPIT TRAFFIC DISPLAYS (FORMAL)* Sheryl L. Chappell, John G. Kreifeldt 545

LIST OF AUTHORS 553

SESSION 1: HUMAN OPERATOR MODELS

Moderator: David L. Kleinman, University of Connecticut

AN INFORMATION THEORETIC MODEL OF THE HUMAN OPERATOR

Charles P. Hatsell, Lt Col, USAF, MC*
Air Force Aerospace Medical Research Laboratory
Human Engineering Division
Wright-Patterson AFB OH

Patricia R. Mineer*

ABSTRACT

Application of information theory to the summing node in the manned system control loop reveals that the human controller must process two quantities of information flow, transinformation $I(X;Z)$ due to dependent input, X, and system state variables, Z; and transinformation $I(E;Z)$ due to dependent summing node output, E, and system state, Z. The measure of controller effectiveness is the entropy decrement, $H = H(X)-H(E)$, due to loop closure. An analysis of the summing node gives,

$$I(X;Z) - I(E;Z) = \Delta H \ . \qquad (1)$$

It is argued that for a highly motivated human operator whose performance on the subject task has plateaued, a task-dependent channel capacity, C_H, will be totally allocated to $I(X;Z)$ and $I(E;Z)$, i.e.,

$$I(X;Z) + I(E;Z) = C_Z \qquad (2)$$

Equations (1) and (2) comprise the information theoretic human operator model.

Model validation was accomplished on data from seven subjects performing a simple one-dimensional, purely compensatory, tracking task. The perturbing function was a 0.5 hz Gaussian random process and two plants were tracked, $1/s$ and $1/s(s+a)$. Agreement with the information theoretic model

*This work was performed while the authors were with the Crew Technology Division, School of Aerospace Medicine, Brooks AFB TX.

was excellent over all subjects and plants. Performance in terms of ΔH had high intersubject variability, reflecting differing $I(E;Z)$ among subjects; furthermore, channel capacity for the tracking task was invariant over subjects and plants. The $1/s(s+a)$ plant was subjectively more difficult to control and always resulted in a higher $I(E;Z)$.

The information-theoretic human operator model is valid regardless of system linearity, stationarity, or noise structure. The model is simply an information-theoretic statement of the summing node together with an experimentally verified conjecture about how the human operator handles information. Additional experiments should identify relationships among C_h, task-type, and exogenous stressors, with obvious application to workload quantification.

AD P001615

RISK AND DECISION PROCESSES IN MANUAL CONTROL BEHAVIOUR

TOM BÖSSER

Psychologisches Institut
Westfälische Wilhelms-Universität
Schlaunstr.2
D-4400 Münster

ABSTRACT

Manual control behaviour with RMS-error-criterion is considered a good approximation to control behaviour with similar symmetric error-weighting-functions, but not asymmetric ones. In order to test this, a compensatory tracking task with a second error condition is defined. Several experiments show that nonlinear (asymmetric) tracking-behaviour arises under these conditions. The adaptation of the control-behaviour to the payoff-conditions can be considered an instance of decision-making on the basis of an expected-utility model.

INTRODUCTION

The 'describing-function'-approach to manual control behaviour does not neccessarily require the quantification of the error according to an RMS-error-criterion (where the error is the sum of squares of the deviation of the system-output from target). However, the statistical methodology available (system identification with stochastic test-signals) would make it rather difficult to use a different criterion to describe performance of human subjects in manual control behaviour. Training in tracking-tasks is usually based on this error-criterion, and therefore subjects adapt to this property of the task. Experimental results can only be generalized to real tasks, when these have the same task-structure, and therefore the same laws may be assumed to govern behaviour under these conditions. It has been suggested that this may not be the case for some tasks, e.g. distance-control in driving a car (BÖSSER 1980).

The basic rationale behind the use of the RMS-error-criterion is that in most conceivable situations a larger deviation from the target is given an unproportionally larger weight than a smaller deviation, and the direction of the deviation does not matter. This view implies that the deviation of system-state from the set-point or target is weighted according to the utility (gain or loss) associated with this deviation. In Fig. 1 this is shown together with other conceivable error-criteria. An error weighted proportional to the deviation from the target may well arise e.g. in process-control, where a deviation from target may simply mean a loss of material proportional to this deviation. A discrete type of error-criterion is the fixed-limit criterion, where only the trans-

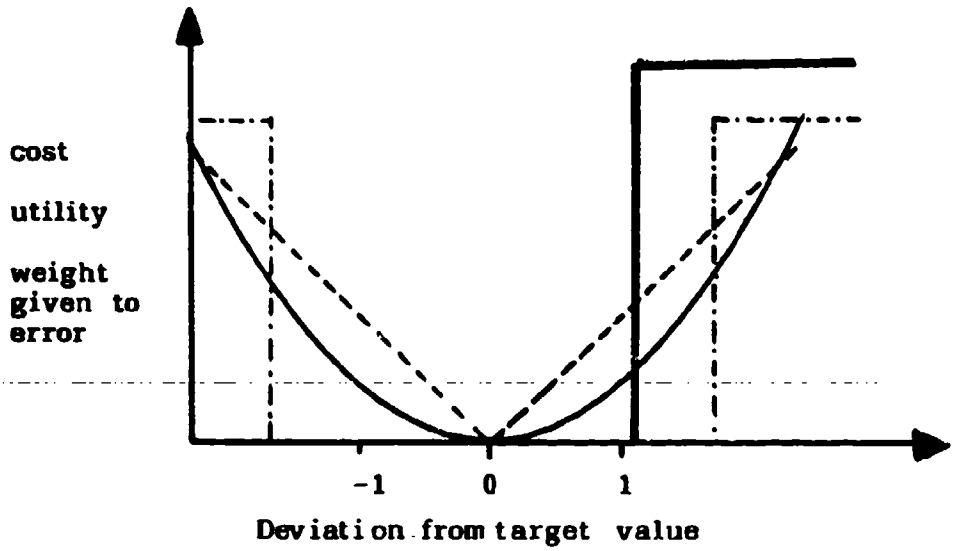


Fig.1: Error-weighting functions

gression of a borderline matters, and weight given to the error is 0 within limits, constant outside of limits.

As can be seen from Fig.1, an RMS-error-criterion does not deviate too much from a constant-weight-criterion or from a fixed-limit-criterion. Because human response is comparatively slow in most situations, it is difficult to respond like a perfect relay-type controller, this would give rise to a strategy where the human controller anticipates approaches to the borderline and infers a value (probability of being driven across the limits) which may be called 'danger', and which will give rise to compensatory action. The subjective weight given to these deviations from target, and the associated tendency to compensate this error may correspond closely to the RMS-type error-criterion.

These considerations suggest that the RMS-error-criterion is a reasonable approximation to various symmetric types of weighting-functions for error in manual control situations. It is quite obvious, however, that this does not apply to asymmetric weighting functions for deviations from the target value, like the fixed limit in one direction shown in Fig. 1. There are obvious examples for weighting functions of this type: Distance control in driving an automobile, where only approach, but not increasing distance from a leading car may have grave consequences, operation of machines of various kinds or the drill of the dentist have consequences, where transgression of a borderline, but not staying short of the line, has unproportionally large negative weight.

In order to investigate manual control behaviour in tasks with asymmetric weighting-function, a standard compensatory tracking task was augmented with a second error-condition represented by a borderline on the display (Fig.2). Crossing of the line is signalled by a buzzer. Subjects were required, in additon to keeping the error-marker close to the target, not to cross the borderline.

Previous experiments have shown that subjects have a strategy available to deal with this task, which involves two aspects: Performance in relation to the line-criterion is improved by moving the mean of the control-position away from the border-line, thus generating a larger RMS-error. (BÖSSER 1982). Secondly the tracking behaviour may be made more efficient in this task by giving it a dynamic asymmetry, which becomes larger with higher frequency of the track and with shorter distance d of the borderline. It was shown that subjects adapt these properties of their tracking behaviour in a systematic fashion depending on the experimental conditions.

Here several experiments are reported where the principles of the adaptation of the human operator in the tracking task with additional error-criterion are investigated. The adaptation described by McRUER & KRENDEL (1974) in their 'Successive Organization of Perception'-Model relates to learning processes, whereas the weighting of error-conditions or payoff-conditions belongs to the domain of motivation.

METHOD

A standard compensatory tracking-task was used, augmented by the borderline at distance d from the target. The disturbance (input-signal) was derived from a 50Hz 3-level Pseudo-Random-ternary signal (period 40. secs), which was choosen because of the known superior statistical properties for the identification of nonlinear transfer-characteristics. The original signal was filtered with a zero-phaseshift filter with various cutoff-frequencies. By addition of a number of these filtered input signals with equal variance were generated with good properties for the identification procedure used and which tested the subjects behaviour in a useful range. The spectra of the easiest and the most difficult signal are shown in Fig.3

A zero-order (position-) control-system with isotonic control-lever and a 15 by 24 cm oscilloscope-display was used, all control and data-collection was done on-line by computer.

The data were analyzed by calculating statistical parameters of the measured signals and by calculating the transfer-characteristics by FFT-methods. Nonlinear transfer-characteristics were investigated by calculating appropriate statistical parameters derived from WIENERs (1958) theory of nonlinear systems-identification. For a satisfactory estimation of parameters the amount of data available was rather small, consequently the estimates of the kernels are subject to much statistical variation. The cross-covariance calculated in the time-domain was judged to be most illustrative under these conditions and is presented here for various conditions.

EXPERIMENT I

An asymmetry in manual control behaviour may be caused by proper of the motor-system rather than higher-level control. In order to test this, the position of the borderline d was varied (above and below target), the orientation of control and display, and the frequency-composition of the input-signal (break-frequency 0.57 Hz and 0.14 Hz). Five subjects (psychology-students) participated. They were paid for participation in the experiment and were instructed to give equal weight to both error-conditions (deviation from target and crossing of the borderline). After several hours of training the experimental sessions 307 secs each were presented in random order.

Table 1 and 2 show the constant position error generated by the subjects, one part of the strategy subjects follow to prevent crossing the borderline. Correspondingly the RMS-error rises, and the number and duration of crossings of the borderline decrease. All subjects exhibit the same pattern of dependence of the performance-variables on experimental conditions. These results show clearly that the constant bias depends mainly on the position of the borderline above or below target. A higher frequency of the disturbance - the consequence of which is a higher probability of crossing the line - induces a larger bias. Reversal of the orientation of control-lever and display does not show a considerable influence, although most variables (some not shown here) seem to suggest that a compatible arrangement with borderline below target gives best performance.

The distributions of the error (Fig. 4) demonstrate the same relationship, the distributions exhibit a marked asymmetry. The distributions of the speed of the control-lever indicate the subjects strategy concerning the dynamics of the control-behaviour: The highest speeds of the control-lever movement occur in movements away from the borderline.

For three conditions (borderline above and below target and no borderline) the bi-crosscovariances between input-signal and remnant (linear proportion of the output variance subtracted) are shown (Fig.5) and demonstrate that the nonlinear proportion of tracking behaviour is dependent on the experimental conditions. The bi-crosscovariance (a) represents behaviour with the same input-signal as (b) and (c), but without additional error-condition (no borderline). The much higher variance due to introduction of the line-error-condition is evidence of the fact that subjects are able to adjust their control-strategy such as to include a specific nonlinear control strategy (asymmetric dynamics) which is represented by the second-order kernel of the WIENER-identification scheme. Orientation of the borderline, and thus the weight given to error-conditions, determines the direction of this asymmetry, as can be seen from the comparison of bi-crosscovariances (b) and (c).

Often it is argued that nonlinear transfer-characteristics may not be of great importance in manual control. We do not agree with this argumentation, because even small proportions of the variance observed

may represent particularly efficient behaviour in extreme conditions. Also the total variance may not be an adequate reference-value when the variance accounted for by the nonlinear transfer-characteristics is in the cross-over region, where it may account for a larger proportion of variance.

From these results it can be concluded that motivation - weight given to various conditions which are the outcome of actions - governs the principles of the control-strategy used. If nonlinear behaviour can be induced, then also linear behaviour is a function of the specific experimental conditions. Further experiments are directed towards identifying the rules which govern the selection and preference of the strategies available to the human operator.

EXPERIMENT II

Previous experiments (BÖSSER 1982) have shown that control behavior in the experimental situation investigated is adapted depending on priority given to either of the two error-conditions. In this experiment adaptation to various mixtures of weight given to the two error-conditions were investigated. Data are shown from one very well trained and highly motivated subject. Experimental conditions were the same as described above. There was just one type of input-signal and on distance of the borderline in all parts of the experiment. On four consecutive days the subject was asked to give different weight to the two error-conditions (RMS-error and line-crossing) in nine daily experimental sessions, from 10/90% to 90/10%. The subject was able to follow this 'fuzzy' instruction well.

On the first two days the subject was asked to perform at his best, on the third day he was asked to do the task in such way as to make it 'easy' for him. Finally on the last day an additional task (counting car in a TV-film) was given to the subject, in addition the radio was turned on loud and various other types of disturbance were generated.

The results are presented in the form of a performance-operating-curve: On the assumption that the total ressources allocated at any one time are constant, this curve represents the relative proportion of the ressources available to the subject which are allocated to the two dimensions of the task. From Fig.6 the extremely orderly curves of this subject's perfomance can be seen, in view of the unspecific instruction even more remarkable. Although this subject was outstanding in his performance as compared to others, this result demonstrates clearly that within a wide range performance in the task investigated may be adapted in a continuous manner to the specific requirements imposed. For other subjects more explicit instructions and longer training may be required to achieve the same result.

The obvious question to be asked next is: What are the conditions and rules governing the adaptation of control-behaviour to payoff-conditions?

EXPERIMENT III

In all previous experiments subjects were shown to be able to vary their tracking behaviour in response to explicit instructions to give priority to either of the two error-conditions defined in our task, or to weight these error-conditions according to certain ratios. In comparison real tasks would be expected to have a larger number of dimensions relevant for the evaluation of the system-performance of the complete man-machine-system. The human operator weights the states of the system according to weighting-functions which relate states of the system to a utility associated with these states. On this basis certain states are preferable to others, and the human operator adapts his performance within the range of his capabilities in such a way as to optimize the total outcome of his actions. These assumptions are largely of an axiomatic nature and not empirically testable.

In order to predict behaviour at selecting between different behavioural strategies, we have to know the following:
* What are the relevant dimensions of this utility-space, and what are the weighting-functions associated with the states of the system, and
* What are the capabilities of the human to extract the information neccessary to choose the correct strategy? Where the environment of the system consists of random-variables, the estimation of probability-distributions is of major importance.

An experiment was designed to test the general applicability of suc a model. In the experimental situation described above one of the task variables, frequency-spectrum of the input-signal, was variied. At higher frequency of the disturbance the probability to cross the borderline becomes larger, and corrective action (moving average value further away from the line) generates more error on the other error-criterion. Subjects were paid according to their performance, and the second variable which was variied experimentally were the payoff-conditions.

Payment for the subjects was based on a rule of the following kind:

$$\text{earnings} = a \cdot (\text{RMS-error}) - b \cdot (\text{number of line-crossings})$$

The amount deducted for line-crossings (b) was varied from 0 to .1 DM and .50 DM per line-crossing. in condition 4 one line-crossing was free, the second one was followed by total loss of the trial's earnings. Subjects were not told the payoff-conditions, but were informed of the result immediately following each trial. Gains and losses were accumulated and the total amount paid at the end of the experiment.

Subjects were trained extensively without payment under all the experimental conditions. Experiments were conducted on four conscutive days, cost of borderline-transgressions increasing. On each day each subject did twelve individual trials of 307 secs each, four with the same input-signal in random order. After each trial of 307 secs the subject was informed about the amount of money gained or lost in that trial.

Results are summarized in Fig.7: In response to lower gains, subjects adopt a control-strategy where they reduce the number of linecrossings at the expense of a larger RMS-error, and also a larger mean-deviation from target. Data show the same relationship for all subjects. Under the experimental conditions induced in this experiment, we can be sure that the subjects try to maximize earnings. The results quite clearly support the hypothesis that the human controller adapts his control-strategy such as to maximize the utility and to minimize expected cost.

DISCUSSION

In the experiments described it was shown that the human operator has available to him a nonlinear control-strategy which may be used when the task is structured such as to require this property. The WIENER-method of nonlinear-system-identification is suited to the description of this property of tracking-behaviour. The dimensions of performance are not independent of each other, they are limited by the number of behaviour alternatives available to the human controller.

The behaviour in the situation investigated may also be regarded a behaviour involving risk-taking and decision-making. Decision-experiments are carried out mostly with clearly stated risks (probability of certain resulting states multiplied by value of the expected outcome). These are not the situations typical for real life: Distance-control as part of the driving-task is evaluated according to the following (and others) dimensions of the utility-space: Speed (loss of time), operating costs, comfort and probability of accident. The expected values involved can be regarded as the product of the expected frequency-distribution of states of the system, given a certain control-strategy, multiplied by an appropriate weighting-function. For each behaviour-alternative (strategy) an expected utility may be calculated in this way.

The expected utility EV_i for behaviour alternatives $a_1, a_2, ..a_i$ and dimensions of the utility-space $z_1, z_2, ...z_j$ will be

$$EV_i = p_{i1} \cdot u_{i1} + p_{i2} \cdot u_{i2} + \cdots + p_{ij} \cdot u_{ij}$$

where p_{ij} are the probability-distributions expected for the future state of the system, and u_{ij} are the utilities associated with these states.

The set of states that may be reached is limited by the possible behaviour alternatives the human is capable of. He has to choose between the possible strategies, as exemplified for example by Fig.6. Does the human optimize according to the hypothesis outlined? In the estimation of the probability-distributions p much error will be involved, but this does not mean that the human does not select the best alternative on the basis of his faulty knowlege. The problem is rather the determination of the number of dimensions of the utility-space and the weighting-functions associated with the states of the system. (Motivation seems to be exhaustively described this way.) If fixed weighting-functions according to external criteria are assumed, the decision-rules become complex, when

it is assumed axiomatically that the human does select an optimal behaviour strategy, the determination of the weighting functions becomes difficult because of the large number of dimensions involved. The latter point of view seems to be preferable.

Of particular interest is the case, where one of the negative utilities involved is the risk of accidents. Results of LICHTENSTEIN et al. (1978) indicate that the associated probabilities can not be estimated correctly in the case of rare events. The same may apply to the estimation of the cost associated, which may be unrealistic before exposure to a grave accident in reality. According to the point of view advocated here, the negative utility of accidents can be compensated by gains in other dimensions of the utility-space, or, to state this more profanely, risk of accident will be accepted when there is a corresponding gain in some other dimension of the utility-space.

We conclude that linear manual control behaviour is a special case restricted to certain error-weighting-conditions. If also asymmetric weighting-functions are considered - some of them, like the tendency to avoid accidents, can be assumed to exist - nonlinear control-strategies arise and the resulting behaviour can be regarded as the result of a decision-process to select an optimal behaviour-strategy.

LITERATURE

BÖSSER, T. Effects of motivation on car-following. Proc. 16. Annual Conference Manual Control, MIT, Cambridge, Mass. 1980.

BÖSSER, T. Tracking performance with varying error-criteria. IFAC Conference on Analysis, Design and Evaluation of man-machine systems. Baden-Baden 1982.

LICHTENSTEIN,S., SLOVIC,P., FISCHHOFF,B., LAYMAN,M & COOMBS,B. Judge frequency of lethal events. J. Experimental Psychology: Human Learning and Memory,4,1978, 551-578.

McRUER,D.T. & KRENDEL,E.S. Mathematical models of human pilot behaviour. AGARDograph No. 188. AGARD-AG-188. London 1974.

WIENER,N. Nonlinear Problems in Random Theory. Wiley, New York. 1958.

This work was supported by Deutsche Forschungsgemeinschaft. Elke Melchior, Peter Schäfer and Martin Schütte contributed to this research.

TABLE 1

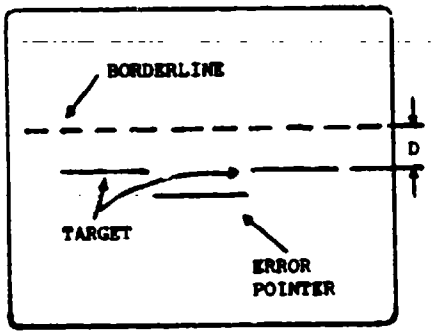
BORDER-LINE / CONTROL	ABOVE TARGET	BELOW TARGET
COMPATIBLE	- 0.25	0.16
NON-COMPATIBLE	- 0.19	0.26

CONSTANT POSITION ERROR (CM)
BREAK-FREQUENCY 0.14 HZ, DISTANCE OF BORDERLINE 0.5 CM
N = 5 SUBJECTS
COMPATIBLE: SHIFTING CONTROL AWAY FROM SUBJECT RAISES ERROR-POINTER

TABLE 2

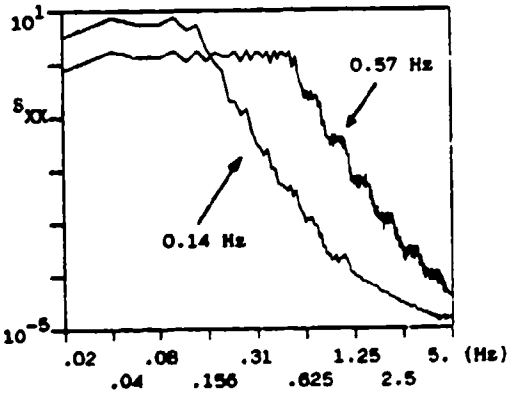
BORDER-LINE / CONTROL	ABOVE TARGET	BELOW TARGET
COMPATIBLE	- 1.47	0.86
NON-COMPATIBLE	- 1.47	1.07

CONSTANT POSITION ERROR (CM)
BREAK-FREQUENCY 0.57 HZ, DISTANCE OF BORDERLINE [illegible].
N = 5 SUBJECTS
COMPATIBLE: SHIFTING CONTROL AWAY FROM SUBJECT RAISES ERROR-POINTER

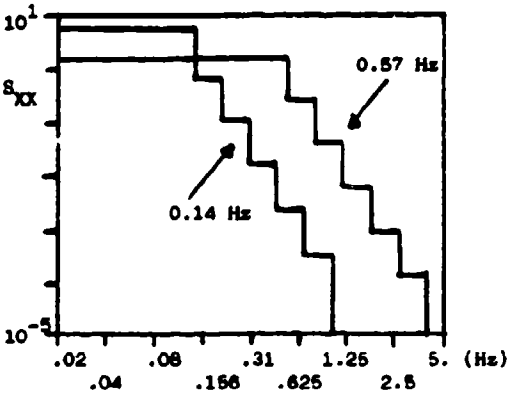


COMPENSATORY DISPLAY USED IN TRACKING EXPERIMENTS

Fig.2

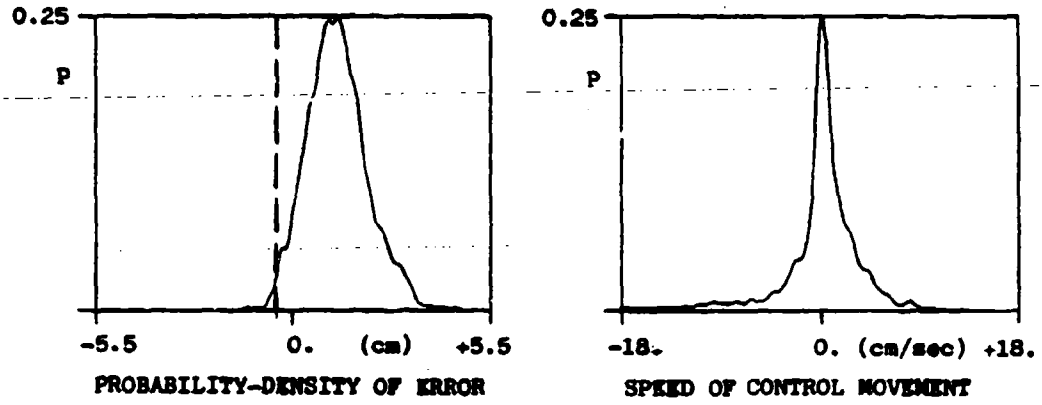


POWER SPECTRA S_{XX} OF TRACK, BREAK-FREQUENCY (-3db) 0.14 AND 0.57 HZ

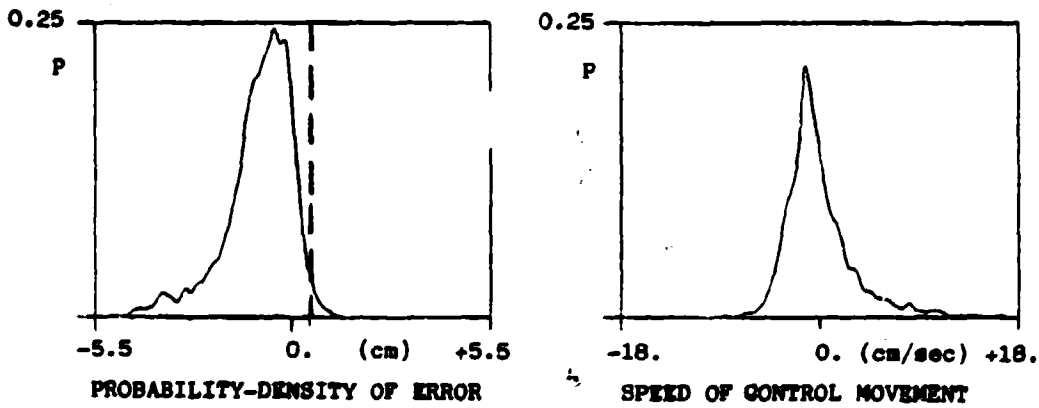


"IDEAL" SPECTRA S_{XX} OF TRACK, BREAK-FREQUENCY (-3db) 0.14 AND 0.57 HZ

Fig.3 Spectra of input-signals, easiest and most difficult signals.



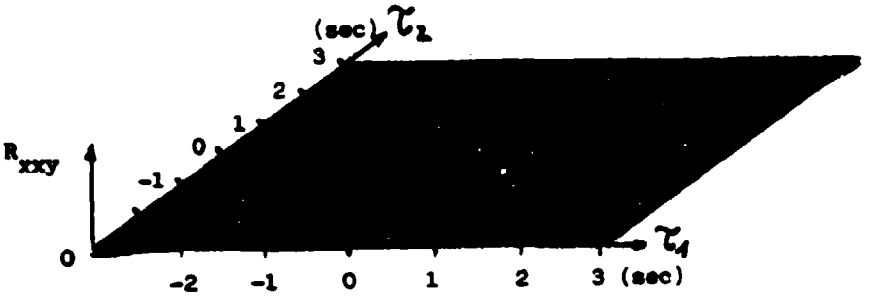
BREAK-FREQUENCY 0.57 Hz
DISTANCE OF BORDERLINE 0.5 CM ABOVE TARGET
N = 1 SUBJECT



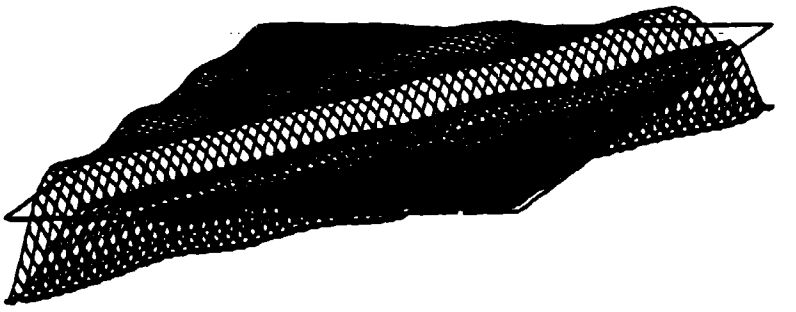
BREAK-FREQUENCY 0.57 Hz
DISTANCE OF BORDERLINE 0.5 CM BELOW TARGET
N = 1 SUBJECT

FIG. 4

(a)



(b)



(c)

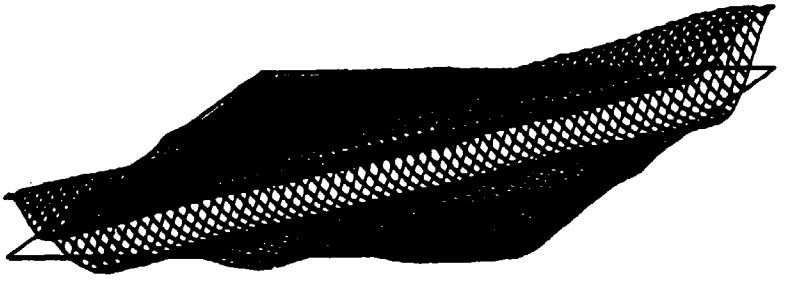


FIG. 5 SECOND ORDER CROSS-COVARIANCE

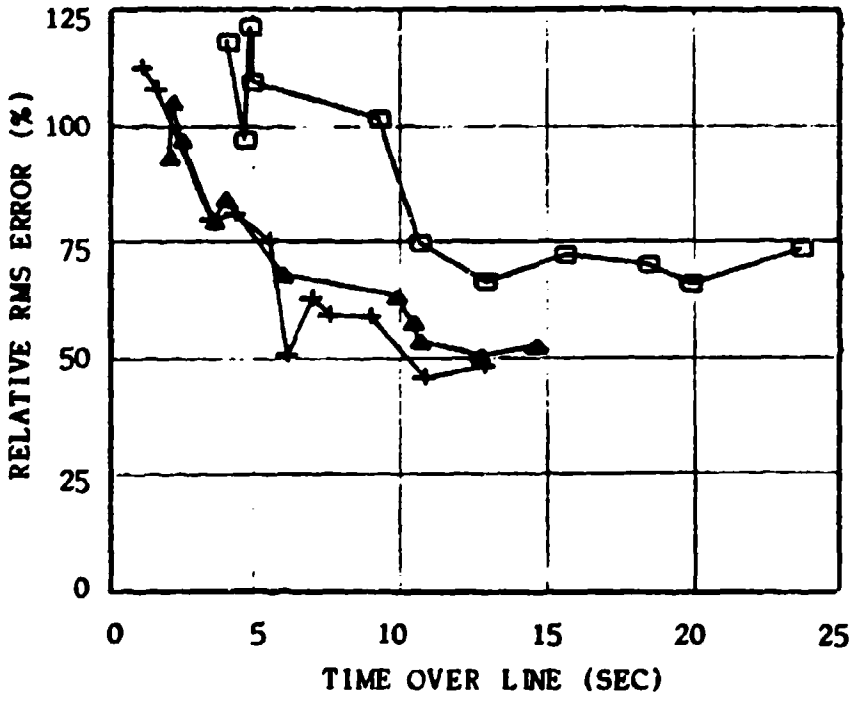
$$R_{\tau_1,\tau_2} = 1/n \sum_{i=1}^{n} x_{i-\tau_1} \cdot x_{i-\tau_2} \cdot z_i$$

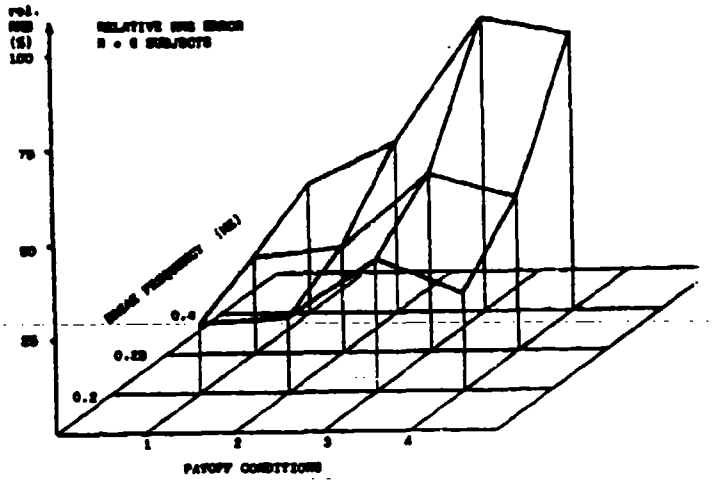
(z = remnant)

(a) NO LINE
(b) DISTANCE OF BORDERLINE 0.5 CM ABOVE TARGET
(c) DISTANCE OF BORDERLINE 0.5 CM BELOW TARGET
BREAK-FREQUENCY 0.57 HZ
N = 1 subject

FIG.6

PERFORMANCE OPERATING CURVE
TASK: + MAXIMUM EFFORT
▲ EASY
□ WITH ADDITIONAL TASK
SUBJECT J.G.




0.4
0.2
1
2
3
4
PAYOFF CONDITIONS

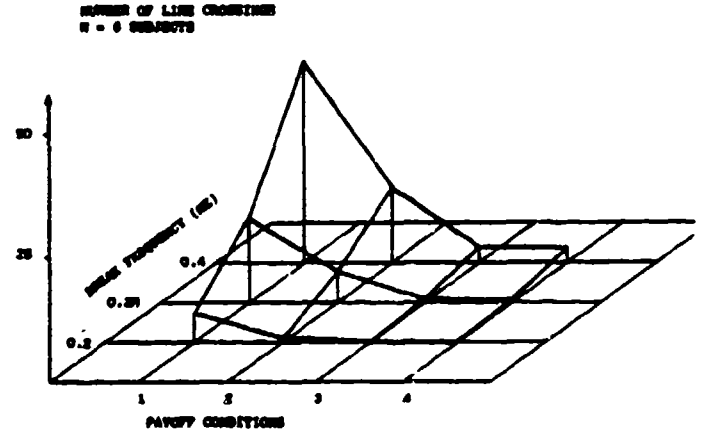
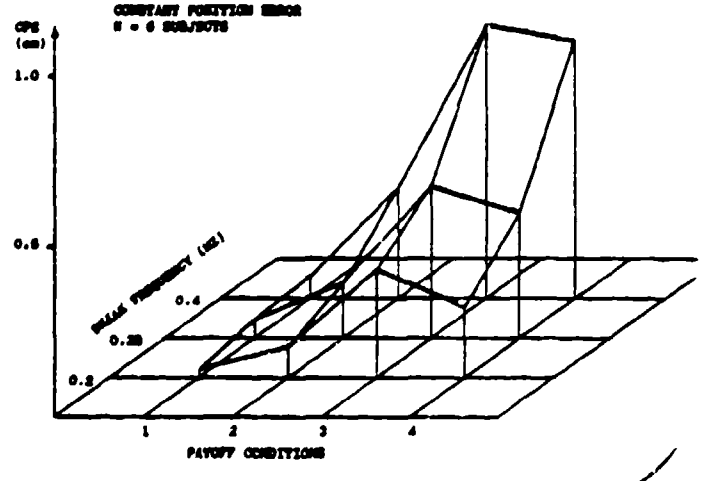

0.4
0.2
1
2
3
4
PAYOFF CONDITIONS

FIG..7


1.0
0.4
0.2
1
2
3
4
PAYOFF CONDITIONS

AD P001616

TIME DOMAIN IDENTIFICATION OF PILOT DYNAMICS AND CONTROL STRATEGY

by

David K. Schmidt
School of Aeronautics and Astronautics
Purdue University
West Lafayette, IN 47907

Abstract

In this paper a method is proposed for the identification of the pilot's control compensation using time domain techniques. From this information we hope to infer a quadratic cost function, supported by the data, that represents a reasonable expression for the pilot's control objective in the task being performed, or an inferred piloting 'strategy'. (Note: here that we are using the term 'strategy' as synonomous with control objective, and not with control law.) The ultimate goals of this research topic include a better understanding of the fundamental piloting techniques in complex tasks, such as landing approach; the development of a metric measurable in simulations and flight test that correlate with subjective pilot opinion; and to further validate pilot models and pilot-vehicle analysis methods. To be presented along with the methodology will be some preliminary numerical results.

Note, this paper was presented at the AFFTC/NASA Dryden/AIAA Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics, Edwards AFB, CA 93523, January 19-21,1982.

Introduction

In this paper, we will propose a method for the identification of the pilot's control compensation using time domain techniques. From this information we hope to infer a quadratic cost function, supported by the data, that represents a reasonable expression for the pilot's control objective in the task being performed, or an inferred piloting "strategy". (Note here that we are using the term strategy as synonomous with control objective, and not with control law.)

The ultimate goals of this research topic include a better understanding of the fundamental piloting techniques in complex tasks, such as landing approach; the development of a metric measurable in simulations and flight test that correlate with subjective pilot opinion; and to further validate pilot models and pilot-vehicle analysis methods. At this time we will present the methodology and some preliminary numerical results.

The Pilot Model and Objective Function

The analyses relies on the well-known [1] optimal-control theoretic technique for modeling the human pilot's manual control function. The hypothesis upon which it is based is that the well trained, well motivated pilot chooses his control inputs (e.g. stick force) to meet the pilot's (internal) objective in the task, subject to his human limitations. His objective is further assumed to be expressible in terms of a quadratic "cost" function

$$J_p = E\left\{\lim_{T\to\infty} \frac{1}{T}\int_0^T \left(Y_p^T Q Y_p + u_p^T R u_p + \dot{u}_p^T G \dot{u}_p\right) dt\right\} \qquad (0)$$

where Y_p = vector of pilot's observed variables (e.g., attitude, acceleration)

u_p = vector of pilot's control inputs

Q,R,G = Pilot-Selected (internal) weightings

The human limitations modeled include information-acquisition and processing time delay, observation and control input errors, and neuromuscular dynamics. A block diagram of the resulting model structure is shown in Figure 1.

The components of this model may be grouped into two parts, one dealing with the information acquisition and state estimation, and one related to the control law or control policy operating on the estimated state. As has been shown in the references on this modeling approach, the "solution" for the pilot's control inputs, as predicted by the model, is expressed as

$$\dot{u}_p = g_x^T \hat{x} + g_u^T u_p + \nu_u$$

where $\hat{x}$ = internal estimate of the system state

g_x, g_u = control gains

ν_u = motor noise, or control input errors

(Readers unfamiliar with the further details of the model are referred to the reference.)

The key points germain to this analysis are that the above equation is a mathematical expression representing the pilot's overt control actions (u_p), and these control actions are measurable experimentally. Furthermore, the gains g_x and g_u are functions of the plant (vehicle) dynamics and his objective function, and thereby represent his control "techniques", level of skill, and familiarity with the vehicle dynamics.

Another factor of importance is that not only is the objective function, from which the gains are determined, a mathematical part of a pilot control model, but it's resulting magnitude obtained from exercising the model has been found to correlate with the subjective pilot opinion obtained from simulation and flight test. Such a correlation is shown in Figure 2, as an example, taken from Refs. 2 and 3. This of course assumes one has been able

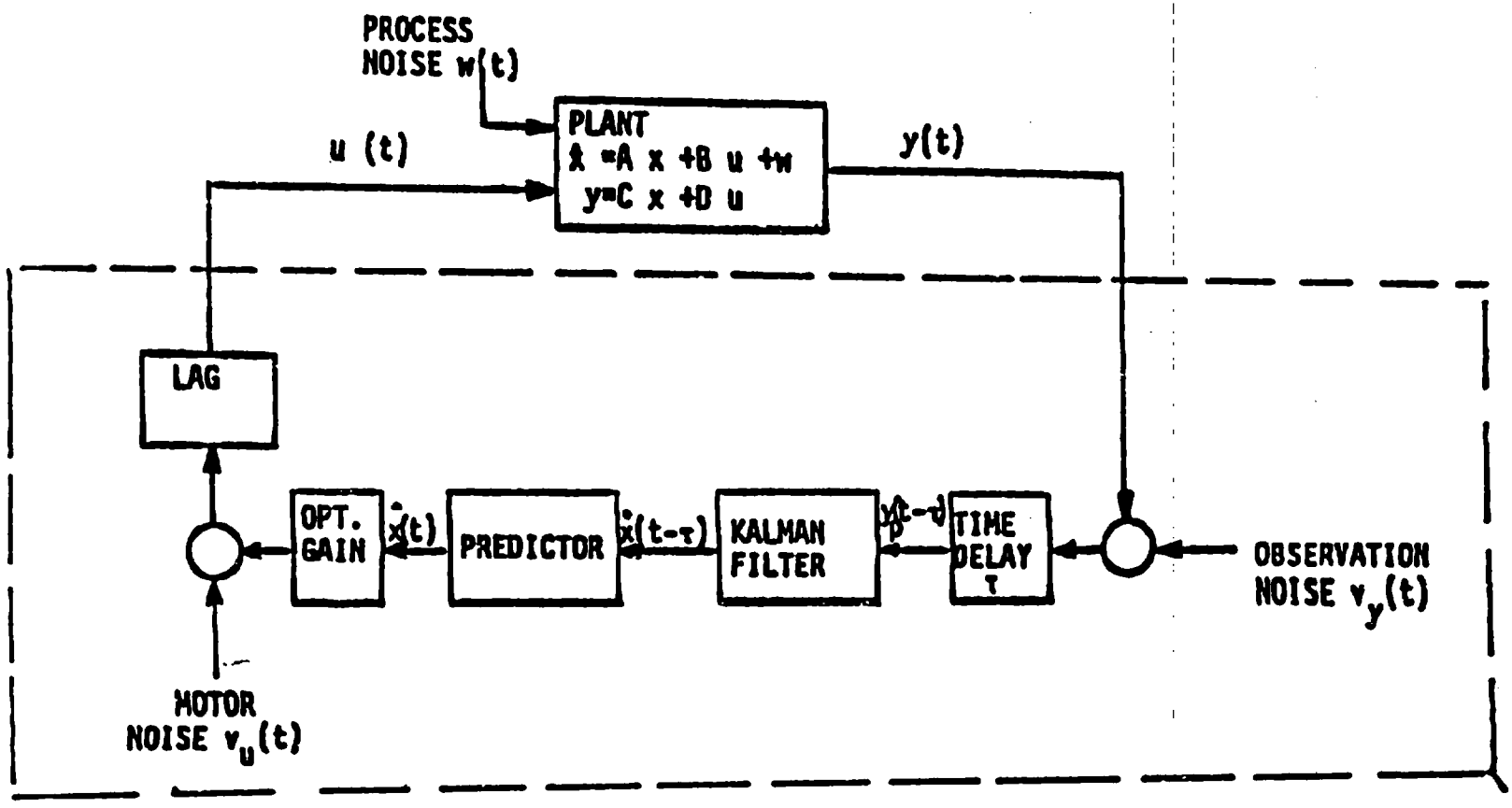


Figure 1. Model Structure

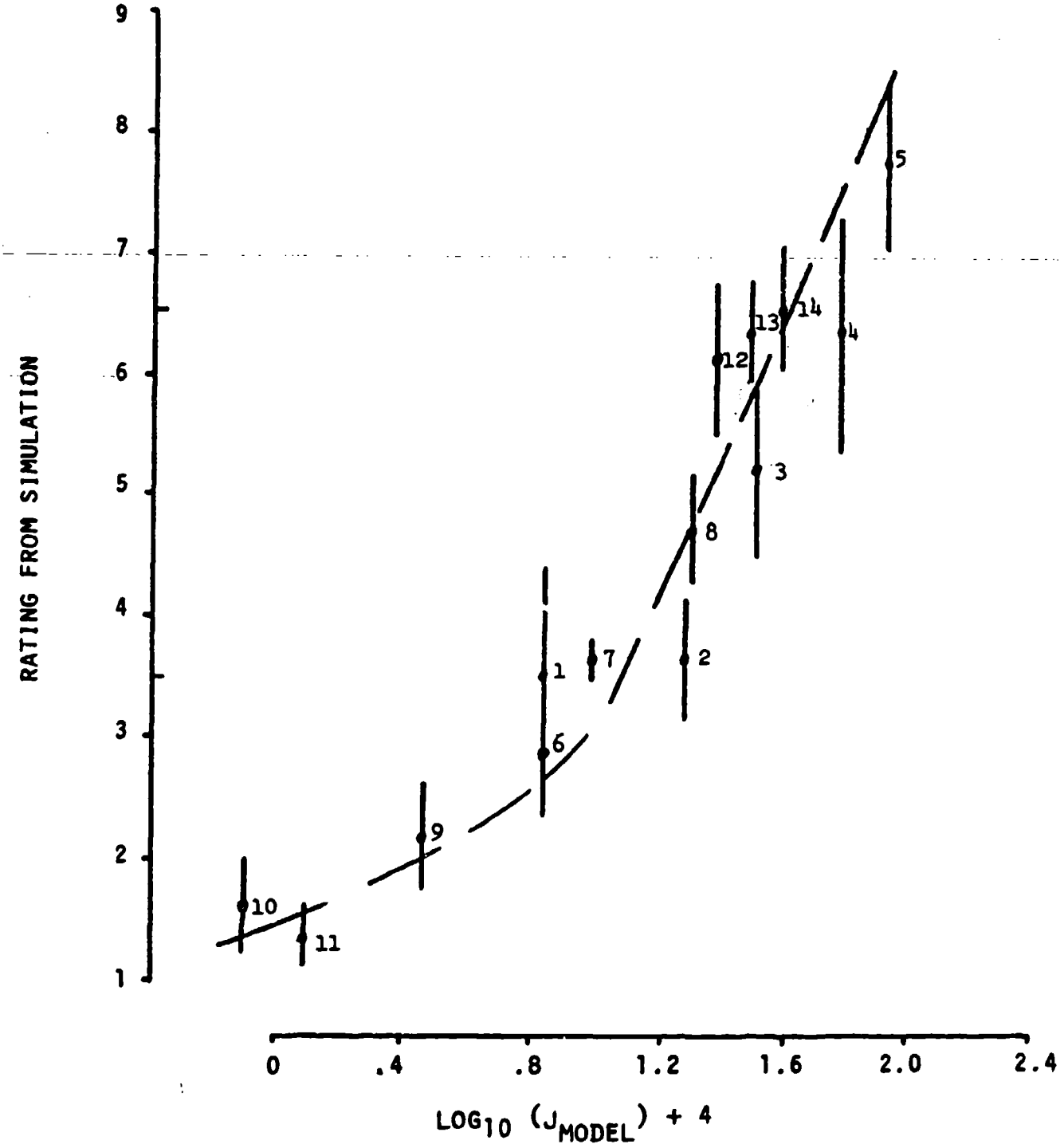


Figure 2, Rating Correlation

to correctly express the pilot's (internal) cost function, which is in fact his strategy that depends on his perception of the task. Now this is easy to do in simple laboratory tasks in which the subject has been instructed to minimize some displayed error, for example. But it is not at all clear just what flight parameters are being "regulated" or "tracked", other than ILS glide slope and localizer error in the case of landing approach. This is but one example, other complex piloting tasks might be considered equally as well.

The Identification Procedure

We seek then a method by which we may identify those pilot parameters that reflect his control techniques, or control strategy. Referring back to the pilot model control law, or

$$\dot{u}_p = g_x^T \hat{x} + g_u^T u_p + \nu_u$$

we note that the gains g_x operate on the estimated state $\hat{x}$. Now the separation principle of optimal estimation and control theory states that the control gains (g_x, g_u) are independent of the state estimation process. Further, the optimal state estimator, in general and in the pilot model, is independent of the overall objective function being minimized by the controller (estimator and control) law. Therefore, if we are mainly after the pilot's control strategy as expressed by, or at least a function of, his objective function, we need only to focus on the gains (g_x, g_u) and not on those variables related only to the state estimator. These latter variables include the time delay, and observation and motor noise covariance matrices, parameters of interest in the identification technique of Levison [4], for example. If our approach is successful, fewer parameters must be identified from the data, which is always an advantage, but the parameters affecting the estimation process are assumed.

The identification method proposed is as follows. The control law expressed previously, may be rewritten as

$$\dot{u}_p = g_x^T x - g_x^T \epsilon + g_u^T u_p + \nu_u$$

where ϵ = error in estimating the true (actual) state x. Note that along with the pilot's control u_p, these true states, such as angle of attack or pitch attitude are measureable, but the state estimate, $\hat{x}$, is a quantity internal in to pilot, as modeled. Hence $\hat{x}$ is not measurable ---nor are ϵ or ν_u. Transposing the above, multiplying by $x = \text{col}\left[x, u_p\right]$, and taking expected values yields

$$\begin{bmatrix} E(x\dot{u}_p^T) \\ \hline E(u_p\dot{u}_p^T) \end{bmatrix} = \left\{ \left[\begin{array}{c|c} E(xx^T) & E(xu_p^T) \\ \hline E(u_p x^T) & E(u_p u_p^T) \end{array}\right] - \left[\begin{array}{c|c} E(x\epsilon^T) & 0 \\ \hline E(u_p\epsilon^T) & 0 \end{array}\right] \right\} \begin{bmatrix} g_x \\ \hline g_u \end{bmatrix}$$

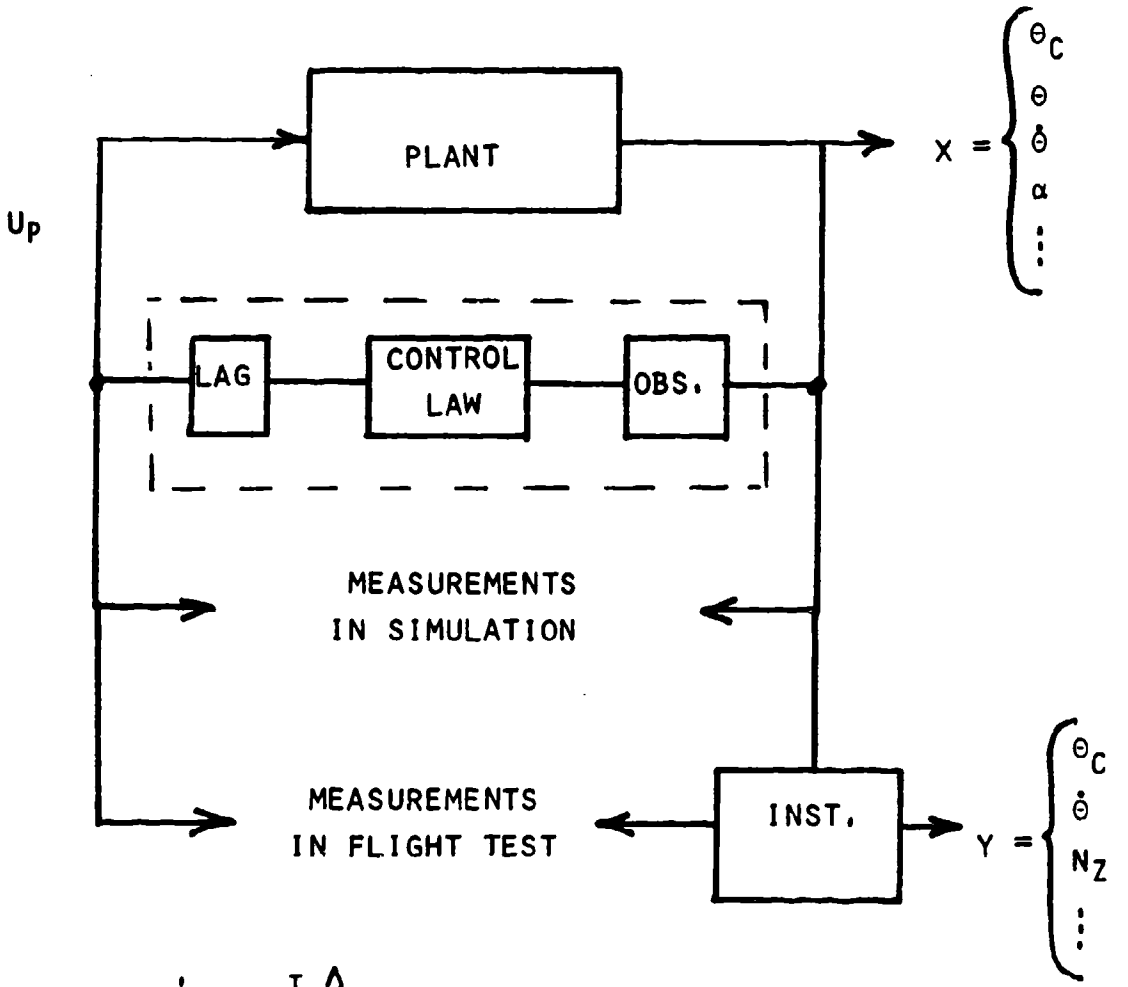
$$+ \begin{bmatrix} E(x\nu_u^T) \\ \hline E(u_p\nu_u^T) \end{bmatrix}$$

$$\text{or } N_{\dot{u}} = M \begin{bmatrix} g_x \\ \hline g_u \end{bmatrix} + N_{\nu_u} \qquad \text{(I)}$$

Now to evaluate these matrices we note first that, in a simulation at least, the vectors x(t) and $u_p(t)$ are measurable, so estimates of their covariance matrices (e.g., $E(xx^T)$) may be obtained from measurements of sampled time histories. (Also, in this paper we assume that good estimates of $\dot{u}_p$ are available from filtered measurements of u_p. The details of accomplishing this filtering are under current investigation, but digital techniques as well as analog methods are still available.) For reference, refer to Figure 3.

EXPERIMENTAL CONSIDERATIONS

TRACKING TASK - ATTITUDE, ACCEL. ...



$$\dot{U}_p = G_X^T \hat{X}_p + G_U U_p + V_U$$

Figure 3

With regard to the remaining terms involving ε and ν_u, both are not measurable and need attention. To resolve this consider the complete system dynamics model by the relation

$$\dot{x} = Ax + Bu_p + w$$

and

$$\dot{u}_p = g_x^T \dot{x} - g_x^T \varepsilon + g_u^T u_p + \nu_u$$

where the relation between state and estimate, or $\hat{x} = x - \varepsilon$ has been employed. The pilot's internal state estimation error, ε, is treated as follows. Define $\varepsilon_u = u_p - \hat{u}_p$ to be the error in estimation of the pilot's own control input, and then let

$$\bar{\varepsilon} = \text{col}\left[\varepsilon,\ \varepsilon_u\right]$$

Now the covariance of $\bar{\varepsilon}$ may be shown to be governed by the relation

$$\text{cov}\ (\bar{\varepsilon}) = E(\bar{\varepsilon}\ \bar{\varepsilon}^T) \triangleq P$$

$$\dot{P} = A_1 P + PA_1^T + W_1$$

Also we have

$$A_1 \Sigma + \Sigma A_1^T + W_1 - \Sigma C^T V_y^{-1} C\Sigma = 0;\ \Sigma = \text{cov}\ (e_{KF})$$

and

$$A_1 = \left[\begin{array}{c|c} A & B \\ \hline 0 & -\frac{1}{\tau} \end{array}\right] \qquad \begin{array}{l} C = \text{pilot's observation matrix} \\ Y_p = C\begin{bmatrix} x(t-\tau) \\ u(t-\tau) \end{bmatrix} + \nu_y \end{array}$$

These relations are all obtained from Ref. (5) and from the pilot model equations given in Ref. (1). Here e_{KF} is the Kalman filter estimation error for the delayed state, Σ the covariance of e_{KF}, and

$$W_1 = \left[\begin{array}{c|c} W & 0 \\ \hline 0 & V_u \end{array}\right]$$

Also W is the covariance of the plant disturbance w, and V_u and V_y are motor noise and measurement noise covariance, respectively, all assumed known. Now the $\dot{P}$ equation may be integrated over the time delay τ, with the initial condition on P from $P(0) = \Sigma$, the Kalman filter error covariance. Now, since the predictor has the property that $E(\hat{x}\,\bar{\varepsilon}^T) = 0$, we have

$$E\left\{\left[\begin{array}{c} x \\ \hline u_p \end{array}\right]\bar{\varepsilon}^T\right\} = E(\bar{\varepsilon}\,\bar{\varepsilon}^T) = P$$

So then the terms $E(x\,\varepsilon^T)$ and $E(u_p\varepsilon^T)$ are available from P, and these are required to form M.

Finally, it can be shown (Ref. (5)), pg. 331) that with the processes w and ν_u uncorrelated we have in this case

$$E(x\nu_u^T) = 0$$
$$E(u_p\nu_u^T) = \frac{1}{2}V_u$$

Returning then to the estimation of the gains (equation I), we see that all the terms in the matrices $N_{\dot{u}}$, N_{ν_u} and M may be calculated, either analytically or from the measurements of x, u_p (and $\dot{u}_p$). The estimate for the gain vector is then

$$\left[\begin{array}{c} g_x \\ \hline g_u \end{array}\right]_{est} = M^{-1}\left[N_{\dot{u}} - N_{\nu_u}\right] \qquad \text{(II)}$$

Note finally that the matrix M is formed from two matrices

$$M = M_x - M_{cor}$$

where the M_{cor} and N_{ν_u} matrices may be thought of as corrections added to a basic least-squares technique. The potential importance of these terms (M_{cor} and N_{ν_u}) will be demonstrated in an example later.

The algorithm is as follows:

1) Select noise covariance matrices, W, V_u, and V_y

2) Select a time delay τ, neuromuscular time constant τ_N (or matrix $T_N = g_u^{-1}$).

3) Form A_1 and solve for Kalman Filter error covariance Σ.

4) Solve for covariance matrix $P(\tau)$ and then the $E(\bar{\varepsilon}\,\bar{\varepsilon}^T)$ is available. (Note, all these steps may be accomplished before or after the experimental data is obtained.)

5) Perform experiment to obtain state and control (and control rate) time histories.

6) From the time histories, obtain estimates for $E(xx^T)$, $E(xu_p^T)$, $E(u_p u_p^T)$, $E(x\dot{u}_p^T)$ and $E(u_p \dot{u}_p^T)$, or the matrices M_x and N_u.

7) Identify M_{cor} and N_{v_u} in $E(\bar{\varepsilon}\,\bar{\varepsilon}^T)$ found in step 4.

8) Form $M = M_x - M_{cor}$ and determing $\left[\frac{g_x}{g_u}\right]_{est}$ from Equation II.

9) Check g_u vs T_N^{-1} (selected in 2 above) and iterate (steps 2-8) again as necessary. Note now that selecting τ_N affects the solution for Σ and $P(\tau)$, along with the effective V_u or

$$V_{u_{eff}} = T_N^{-1} V_u \left(T_N^{-1}\right)^T$$

while selecting τ only affects $P(\tau)$ in the procedure.

Comparison to Classical Results

It is interesting to note that the "corrections" performed by including M_{cor} and N_{v_u} are qualitatively related to an identification technique (discussed in Ref. 6, and elsewhere) used to determine the human describing

function in a compensatory task, which goes back to the development of the "crossover model" of McGruer et al. Shown in Figure 4 is a schematic of this situation, showing the closed-loop tracking of some commanded θ_c. Measurements may be taken of $\theta_c(t)$, $\varepsilon(t)$, $u_p(t)$, and $\theta(t)$ and manipulated in the frequency domain to obtain frequency spectra

$$G_1(j\omega) = U_p(j\omega)/\theta_c(j\omega)$$
$$G_2(j\omega) = \varepsilon(j\omega)/\theta_c(j\omega)$$
$$G_3(j\omega) = U_p(j\omega)/\varepsilon(j\omega)$$

Now, in this model the pilot's control is considered to consist of two parts, one correlated with the input θ_c, the other uncorrelated with the input. The latter component was defined to be "remnant." Mathematically,

$$u_p(j\omega) = Y_p(j\omega)\ \varepsilon(j\omega) + r(j\omega) \text{ and } r(j\omega)/\theta_c(j\omega) \to 0 \text{ in effect.}$$

Block diagram manipulation leads then to the desired relation

$$Y_p(j\omega) = G_1(j\omega)/G_2(j\omega)$$

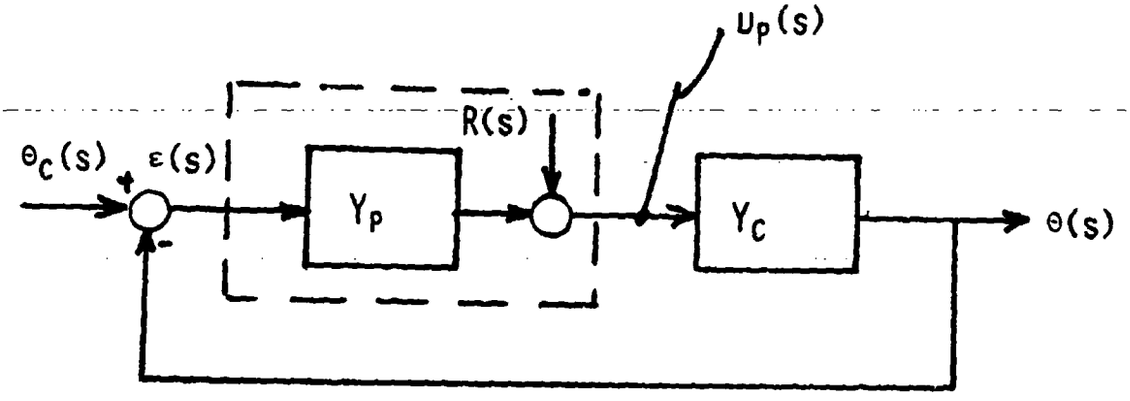
rather than the simpler, and incorrect, expression $Y_p(j\omega) = G_3(j\omega)$. This was due to the presence of remnant $r(j\omega)$ in the <u>measured</u> control input, and the necessity to eliminate it's effect by defining it as the uncorrelated component of u_p, and using this property. Comparing to our control law, transformed just for discussion purposes, we have

$$\dot{u}_p(j\omega) = \underbrace{g_x^T\, x(j\omega) + g_u^T\, u_p(j\omega)} - \underbrace{g_x^T\, \bar{\varepsilon}(j\omega) + \nu_u(j\omega)}_{\substack{\text{unmeasurable} \\ \text{separately}}}$$

compared to

$$u_p(j\omega) = Y_p(j\omega)\ \varepsilon(j\omega) + \underbrace{r(j\omega)}_{\substack{\text{unmeasurable} \\ \text{separately}}}$$

CLASSICAL RESULTS



MEASURE $U_p(s) = Y_p\ \varepsilon(s) + R(s)$

$$Y_p = G_1(JW)/G_2(JW) \neq G_3(JW)$$

WHERE $G_1(s) = U_p(s)/\Theta_c(s)$

$G_2(s) = \varepsilon(s)/\Theta_c(s)$

$G_3(s) = U_p(s)/\varepsilon(s)$

Figure 4

The significant difference is that $r(j\omega)$ was, in effect, discarded in finding Y_p, but $g_x^T \epsilon$ is not uncorrelated with x or u_p and must be accounted for in the identification problem.

A Numerical Example

To evaluate the numerical properties and the sensitivity to the a priori selected parameters (V_y, W, V_u, τ) a fast time simulation of the pilot model equations has been assembled, and the simulated control task is shown in Figure 5 . As shown, the task is that of pursuit tracking with $11.7/s^2$ controlled element dynamics, and the displayed command signal is filtered white noise with the filter transfer function given ($\theta_c(s)/w(s)$). The state vector is shown, the known gain vector to be identified is listed, and the weights in the objective function used are given. A sample time history of the state and simulated pilot's control input is depicted in Figure 6. Such time histories were sampled at 10 msec intervals and the gains estimated from time windows of data 5, 10, 15, 20, 25, 30 and 35 seconds wide. The root-sum-squred percent error of the five estimated gains is shown in Figure 7. Where

$$E_{RSS} = \left[\sum_{i=1}^{5} \left[\frac{g_i - \hat{g}_i}{g_i} \right]^2 \right]^{1/2}$$

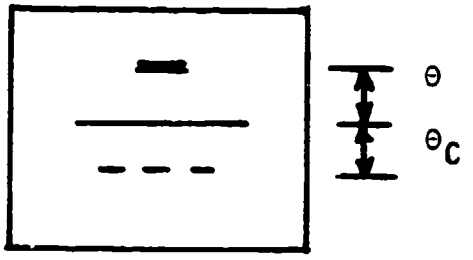
As shown, about 30 seconds of data is required to obtain less than 10% rss error in this example. Other dynamics of higher order, and therefore more gains, will be evaluated in the near future and the convergence will not be as rapid.

The importance of using the proper corrections (e.g., M_{cor} and N_{ν_u}) is shown in Figure 8 , in which the five exact gains, $g_1 \rightarrow g_5$ are shown, along with two sets of gain estimates. The set labeled "uncorrected" was obtained via straight-forward least squares (i.e., M_{cor} and N_{ν_u} not included).

EXAMPLE

PLANT: $\frac{\theta(s)}{\delta(s)} = \frac{11.7}{s}$

$$\frac{\theta_C(s)}{W(s)} = \frac{3.67}{s^2 + 3s + 2.25}$$



$\varepsilon = \theta_C - \theta$

$$X^T = \left[\theta_C \ \dot{\theta}_C \ \theta \ \dot{\theta}\right]$$

$$\dot{X} = AX + BU + N$$

$$\dot{U}_P = G_X^T \hat{X}_P + G_U U_P + V_U$$

$$\left[G_X^T \ \ G_U\right] = \left[5.53,\ 1.86,\ -6.76,\ -3.69,\ -9.28\right]$$

$Q_\varepsilon = 16/.35,\ Q_{\dot{\varepsilon}} = 1/.35,\ Q_{\dot{\delta}} = 1$

Figure 5

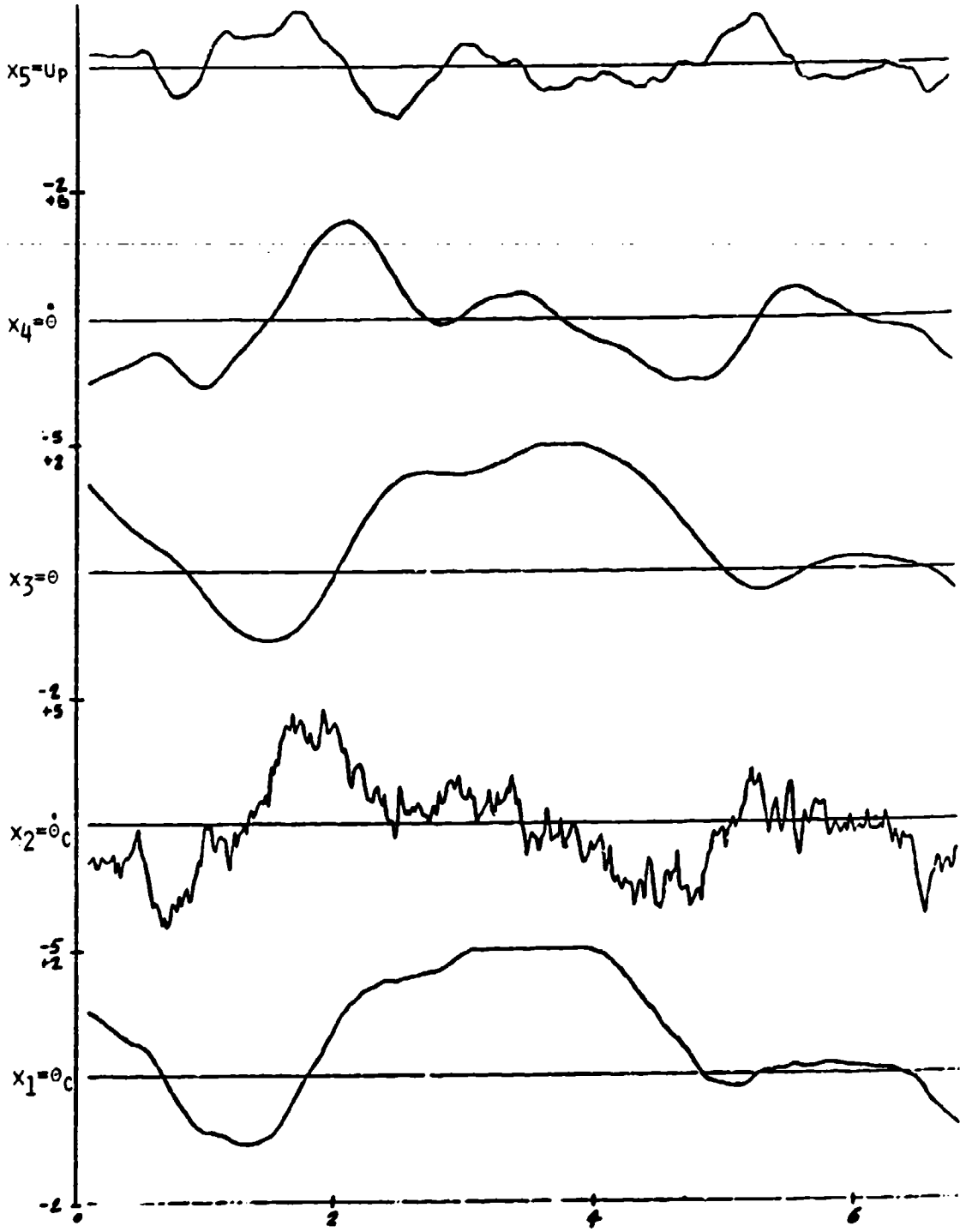


Figure 6 Time Histories

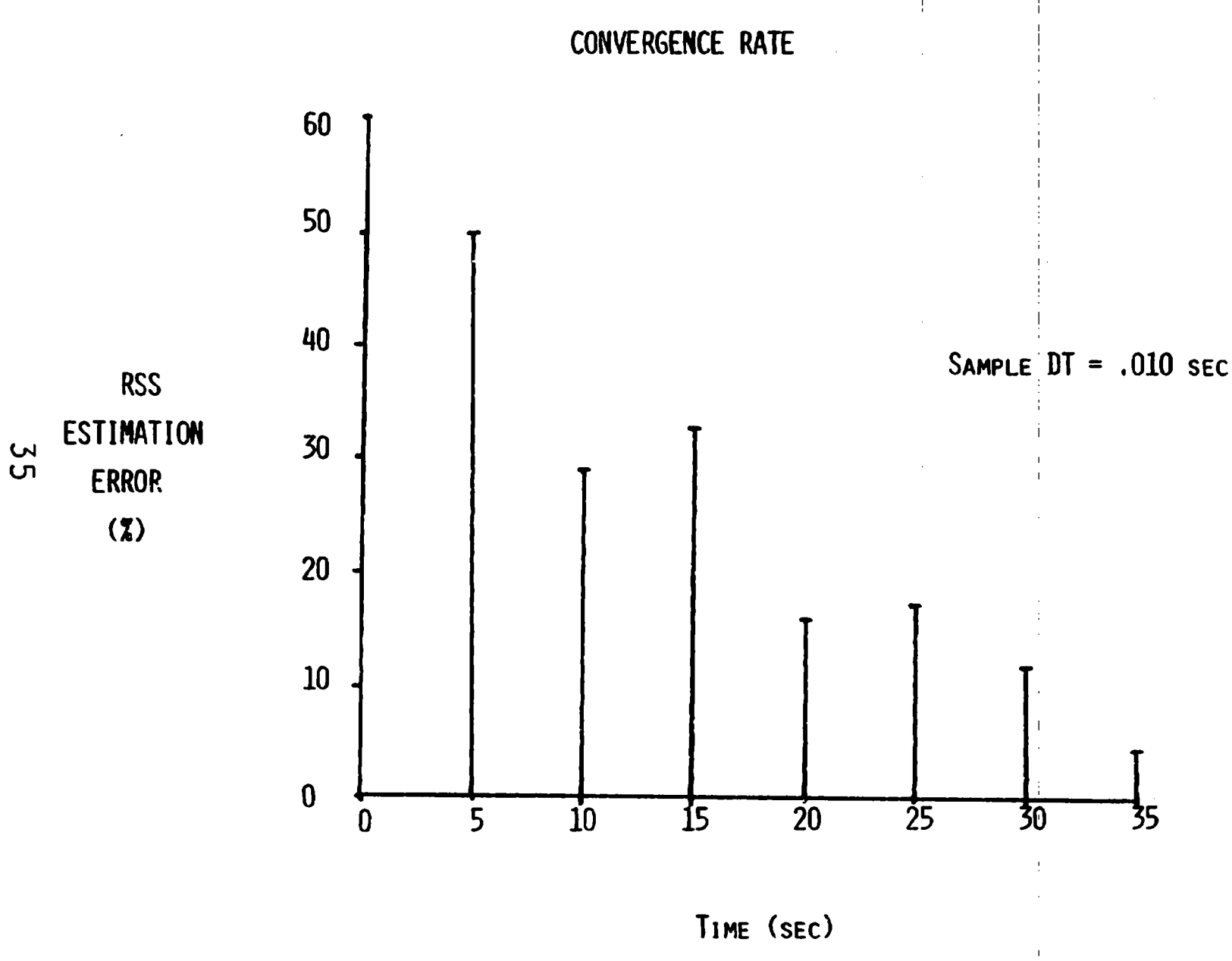
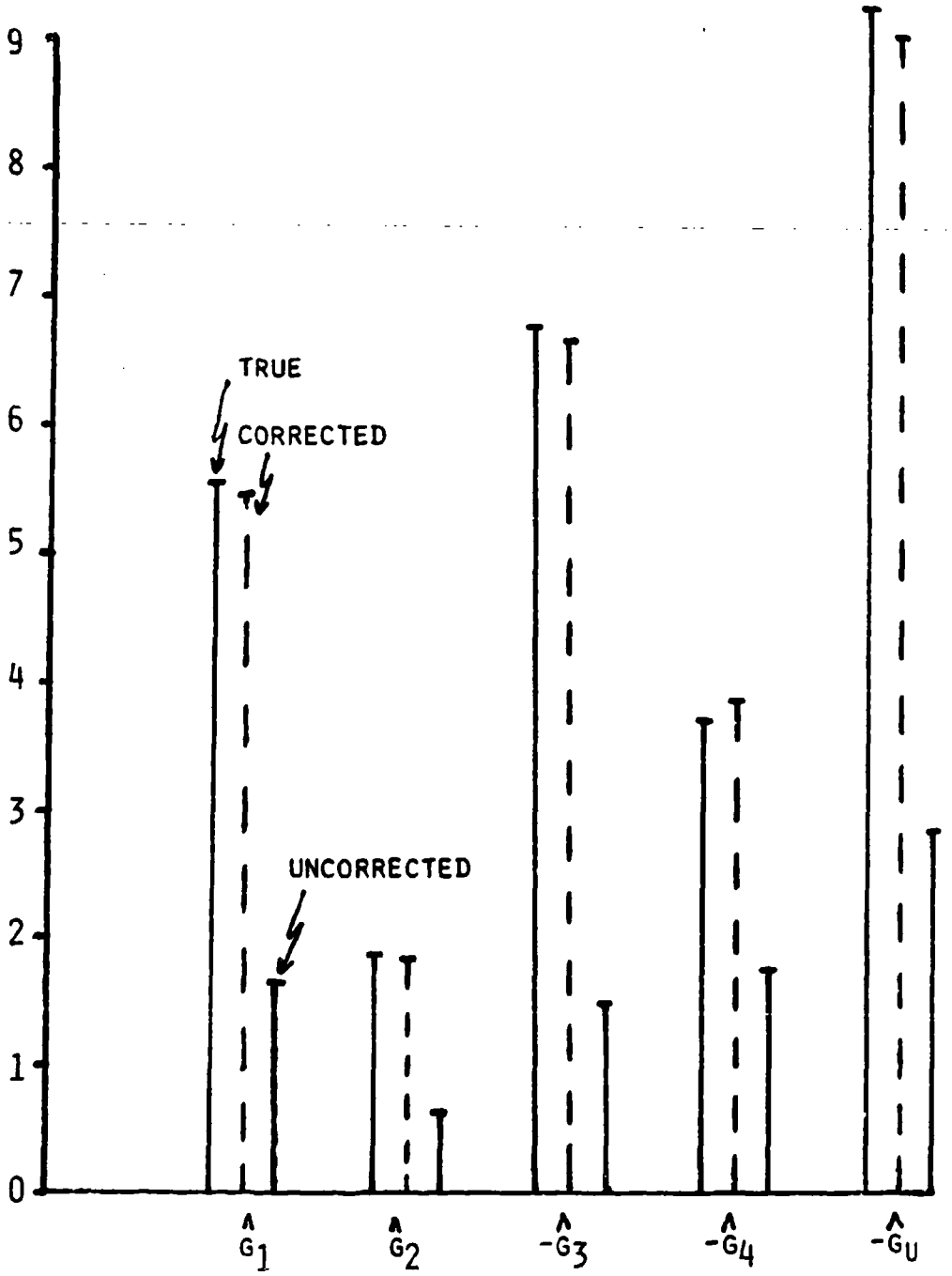

CONVERGENCE RATE
RSS
ESTIMATION
ERROR
(%)
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35
SAMPLE DT = .010 SEC
TIME (SEC)

Figure 7

IMPORTANCE OF CORRECTIONS



GAIN ESTIMATES

Figure 8

Conversely, the "corrected" set used perfectly corrected data, or the actual $\hat{X}$'s in the identification. Both sets of gain estimates are based of 50 seconds of data. Clearly, in this case again, the corrections are important. Further verification of the method is in process.

Inference of the Objective Function

Attention is now turned to estimation of the objective function weightings from the gain estimates just discussed. (Note, this is referred to in the control literature as the "inverse problem".) These weights are related to the gains via the Riccati matrix K, the solution of

$$\tilde{A}^T K + K\tilde{A} + \tilde{Q} - K\tilde{B}G^{-1}\tilde{B}^T K = 0$$

and

$$\left[\begin{array}{c|c} g_x^T & g_u^T \end{array}\right] = -G^{-1}\tilde{B}^T K$$

where

$$\tilde{A} = \left[\begin{array}{c|c} A & B \\ \hline 0 & 0 \end{array}\right] \qquad \tilde{Q} = \left[\begin{array}{c|c} C^T Q_y C & 0 \\ \hline 0 & R \end{array}\right]$$

$$\tilde{B}^T = \left[0 \mid I_u\right] \qquad I_u = \text{identity of dimension equal to } u_p \text{ control vector}$$

And recall that Q_y, R, and G are the weightings defined in Eqn (0). Now due to the structure of the OCM, we are able to reduce the above into some simpler relations. First, noting that letting $G = I_u$, without loss of generality (at least in the case of scalar control input u_p), we obtain

$$K = \left[\begin{array}{c|c} L & -g_x \\ \hline -g_x^T & -g_u \end{array}\right], \quad g_u = g_u^T$$

and

$$R = g_u g_u^T + g_x^T B + B^T g_x \qquad \text{(III.a)}$$

$$0 = g_x g_u - LB + A^T g_x \qquad \text{(III.b)}$$

$$C^T Q_y C = g_x g_x^T - LA - A^T L \qquad \text{(III.c)}$$

Now L can be eliminated in the last two relations if desired, and have

$$B^T C^T Q_y C B = B^T g_x g_x^T B - H - H^T \qquad \text{(IV)}$$

where

$$H = (g_u g_x^T + g_x^T A)AB$$

By observing Equation III (and IV) we can see that if estimates of g_x and g_u are available, and plant and observation matrices A, B, and C are known, the R weighting can be obtained directly, but Q_y requires special attention. From Egns. III.b and .c we see that if L can be obtained by solving III.b, an n x n matrix equation with only $C^T Q C$ unknown results from III.c. But this is only possible if B^{-1} exists, which is only true if the number of independent control inputs (in u_p) equals the number of states (in x)-an unlikely situation.

An alternate attack using Eqn. IV leads to similar results. One could conceivably solve for a diagonal Q_y via a numerical method like Newton-Raphson, but that requires the matrix $CBB^T C^T$ to be invertable. This is possible if the number of control inputs (in u_p) equals the number of outputs (or y), (or the system transfer function matrix is square). Although this is less restrictive than the previous situation, it is also untrue in many applications of interest to us here. So the following conclusions may be stated, that in general a <u>unique</u> set of objective function weights may not be obtainable from gain estimates alone. This result is not new, we've just looked at it in the context of our specific problem.

Although improved methods are currently under investigation in this regard, we may always test assumed objective function weights to determine if they're feasible. This is considered a reasonable alternative since in an actual experiment, the analyst knows several reasonable statements for the objective function, and he may at least test them to see which one is best supported by the data. To pursue this approach, the accuracy of the gain estimate will also be developed such that statistical hypothesis tests may be performed. But for now, this is an important consideration.

In the case of our numerical example, Equation III.a leads to R = 0, and Equation IV yields

$$(11.7)^2(q_\varepsilon^* + q_\theta^*) = -2(11.7)g_u g_{x_3} + (11.7\ g_{x_r})^2$$

where

$$Q_y = \begin{bmatrix} q_\varepsilon & 0 & 0 & 0 \\ 0 & q_{\dot{\varepsilon}} & 0 & 0 \\ 0 & 0 & q_\theta & 0 \\ 0 & 0 & 0 & q_{\dot{\theta}}^* \end{bmatrix}$$

Using the estimated gains we obtain

$$(q_\varepsilon^* + q_\theta^*) \tilde{=} 2.89 \quad \text{(actually } q_{\dot{\varepsilon}} \text{ was } 1'.35 \text{ and } q_{\dot{\theta}} \text{ was 0)}$$

Now "guessing" that q_θ and $q_{\dot{\theta}}^*$ were zero, at first, we may iterate on q_ε and solve III.c, then check with III.b. If no $q_\varepsilon > 0$ led to a solution, then the assumption of q_θ and $q_{\dot{\theta}}^*$ equal to zero would need revision. Finally, note that from Equation III.b, we can actually solve for as many columns (and rows) of L as the number of control inputs (or rank B), and this part of the L matrix may be used to check results from III.c.

Acknowledgement

This work is being performed with the cooperation of NASA Dryden Research Facility under NASA Grant NAG4-1. Mr. Donald T. Berry is the technical monitor, and his support, and that of NASA, is appreciated.

References

1. Kleinman, D., Baron, S., and Levison, W.H., "An Optimal Control Model of Human Response, Part I: Theory and Validation," Automatica, Vol. 6, 1970, pp. 357-369.

2. Hess, R.A., "Prediction of Pilot Opinion Rating Using an Optimal Pilot Model," Human Factors, Vol. 19, Oct. 1977, pp. 459-475.

3. Schmidt, D.K., "On the Use of the OCM's Objective Function as a Pilot Rating Metric," 17th Annual Conf. on Manual Control, UCLA, June, 1981.

4. Levison, W.H., "Methods for Identifying Pilot Dynamics," Proceedings of the USAF/NASA Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics, Edwards AFB, Jan. 19-21, 1982.

5. Bryson, A.E., and Ho, Y.C., Applied Optimal Control, Halsted Press, NY, 1975.

6. McGruer, D.T., and Krendel, E.S., Mathematical Models of Human Pilot Behavior, AGARDograph, No. 188, Jan. 1974.

A TEST OF FITTS' LAW IN TWO DIMENSIONS WITH HAND AND HEAD MOVEMENTS

Richard J. Jagacinski
Ohio State University
Columbus, Ohio 43210

Donald L. Monk
Air Force Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio 45433

ABSTRACT

Subjects used either a joystick or a helmet-mounted sight to capture stationary circular targets that randomly appeared on a CRT screen. Capture times were longer with the helmet- mounted sight. For both hand and head movements, Fitts' Law was found to approximate the pattern of capture times along each of eight radial axes extending from the center of the screen. Namely, capture times along each axis increased as a linear function of the logarithm of movement distance divided by target width. The effective target width was corrected for the noise distribution associated with each control system's hardware. Additionally, for the helmet-mounted sight, mean capture times along the diagonal axes were somewhat longer than along the vertical and horizontal axes. A similar trend was found for the hand movements, but it was smaller in magnitude and not statistically significant. These patterns can be used to evaluate various composition rules for a two-dimensional version of Fitts' Law.

AD P001617

A Nonlinear Internal Feedback Model for the Pupil Control System

By Fuchuan Sun, William C. Krenz, and Lawrence W. Stark

Neurology Unit
University of California, Berkeley

ABSTRACT

The human pupillary control system is a typical biological control system and exhibits a series of interesting nonlinear behaviors, particularly asymmetry, 'pupillary escape,' and 'pupillary capture.' We present here a fairly simple nonlinear model, in which a signal statically proportional to pupil size is fed back internally to cause a change in system parameters related to gains and rates of light adaptation. The model was simulated on a digital computer, improvements over previous pupil models demonstrated, and a variety of experimental data was well matched. Candidate physiological mechanisms for the components of the model are discussed.

Acknowledgement: We gratefully acknowledge partial support from the NCC 2-86 Cooperative Agreement with NASA-Ames Research Center.

INTRODUCTION

We have undertaken a modeling effort to describe the pupillary light reflex system, including the retina, pretectum, Edinger-Westphal nucleus and iris. There are essentially three types of behavior which we wish our model to describe: pupillary escape, pupillary capture, and the asymmetry of the response. Pupillary escape (as named by Lowenstein and Loewenfeld (1969)) is the phenomenon that, when stimulated with a step input of light, the pupil will quickly constrict, then slowly dilate. However, when the pupil is initially small or the intensity of the input is large, it will constrict more slowly and not dilate, a process which has been called pupillary capture (Usui (1974)). Finally, when the pupil is stimulated by a negative step (decrease in light intensity) the response is a small dilation. We feel these are some of the dominant characteristics of the pupillary light reflex system.

Early efforts to model the pupil (Stark(1959), Clynes(1961), Troestra(1968), Webster(1971)) dealt mostly with the pupillary escape and asymmetry phenomena. Semmlow & Stark (1972) and Semmlow & Chen (1977) developed models utilizing saturation of the iris to describe the pupillary capture behavior. Shimizu (1977) proposed a model whereby the time constant of the reponse changed with the intensity of the stimulus.

Each of these models fit the experimental data for a small class of inputs, in a qualitative sense.

The problem with the above models is that none could quantitatively match the pupillary escape and capture responses, particularly when the pupil size was small. Our goal in this effort is to match that data for a variety of initial conditions. Our present model rests on the use of feedback from a signal proportional to pupil size to cause a change in system parameters related to gains and rates of light adaptation.

MODEL DEVELOPMENT

Sobel & Stark (1962) proposed an internal feedback model of the pupil in which the parameters of the system could be changed by the DC gain. We modified this after observing that the pupil response seemed to be a function of the pupil size. We know this because pupillary capture occurs when the pupil is small, whether this is caused by light intensity or others factors, such as accommodation. This led us to develop a model utilizing feedback of tonic pupil size to change the parameters.

The model of Sobel & Stark has been reformulated to include pupil size feedback and to create separate tonic and phasic pathways through which the input travels [Fig. 1]. The phasic (AC) path is a high-pass filter which responds to any positive-going transition. The gain (Kac) of this response

becomes smaller as the tonic pupil size decreases. The tonic (DC) path is a low-pass filter whose gain (Kdc) depends on the pupil size. A large pupil implies a large Kac and small Kdc, while a small pupil decreases Kac and increases Kdc. The remaining elements of the model (LOG operator, third-order integrator with delay, and iris nonlinearity) are well-established for the human pupil.

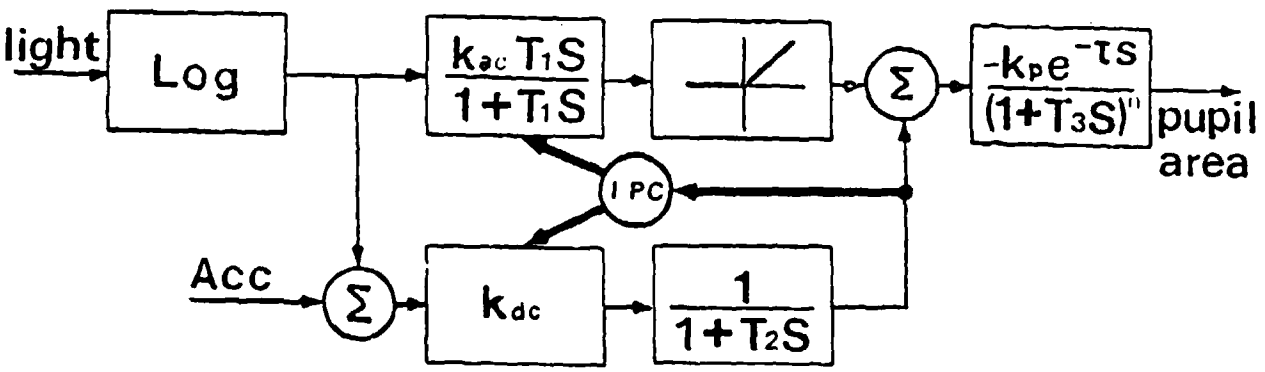


FIGURE 1 : Complete model showing separation of AC and DC paths and internal parameter control.

METHODS

All of our simulations were run on an LSI-11/23 (with floating point hardware and graphics display) and took approximately 5 seconds to simulate six seconds worth of data. The model was reformulated into a state-space representation and Runge-Kutta integration was performed, using multiple-step and sinusoidal input functions. The model parameters were optimized

so as to fit the data. Our data was gathered using an infrared television pupillometer. The prototypical data was the average of 10 records from 5 normal subjects who showed similar responses. The different amplitudes of light input were accomplished by inserting a neutral density filter in front of the light source. These filters had effects of -5 log units (resulting in small inputs), -3 log units (medium inputs), and 0 log units (large inputs).

RESULTS

The response of the model to a step input of light is shown in Figure 2. It is clearly seen that the model responds well to small and large (pupillary escape and capture, respectively) inputs. The response to the intermediate input displayed oscillations, implying a more complex model than that which we have suggested. We feel, however, that within the context of the goal -- describing pupillary escape and capture -- the model matches the data very well.

To test the model under the various condtions for which it was designed, we also applied small and large step inputs of light to small and large pupils (set initially by accommodation). All three comparisons are shown in Fig. 3, and again it is seen that the model provides an excellent fit. Figure 3a compares the responses of pupils with similar initial sizes for large (lower curve) and small (upper curve) inputs. A small

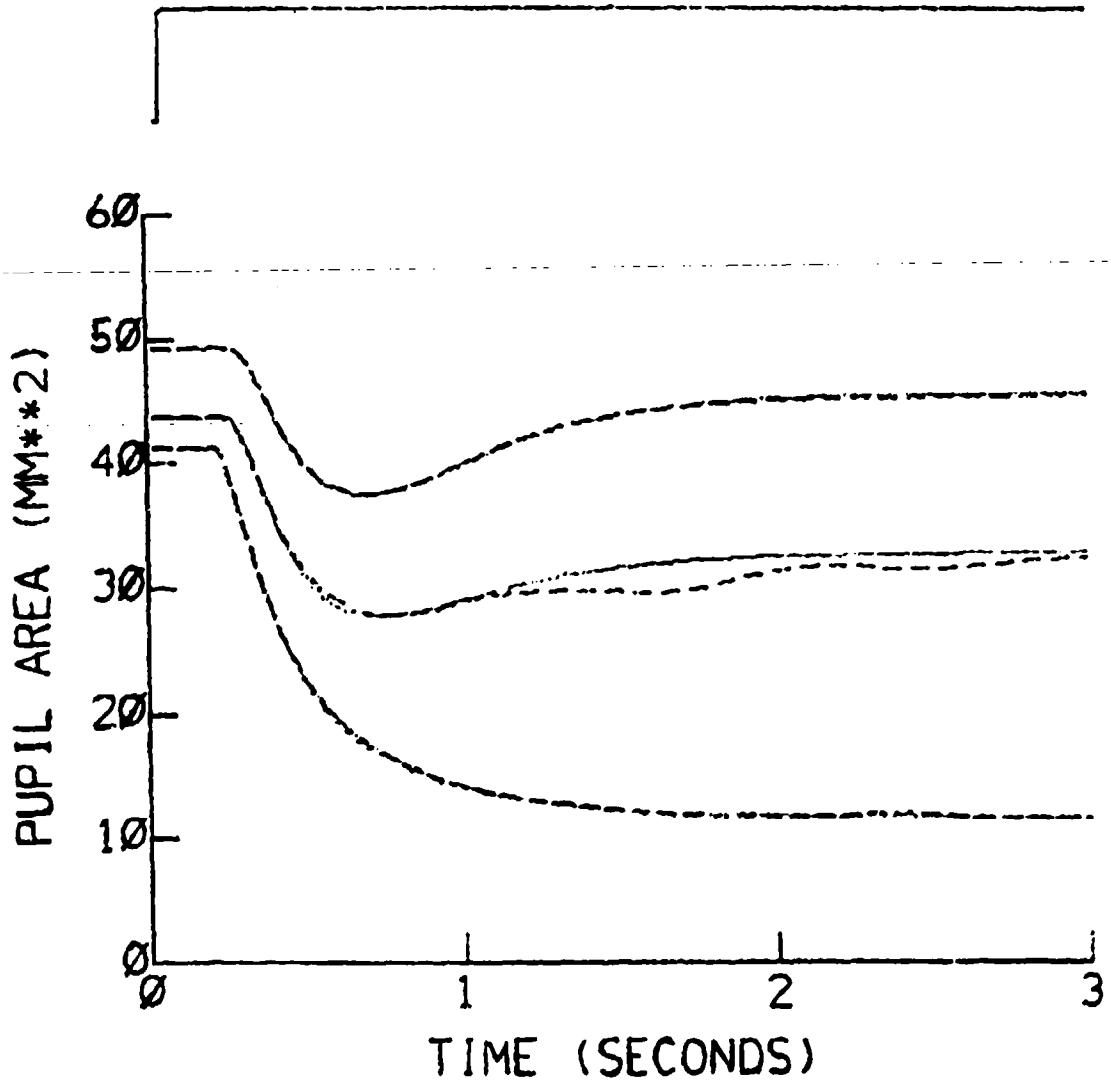


FIGURE 2 : Pupillary responses to a step input of light. Upper graph is the input stimulus for experiment and model. Lower graph shows pupillary response to a small input (upper curve), medium input, and large input (lower curve). Experimental data is represented by dashed lines, model data by solid lines.

input was applied to different initial pupil sizes [Fig. 3b] to show that pupillary capture occurs in the small pupil, while escape occurs in the large pupil, although the inputs were identical. Large inputs were also applied to small and large pupils, yielding the responses shown in Figure 3c. As was expected, capture occurred for each of the initial conditions. We can therefore conclude that the necessary conditions for pupillary escape are small light input and large initial pupil size.

In order to prove that pupillary capture is not merely a saturation effect, as has been proposed; the pupil was stimulated with a large double step of light. If the pupil was saturated, the second input would cause no change. That however, was not the case [Fig. 4]. The figure shows that the pupil continues to constrict when an additional stimulus is presented.

The final comparisons used to tune the parameters of the model were via the velocity and phase plane graphs [Figs. 5 & 6] Again, it is seen that the fit is quite good for the large and small inputs, and satisfactory for the intermediate input.

As was mentioned above, the gains Kac and Kdc are functions of the tonic pupil size. To simplify the model, the relations were approximated by piecewise linear functions. While a more complex function may have provided a better to the data, the limitations of the data at hand dictated a simple relation. With this in mind, the parameters of the final model [Fig. 7]

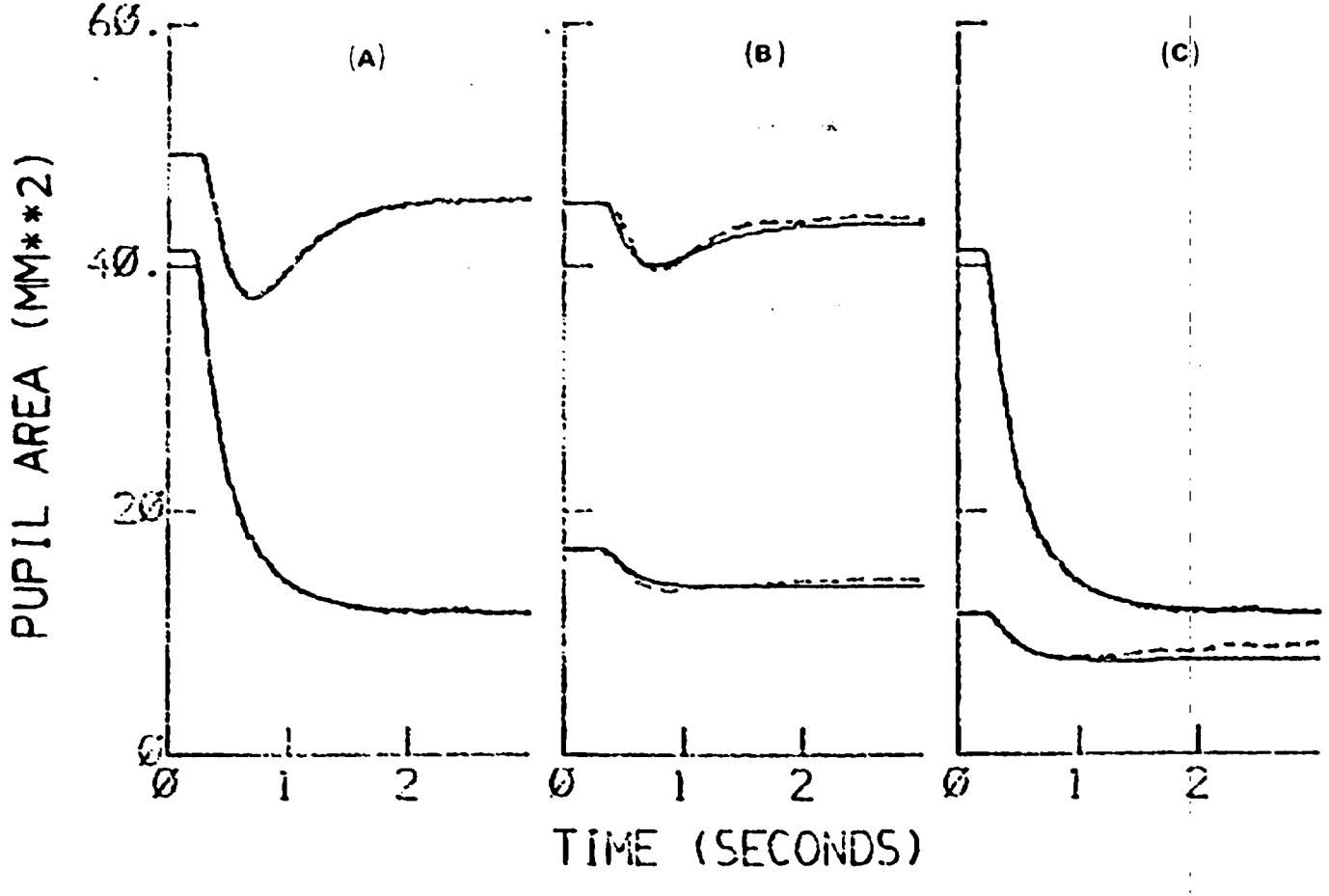


FIGURE 3 : Test conditions for pupillary experimental (dashed line) and model (solid line) responses.
A) Large and small step inputs for large initial pupil size.
B) Small step input for large and small initial pupil size.
C) Large step input for large and small initial pupil size.

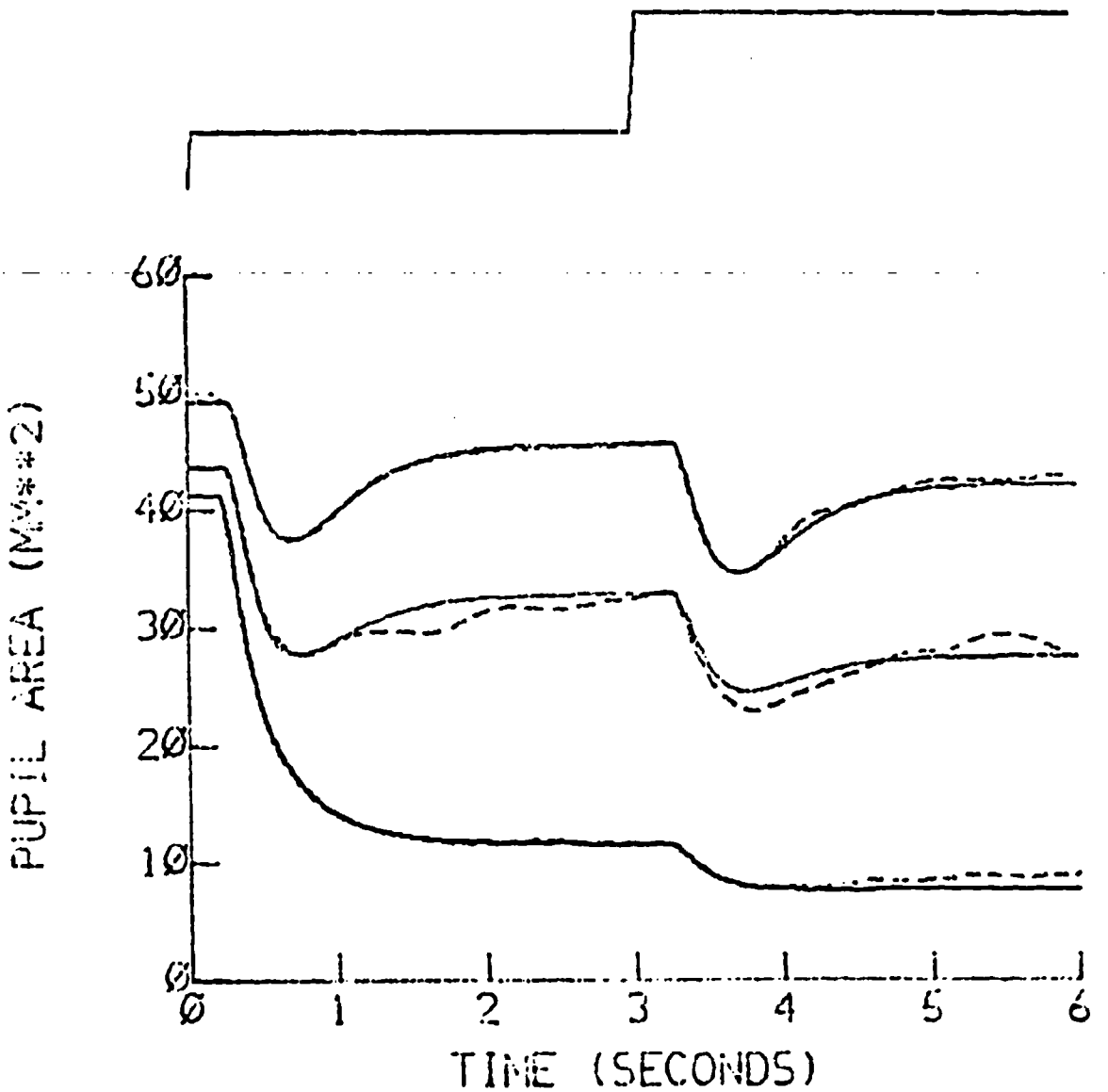


FIGURE 4 : Pupillary response to double step input of light. Upper graph shows double step input for model and experiment. Lower graph shows response to small inputs (upper curve), medium inputs, and large inputs (lower curve).

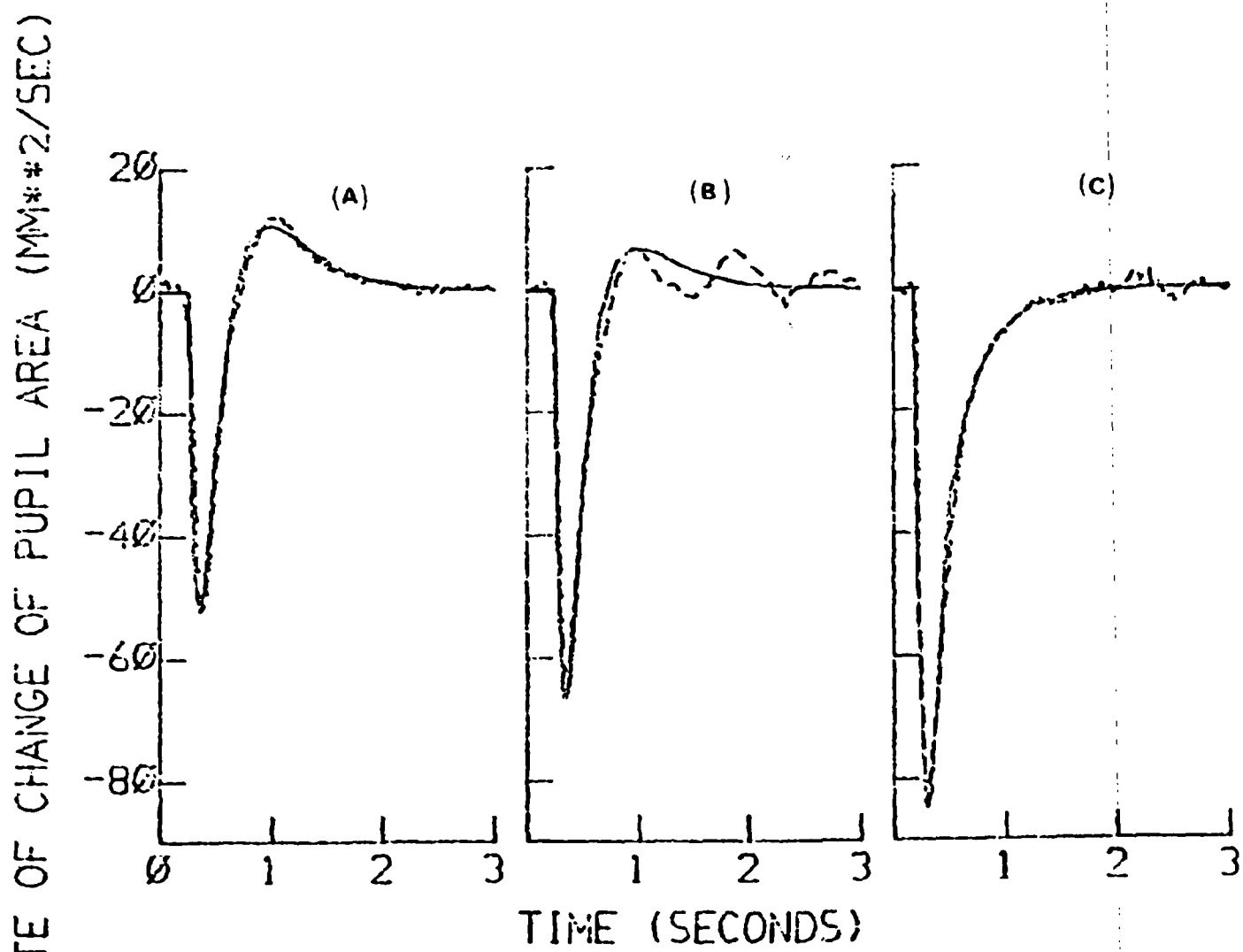


FIGURE 5 : Test of match between model and experimental data using rate of chenge of pupil area. Comparisons were made for small input (A), medium input (B), and large input (C).

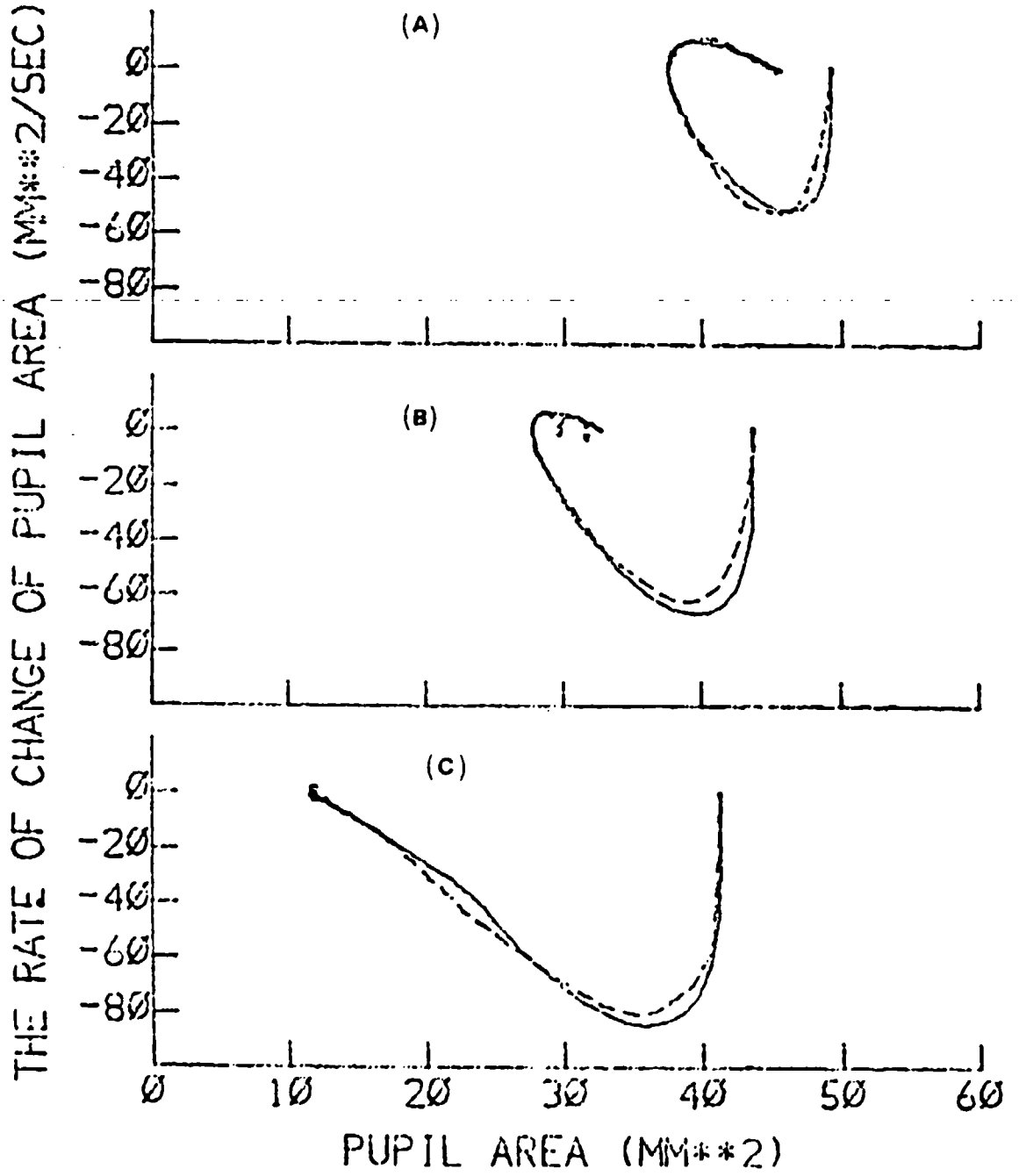


FIGURE 6 : Phase plane trajectories of model (solid) and experimental (dashed) data for small (A), medium (B), and large (C) inputs. Pupillary escape yields the circular form shown in (A), while pupillary capture yields the open form of (C).

are as follows :

T1 = 0.085
T2 = 0.15
T3 = 0.28
Td = 0.2
Kac (large pupil) = 1.0
Kac (med. pupil) = 0.9
Kac (small pupil) = 0.85
Kdc (large pupil) = 0.035
Kdc (med. pupil) = 0.085
Kdc (small pupil) = 0.17

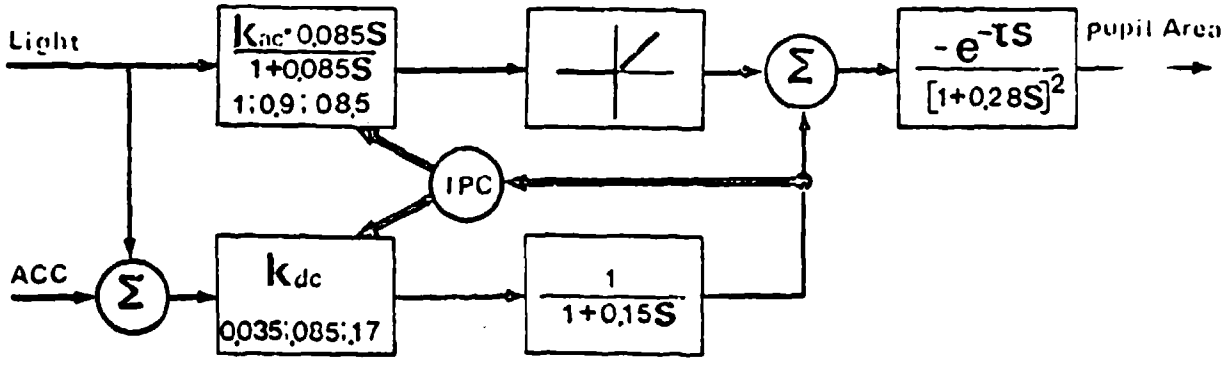


FIGURE 7 : Results of simulation with parameter values shown.

DISCUSSION

Before attempting another pupil model, some of the aforementioned pupil models were used. However, no matter how the parameters were changed, the models would not agree with the available data. In particuar, Shimizu's model, although working for large initial pupil size, could not fit the data when the initial pupil size was small. Our experience with these models lent insight towards what we considered a "better" model, such as separating the AC and DC pathways of the response. It was felt this was valid because the initial (phasic) response to a step seemed relatively independent of the final value which it would attain. Also, it was felt that the main difference between pupillary escape and capture was that the DC path increased its gain for capture while the AC path decreased its gain. Since the data sugggested that pupillary capture was dependent upon tonic pupil size, we could then feed pupil size back to adjust the DC and AC gains, a process which seemed physiologically plausible.

By applying a double-step input of sizable intensity, we showed that the iris saturation model was not plausible. The idea in this model is that pupillary capture is merely an escape which has been saturated. Clearly, from Figure 4, one can see that there is no saturation effect, since the pupil continues to constrict when another input is applied. This suggests that

pupillary capture is a mechanism in its own right, rather than just a special case of pupillary escape.

Throughout these studies, pupil area was used as a measure for pupil size, because it is slightly less sensitive to noise in the data collection process. However, diameter could have been used without ill effects in the model (diameter is often used to measure pupil size). The only difference would be in the functions relating Kdc and Kac to pupil size, but since these are arbitrary nonlinear functions, new functions could be found without harming the contructs of the model.

A model with separate AC and DC pathways, besides yielding better modeling and control, is more realistic physiologically. The AC path corresponds to the (fast) Y cells in the retina, while the DC path corresponds to the (slow) X cells in the retina. Each path reacts differently to different levels of light. For instance, large pupils are usually associated with scotopic (night) vision. In this condition, it is less important for the pupil to regulate the long-term light flux onto the retina via the pupil, implying that the AC path dominates and that pupillary escape occurs in the large pupil. Conversely, a small pupil (about 4-5 mm) is generally indicative of photopic vision conditions. In this case, regulation of light flux into the eye becomes more important, and the DC path becomes prominent. Finally, by using pupil size as a feedback control variable, we can assume this happens at the level of the

Edinger-Westphal nucleus.

CONCLUSION

This model shows that internal pupil size feedback is a valuable tool in determining the pupil response to light. The model of course fits the standard pupillary escape movement, but more importantly, it fits the pupillary capture movement as well. Changing the initial conditions via accommodation shows that the pupil size, not background light intensity, controls the capture mechanism. It was also shown that the old saturation model of capture is not plausible by applying a two-step input. Finally, it was seen that separating the model into tonic and phasic paths is reasonable both physiologically and from a modeling point of view.

It is hoped that, in the future, the model can match the intermediate (partial escape and capture) response of the pupil. The small oscillations present in the data suggest that there is perhaps a delay in the pupil size feedback mechanism. Initial tries at implementing this have not yet given good results. Another possibility is a Volterra series model of this, or perhaps even a hybrid of the two. Overall, however, it is felt that this model is a valuable tool in the study of the pupil.

REFERENCES

1) Lowenstein, O., and Loewenfeld, I.E., "Effect of Various Light Stimuli", THE EYE, p274, ed. by H.Davson, Academic Press,1969.

2) Usui, S., "Functional Organization of the Human Pupillary Light Reflex System", Ph.D. thesis, UC Berkeley, p.86, 1974.

3) Stark, L., "Stability, Oscillations, and Noise in the Human Pupil Servomechanism", Proc. IRE, Vol. 47, pp.1925-1939, 1959.

4) Clynes, M. "Unidirectional Rate Sensitivity: a Biocybernetic Law of Reflex and Humoral Systems as Physiologic Channels of Control and Communication", Ann. NY Acad. of Science #92, pp.946-969, 1968.

5) Troelstra, A., "Detection of Time-varying Light Signals as Measured by the Pupillary Response", J. Opt. Soc. Am #58, pp.685-690, 1968.

6) Webster, J., "Pupillary Light Reflex: Development of Teaching Models", IEEE Trans. BME-18 #3, pp.187-194, 1971.

7) Semmlow, J., and Stark, L., " Simulation of a Biomechanical Model of the Human Pupil", Math. Biol. #11, pp.109-128, 1971.

8) Semmlow, J. and Chen, D., "A Simulation Model of the Human Pupil Light Reflex", Math. Biol. Vol.33, No.1/2, 1977.

9) Shimizu, Y., and Stark, L., "Pupillary Escape", 10th Symp. on the Pupil, New York, June, 1971.

10) Sobel, I., and Stark, L., "Re-evaluation of the Pupil System", Quarterly Progress Report #66, Res. Lab. Elec. MIT, pp.412-419, July, 1962.

AD P001618

COMPUTATIONAL PROBLEMS IN HUMAN OPERATOR ARMA MODELS

by

Gyan C. Agarwal, Hitoshi Miura and Gerald Gottlieb
Department of Industrial and Systems Engineering
University of Illinois at Chicago
Chicago, Illinois 60680

and
Department of Physiology
Rush Medical College
Chicago, Illinois 60612

ABSTRACT

The time series technique as applied to human operator dynamic system identification is examined for a known input-output data sequence. The error in parameter estimates increases with increased sampling interval. In the presence of measurement noise, higher order models are necessary to minimize the residual error.

INTRODUCTION

In the past, much of the engineering concerned with people in systems could be accomplished by trial and error or by common sense. The operator was left to adjust to inadequacies of the system as best he/she could. Indeed, the operator generally adapted to difficult systems very well, but at a cost of fatigue, poor system performance, and error under stress unacceptable in many systems today. It has thus become necessary to take account of the man-machine interaction in the design process, and to do so in such a way that the consequences of design choices can be predicted in terms of system-performance criteria. This in turn requires that the system designer be able to

characterize the way the human operator behaves in a manner consistent with his description of machine functioning, i.e., that he be able to model and predict the performance of the operator as a component of the man-machine system (1).

Since the work of the human operator model by Tustin, many investigators have proposed different models to quantitatively represent human operator dynamics (1-7). Recently a powerful new method utilizing time series analysis has been considered as one of the promising tools among those techniques (2, 7, 9). Shinners (2) and Tanaka, Goto, and Wasizu (7) have applied the time series analysis to their models of the human operator. These investigators used the methods of Box and Jenkins and Akaike's information criterion (AIC) respectively to estimate the order and the value of parameters. Shinners (2) obtained the transfer function of the human operator in compensatory tracking with a controlled element of unity gain in the following form:

$$G_H(z) = \frac{-0.65(1-0.803z^{-1})z^{-1}}{(1-0.386z^{-1})\ (1-0.97z^{-1})} \qquad (1)$$

where z^{-1} is the backward shift operator, namely $z^{-1}x(k) = x(k-1)$.

Using the Normalized Residual Criterion (NRC) of Suen and Liu (8) to estimate the order and the values of the parameters, Agarwal et. al. (9, 14) have obtained the transfer function of the human operator in both compensatory and pursuit tracking modes. One of their models in compensatory tracking is given by:

$$G_H(z) = \frac{-0.106(1-z^{-1})}{(1-z^{-1})\ (1-0.479z^{-1})} \qquad (2)$$

These time series analysis methods have the advantage that they require much smaller data lengths than the techniques based on fourier transform methods. Thus, changes in human operator dynamics could be studied much more effectively than hereto possible. The purpose of the present paper is to investigate how sampling interval and measurement noise would influence the results in using time series techniques.

METHODS AND DATA ANALYSIS

For the purpose of evaluating the time series technique used in the identification of the human operator dynamics a controlled set of input-output data is generated. Figure 1 shows the relationship between input and output of the man-machine system considered in this paper.

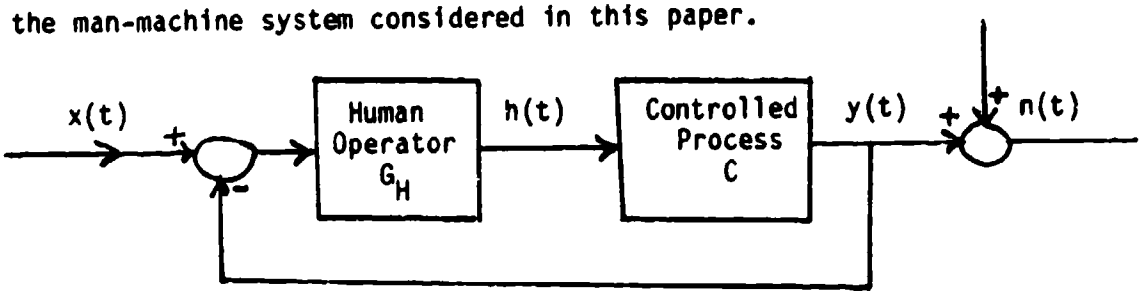


Figure 1

The input x(t) is synthesized by adding five .on-harmonic sine waves:

$$x(n\Delta t) = \sum_{j=1}^{5} A_j \sin[2\pi f_j (n\Delta t)] \qquad (3)$$

$$n = 0, 1, 2, \ldots$$

The frequencies and amplitudes were chosen from the STI data (10) as follows:

f_j =	0.0797	0.199	0.479	0.997	1.670	Hz
A_j =	6.46	2.58	1.07	0.516	0.309	

The human operator transfer function and the controlled process dynamics were chosen from Shirly's work (11) as follows:

$$G_H(s) = \frac{5(8.7-s)}{8.7+s} \; ; \quad C(s) = \frac{1}{s} \tag{4}$$

With the input and the system dynamics defined as above, the output is obtained by solving the linear differential equations by using the fourth order Runge-Kutta method with $\Delta t = 0.01$ seconds. The regression analysis was done by using MINITAB 2 statistical software package on an IBM computer. The first five seconds of the input-output data sequence were not used to allow the initial transient to die. The input-output data sequences were sampled at four different sampling intervals of 0.01, 0.02, 0.1 and 0.2 seconds to investigate the effect of the sampling period. This analysis was repeated with measurement noise n(t) added to the output signal. This Gaussian noise signal was digitally generated with mean $\mu = 0$ and the standard deviation $\sigma = 0.25$ or 0.5. To obtain the parameter values, 100 or 200 observation points were used.

The transfer function model using n lags on output and m lags on input, TF(n,m), is specified by the equation

$$y(t) = \sum_{i=1}^{n} \alpha_i \, y(t-i) + \sum_{j=0}^{m} \beta_j \, x(t-j) + v(t) \tag{5}$$

$$(m \leq n)$$

The parameters n and m were chosen using the Normalized Residual Criterion (8). This method has been outlined earlier in Agarwal et. al. (9,14). The coefficients α_i and β_j are obtained in the regression analysis.The residuals v(t) are minimized in the NRC method and for the chosen model must have white noise properties. The normalized sum of the squares of residuals is given by

$$\frac{\|\underline{V}\|^2}{\|\underline{Y}\|^2} = \frac{\|y-\Sigma\alpha_i y-\Sigma\beta_j x\|^2}{\|\underline{Y}\|^2} = \varepsilon\ (n,\ m,\ T) \qquad (6)$$

Where $\|\underline{V}\|$ denotes the sum of the squares of residuals and $\|\underline{Y}\|$ denotes the sum of squares of the output. The quantity, $\varepsilon\ (n, m, T)$, is proportional to the normalized variance of the regression for a given n and m. If this ratio is minimized over n and m, then the data fit as measured by the correlation coefficient ρ will be maximized, i.e.,

$$\hat{\rho} = [1 - \hat{\varepsilon}\ (n, m)\ ]^{1/2} \qquad (7)$$

where T, being a constant for the data, is omitted in the expression, and $\hat{\varepsilon}\ (n, m)$ is the minimum value for $\varepsilon\ (n, m, T)$.

Values of $\varepsilon\ (n, m)$ were computed for several sampling intervals, with and without noise and plotted against different values of n as shown in Figure 2 for two different sampling intervals, without noise. In a typical example of Figure 2-a, all curves turn flat at n = 2, and this is chosen as the optimal value of n. After determining the value of n, the optimal value of m is selected. Since there is no apparent explanatory difference provided by

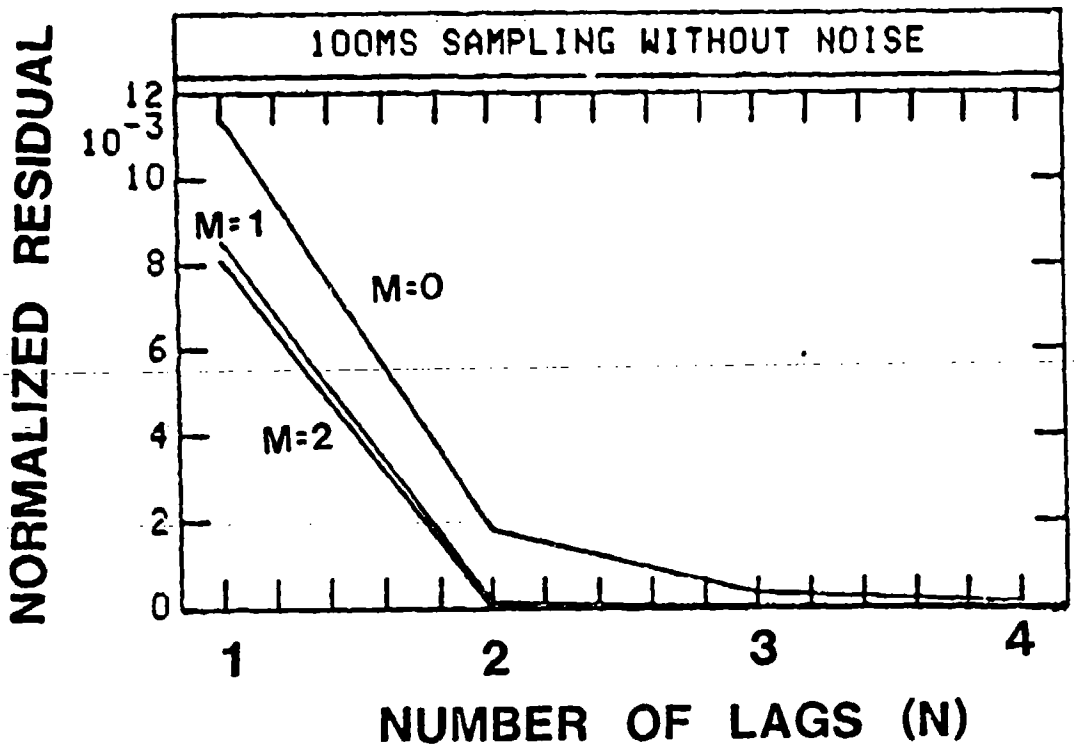

100MS SAMPLING WITHOUT NOISE
NORMALIZED RESIDUAL
10^{-3}
12
10
8
6
4
2
0
M=1
M=0
M=2
1
2
3
4
NUMBER OF LAGS (N)

Fig. 2-a

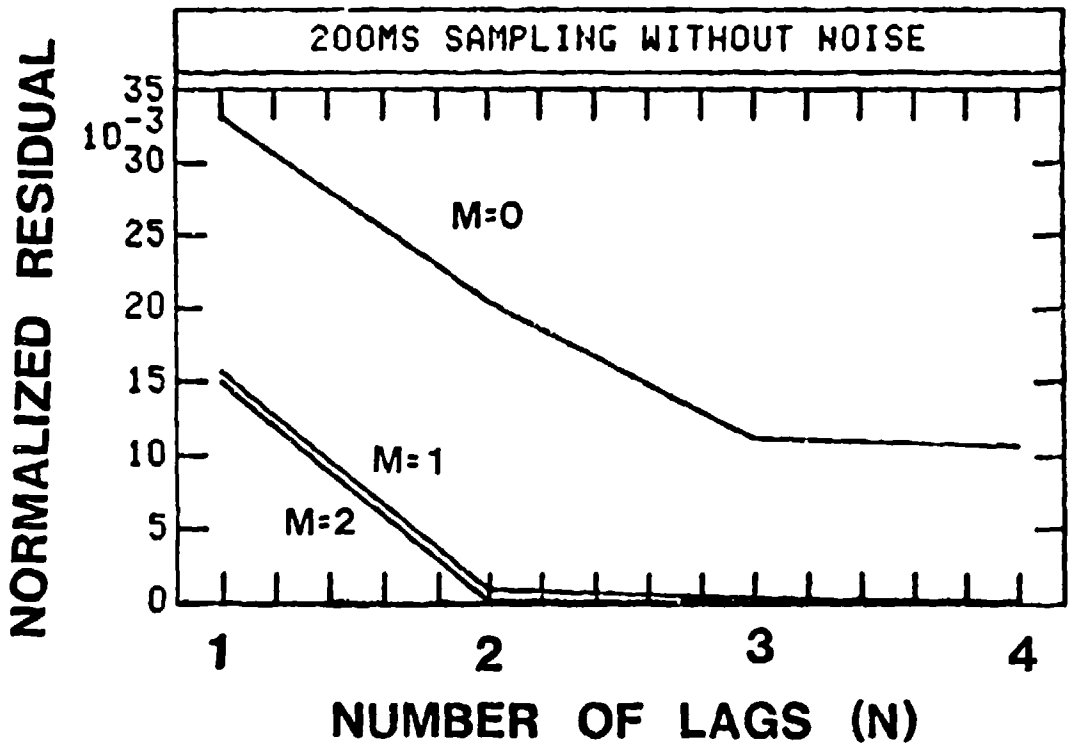

200MS SAMPLING WITHOUT NOISE
NORMALIZED RESIDUAL
10^{-3}
35
30
25
20
15
10
5
0
M=0
M=1
M=2
1
2
3
4
NUMBER OF LAGS (N)

Fig. 2-b

models with m > 1, m = 1 is chosen as the best value. To add statistical significance to this selection of model order, note that the critical test is whether for m > 1, one can justify going from n = 1 to n = 2. Since (13),

$$\Delta = \left(\frac{\|\underline{v}\|_1^2 - \|\underline{v}\|_2^2}{\|\underline{v}\|_2^2} \right) \left(\frac{T - n_2}{n_2 - n_1} \right) \qquad (8)$$

is $F(n_2 - n_1, T - n_2)$ distributed for large T. We tested this difference in going from $n_1 = 1$ to $n_2 = 2$. For example, in Figure 2-b, for m = 1 in TF model, using T = 98, $n_1 = 1$, and $n_2 = 2$, $\Delta = 29300$ is obtained. Since the F distribution tables gives $F_{0.01}(1, 120) = 6.85$, it is clear that the data decisively rejects the hypothesis of no significant difference in normalized residual for n = 1 and n = 2.

The resulting estimated model of the human operator, in the case without measurement noise, can be written in the following form:

$$y(t) = \hat{\alpha}_1 y(t-1) + \hat{\alpha}_2 y(t-2) + \hat{\beta}_0 x(t) + \hat{\beta}_1 x(t-1) + \hat{v}(t) \qquad (9)$$

This form of the model is the same as the one for high frequency case of Agarwal et. al. (9). This analysis was repeated at sampling intervals of 0.01, 0.02, 0.1 and 0.2 seconds.

Effect of Noise on Model and Parameter Estimation

Identification of the model and estimation of the parameters are performed again with input data x(t) and the output data corrupted by noise, y(t)+n(t), where the measurement noise n(t) is chosen from the normal

distribution of zero mean and S.D. of 0.25 or 0.5.

The output with noise and its predicted value by the estimated equation are shown for a sampling interval of 0.1 seconds in Figure 3. Note that the optimal model order has been changed to TF (4,2) with measurement noise present.

Residual Analysis

Several statistical tests are performed on the estimated residuals $\hat{v}(t)$ which were assumed to have a normal distribution and statistically independent. The normality can be examined quantitatively by the goodness-of-fit test based on the quantity:

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \qquad (10)$$

where o_i and e_i represent the observed and expected frequencies, respectively for i th cell. The sampling distribution is approximately the Chi-square distribution with K - 1 degrees of freedom, where K is the number of intervals for a given histogram. For example, for 100 ms sampling using a 7 interval histogram $\chi^2 = 10.70$. Since this is less than 12.592, the value for $\chi^2_{0.05}$ for 6 degrees of freedom, the hypothesis that the residuals come from a normal population cannot be rejected at the 0.05 level of significance. From this test it is concluded that the residuals are almost certainly from a normal distribution.

Although $\hat{v}(t)$ values pass the normality test well, they might be autocorrelated. From several autocorrelation tests, the standard 'runs'

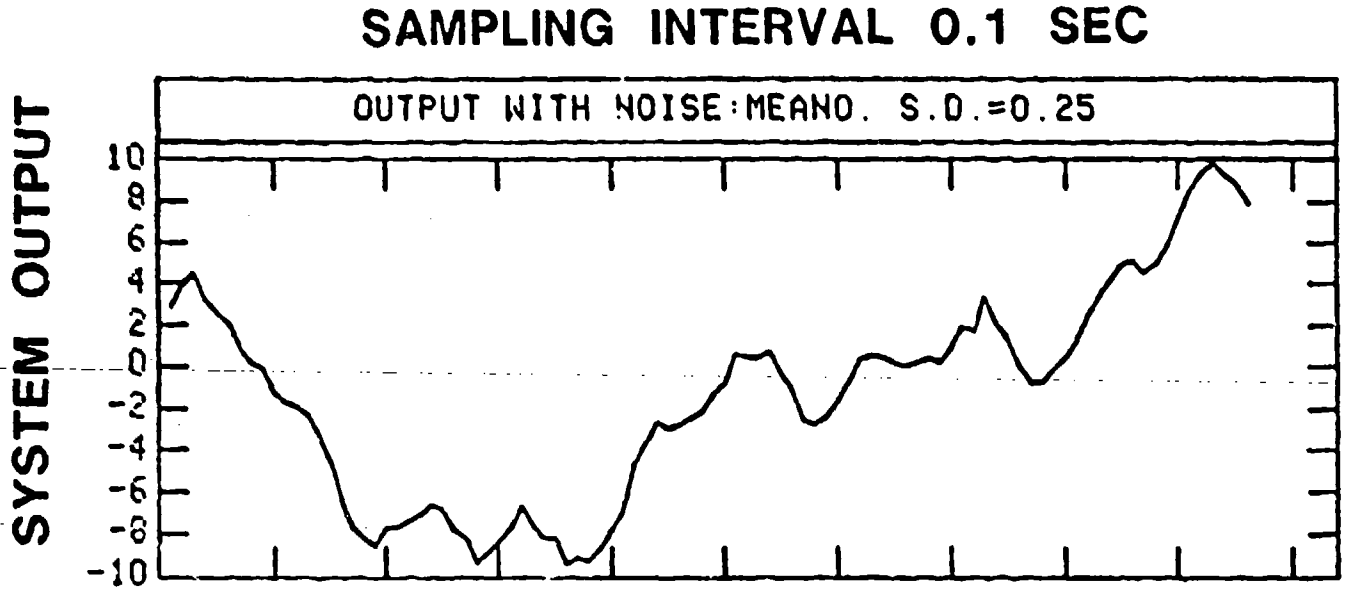

SAMPLING INTERVAL 0.1 SEC
OUTPUT WITH NOISE:MEANO. S.D.=0.25
SYSTEM OUTPUT
10
8
6
4
2
0
-2
-4
-6
-8
-10

Fig. 3-a

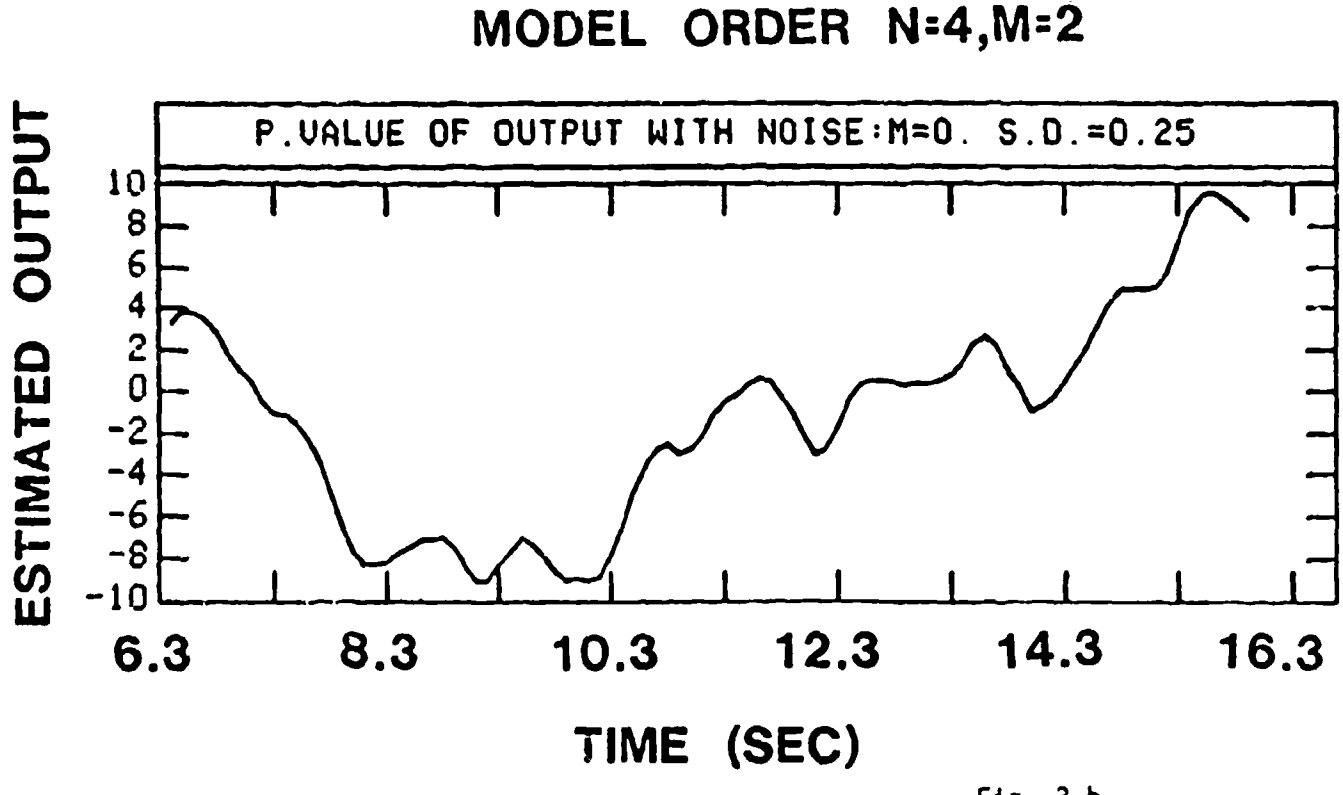

MODEL ORDER N=4,M=2
P.VALUE OF OUTPUT WITH NOISE:M=0. S.D.=0.25
ESTIMATED OUTPUT
10
8
6
4
2
0
-2
-4
-6
-8
-10
6.3
8.3
10.3
12.3
14.3
16.3
TIME (SEC)

Fig. 3-b

randomness test was selected to apply because this method is so simple and reliable. In this case, since the $v(t)$ has zero mean, replacing each residual by the sign (+) if it is greater than 0, by the sign (-) if it is less than 0, the sequence of (+) and (-) are obtained. When n_1 is the number of signs (+), n_2 the numbers of signs (-) respectively, the test whether the arrangement of the signs is random is based on the statistic:

$$Z = \frac{u - \upsilon}{\sigma} \tag{11}$$

where:

$$\upsilon = \frac{2n_1n_2}{n_1 + n_2} + 1, \qquad \sigma = \sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1 + n_2)^2 (n_1 + n_2-1)}}$$

and u is the number of runs.

For example, for 100 ms sampling with noise: ($\mu = 0$, $\sigma = 0.5$), this sequence has 46 runs and $n_1 = 45$ and $n_2 = 51$. Since $Z = -0.58$, which is less than $Z_{0.05} = 1.96$, the hypothesis that the arrangement of signs (and hence, the residual) is random can't be rejected at the 0.05 level of significance.

Transfer Function of System and Human Operator

A linear difference equation such as eq. (9) is transformed using Z transforms, into the following equation:

$$y(z) = \frac{\beta_0+\beta_1z^{-1}+.....+\beta_mz^{-m}}{1-\alpha_1z^{-1}-.....-\alpha_nz^{-n}} \quad x(z) = H(z)x(z) \tag{12}$$

where H(z) is the transfer function of the discrete system. This equation

where H(z) is the transfer function of the discrete system. This equation defines the closed-loop transfer function of the system. For example, for 100 ms sampling without noise,

$$H(z) = \frac{-0.315 + 0.628z^{-1}}{1 - 1.384z^{-1} + 0.699z^{-2}} = \frac{G_H(z)C(z)}{1 + G_H(z)C(z)} \qquad (13)$$

where $G_H(z)$ is the human operator transfer function and C(z) is the plant's transfer function respectively. From (13),

$$G_H(z)C(z) = \frac{-0.240 + 0.478z^{-1}}{1 - 1.530z^{-1} + 0.532z^{-2}} \qquad (14)$$

The z transform may be converted into a Laplace transform using the following relations:

$$L\,(e^{-at}) = \frac{1}{s + a} \qquad (15)$$

$$Z\,(e^{-akT}) = \frac{z}{z - e^{-aT}}$$

where T is the sampling interval.

In Laplace transform, equation (14) becomes:

$$G_H(s)C(s) = \frac{0.240\,(13.39 - s)}{(s + 0.04)\,(6.27 + s)} \qquad (16)$$

The plant's dynamics is given as $C(s) = \frac{1}{s}$, therefore

$$G_H(s) = \frac{0.240s\ (13.39 - s)}{(s + 0.04)\ (6.27 + s)} \qquad (17)$$

The terms of s and (s + 0.04) are fairly close and may be cancelled. Therefore an approximate model of the human operator is obtained in the following form:

$$G_H(s) = \frac{0.240\ (13.39 - s)}{6.27 + s} \qquad (18)$$

The results of this analysis for sampling intervals of 0.01, 0.02, 0.1 and 0.2 seconds are given in Table 1 for cases without measurement noise and in Table 2 with measurement noise.

In Table 1, the gain value increases from 0.047 to 0.24 as the sampling interval is increased, but it is very different from the given value of k = 5. The chosen pole and zero locations were -8.7 and 8.7, respectively. The sampling interval of 10ms yielded the best approximation to these values, i.e., -8.45 and 8.87. With increasing sampling interval the location for zero moved to 9.63, 13.38 and finally to 18.20 at 0.2 second sampling. The movement of the pole position was much smaller.

The transfer function models with measurement noise were much more complex and most of these could not be converted to the Laplace transform domain. In the Z-domain these transfer functions are significantly different from the equivalent zero noise cases.

TABLE 1. TRANSFER FUNCTION MODEL: WITHOUT NOISE

Sampling Interval	System $G_H(z)C(z)$	Human Operator $G_H(s)$	
		Exact Model	Approximate Model
10 ms	$\frac{-0.047+0.051z^{-1}}{1-1.918z^{-1}+0.919z^{-2}}$	$\frac{0.047(8.87-s)}{8.45+s}$	
20 ms	$\frac{-0.084+0.099z^{-1}}{1-1.852z^{-1}+0.853z^{-2}}$	$\frac{0.084(9.63-s)}{8.01+s}$	
100 ms	$\frac{-0.240+0.478z^{-1}}{1-1.530z^{-1}+0.532z^{-2}}$	$\frac{0.240s(13.38-s)}{(0.04+s)(6.27+s)}$	$\frac{0.240(13.38-s)}{6.27+s}$
200 ms	$\frac{-0.235+0.904z^{-1}}{1-1.348z^{-1}+0.353z^{-2}}$	$\frac{0.235s(18.20-s)}{(0.03+s)(5.18+s)}$	$\frac{0.235(18.20-s)}{5.18+s}$

	TABLE 2. TRANSFER FUNCTION MODEL: WITH NOISE		
Sampling Interval Noise S.D.	System $G_H(z)C(z)$	Human Operator $G_H(s)$	
		Exact Model	Approximate Model
10ms S.D.=0.25	$\frac{-1.018+2.090z^{-1}-1.075z^{-2}}{1+2.084z^{-1}+1.078z^{-2}+0.004z^{-3}+0.001z^{-4}}$		
100ms S.D.=0.25	$\frac{0.259-1.244z^{-1}+2.077z^{-2}}{1+0\ 913z^{-1}-2.416z^{-2}-0.013z^{-3}+0.465z^{-4}}$		
200ms S.D.=0.25	$\frac{-1.700+0.750z^{-1}+0.390z^{-2}}{1-1.001z^{-1}-0.005z^{-2}}$		
100ms S.D.=0.5	$\frac{0.637-2.403z^{-1}+3.234z^{-2}}{1-1.615z^{-1}-3.167z^{-2}-0.077z^{-3}+0.432z^{-4}}$		
200 ms S.D.=0.5	$\frac{-0.142+0.722z^{-1}+0.373z^{-2}}{1-1.039z^{-1}+0.045z^{-2}}$	$\frac{s(0.142(99.17-s)+6.08e^{-0.2s})}{(0.03+s)\quad(15.5+s)}$	$\frac{0.142(99.17-s)+6.08e^{-0.2s}}{(15.5+s)}$

Although the time series analysis is a powerful and reliable technique, the sampling interval and measurement noise will influence the model order and the parameter values obtained. Modeling of the closed-loop systems makes the problem even more difficult.

REFERENCES

(1) Sheridan, T. B., and Ferell, W. R., Man-Machine Systems: Information, Control, and Decision Models of Human Performance, M.I.T. Press, Cambridge, MA, 1974.

(2) Shinners, S. M., "Modeling of Human Operator Performance Utilizing Time Series Analysis", IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-4, No. 5, pp. 446-458, September, 1974.

(3) Taylor, Jr., L. W., "A Comparison of Human Response Model in the Time and Frequency Domains", Third Annual NASA-University Conference on Manual Control, NASA SP-144, pp. 137-153, 1967.

(4) Elkind, J. I., and Forgie, C. D., "Characteristics of the Human Operator in Simple Manual Control System", IRE Transactions on Automatic Control, Vol. AC-4, No. 1, pp. 44-55, May, 1959.

(5) Young, L. R., Green, D. M., Elkind, J. I., and Kelly, J. A., "Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control", IEEE Transactions on Human Factors in Electronics, Vol. HFE-5, pp. 6-13, 1964.

(6) McRuer, D. T., and Jex, H. R., "A Review of Quasi-Linear Pilot Models", IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No. 3, pp. 231-249, September, 1967.

(7) Tanaka, K., Goto, N., and Washizu, K., "A Comparison of Techniques for Identifying Human Operator Dynamics Utilizing Time Series Analysis", Proceedings of the Twelfth Annual Conference on Manual Control, NASA TM X - 73170, pp. 673-693, 1976.

(8) Suen, L. C., and Liu, R., "A Normalized Residual Criterion for the Determination of Structure of Linear Systems", Policy Analysis and Information Systems, Vol. 1, pp. 1-22, July, 1977.

(9) Agarwal, G. C., Osafo-Charles, F., and O'Neill, W. D., "Modeling on Human Operator Dynamics in Simple Manual Control Utilizing Time Series Analysis", Sixteenth Annual Conference on Manual Control, pp. 1-28, M.I.T., May, 1980.

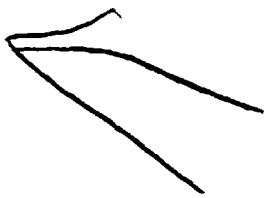
(10) Jex, H. R., and Allen, R. W., "Research on a New Human Dynamic Response Test Battery", Sixth Annual Conference on Manual Control, pp. 743-777, April, 1970.

(11) Shirly, R. S., "A Comparison of Techniques for Measuring Human Operator Frequency Response", Sixth Annual Conference on Manual Control, pp. 803-869, April, 1970.

(12) Stark, L., Iida, M., and Willis, P. A., "Dynamic Characteristics of the Motor Coordination System in Man", Biophysical Journal, Vol. 1, pp. 279-300, 1961.

(13) Astrom, K. J., and Eykhoff, P., "System Identification -- A Survey", Automatica, Vol. 7, pp. 123-162, 1971.

(14) Osafo-Charles, F., Agarwal, G. C., O'Neill, W. D., and Gottlieb, G. L., "Application of Time-Series Modeling to Human Operator Dynamics", IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-10, pp. 849-860, 1980.



2

SESSION 2: PHYSIOLOGICAL MODELS; WORKLOAD AND PERFORMANCE MEASURES

Moderator: Ronald O. Anderson, Air Force Wright Aeronautical Laboratories

AD P001619

CHANGES IN HUMAN JOINT COMPLIANCE DURING PERIPHERAL ISCHEMIA

by
Robert J. Jaeger, Gyan C. Agarwal, and Gerald L. Gottlieb
Department of Physiology, Rush Medical College
Chicago, IL 60612
and
Bioengineering Program, University of Illinois
Chicago, IL 60680

ABSTRACT

Joint compliance was measured at both the wrist and ankle by applying various torque waveforms (step, sinusoidal, and random) about the joint while measuring the resulting joint angle. Stretch-evoked electromyographic activity (EMG) was recorded from agonist-antagonist muscle pairs at the joint being tested. Measurements were made on normal human subjects before and after the development of peripheral ischemia in the limb. Ischemia is thought to preferentially block conduction in large diameter nerve fibers. Before ischemia developed, feedback from primary muscle spindles was intact. As ischemia developed, this feedback was reduced or eliminated.

Under normal conditions, the compliance measured by sinusoidal techniques is dependent on input amplitude and the resting level of contraction in the muscle. However, for any given mean level of contraction and RMS level of perturbation, the compliance is often well fitted by a second order linear system. When ischemia has developed, changes in compliance were seen at low frequencies. In terms of a second order mechanical system model, these changes were observed in the stiffness and viscosity parameters, and not in the moment of inertia. The observations are strongly dependent on the level of muscle fatigue during the the development of ischemia. There was also a concomitant loss of stretch-evoked EMG during the ischemia. The functional role of stretch-evoked EMG has always been controversial. These data suggest a relationship between mechanical characteristics of the joint and the stretch-evoked EMG, and quantify the mechanical contributions stretch-evoked (i.e. reflex) responses make to joint compliance.

INTRODUCTION

In studying man-machine interactions one is concerned with a number of issues. One of these is how quickly a human operator can voluntarily respond to disturbances in the system. Another concerns determining a mathematical transfer function of the human operator which also incorporates the

"involuntary" component of the response. Both of these issues are difficult to resolve because of the complexities of the system. The particular complexity of the human operator we will focus on in this paper is one of the internal feedback systems within the human operator - the muscle spindle loop. One way in which the operation of this system may be observed and studied is by the stretch-evoked myotatic response, and possibly the resulting mechanical responses.

We have attempted to study this system in both the normal, closed loop case and in an altered case with the muscle spindle feedback pathway disabled. In the latter case, this was done by inducing ischemia in the limb to block conduction in large diameter afferent nerve fibers from the spindles.

With regard to how quickly a human operator can respond to perturbations in the system which manifest themselves in sudden movements of the control manipulandum, it is clear that reflex responses have the potential to act most rapidly. There has been considerable controversy regarding the function of early stretch-evoked EMG responses. An early view, proposed by Merton (1953) was that stretch evokes a myotatic response in the muscle, mediated by the primary muscle spindle, which tends to return the muscle to its original length. This is the classic concept of load compensation, which states that length is the regulated variable. While this hypothesis has fallen out of favor due to various demonstrations that length is not "well regulated", it must be kept in mind that the design criteria and the definition of good performance for the human motor system are not available to us.

Another view is that the myotatic response is produced by motor units that were about to fire irrespective of the stretch stimulus, and as such, this response is an "epiphenomenon of little physiological consequence" (Hammond et al, 1956). This view was strengthened by the observation that the myotatic response in the soleus produced insignificant muscle tension compared with the later response, which was termed the "functional stretch reflex" (FSR). The implication here was that the myotatic response was not functional (Melvill Jones and Watt, 1971).

However, anatomical structures, such as the muscle spindle and the monosynaptic reflex arc, rarely exist without a specific physiological function or functions. While the first of the two preceeding views suggested the function of length regulation, the second rejected any functional role for the myotatic response. The truth probably lies somewhere in between. Much of the theoretical thinking concerning the human motor system has been guided by supposed evolutionary considerations which suggest that the monosynaptic reflex arc is phylogenetically old, and that higher centers have evolved around it. Nauta and Feirtag (1979) have speculated that the monosynaptic reflex arc may be relatively new on the presumed evolutionary scene. This speculation implies some important functional role for the muscle spindles and the myotatic

response, but what that might be remains unclear.

One of the more recent hypotheses concerning the function of early stretch evoked EMG activity is that this activity regulates the internal properties of the muscle, specifically its stiffness. In this hypothesis, segmental mechanisms act to maintain the stiffness of the muscle in the face of changing external loads and descending motor commands, rather than simply regulating the length of a muscle at some set point (Houk, 1979).

With respect to determining a correct mathematical transfer function of the human operator, attempts to model and simulate the neuromuscular system have always been frustrated by the lack of adequate block diagrams of the system. This reflects our lack of understanding of the neural circuitry involved. For example, attempts at modeling ballistic movements to determine the role of spindle feedback have yielded conflicting results (Dijkstra and Denier van der Gon, 1973; Van Dijk, 1978). It is for this reason that functional changes in the presence and absence of spindle feedback were studied here using only torque input and joint angle output rather than resorting to neuroanatomical block diagrams.

Studies which address the function of stretch-evoked EMG activity involve careful quantification of the EMG, torque perturbation, and joint rotation. Attempts to determine the biomechanical properties of different joints have used a variety of torque waveforms as inputs and measured the joint angle as the output. At a minimum, this requires a linear second order model which includes moment of inertia, viscosity, and stiffness terms. Compliance is shown in the following equation:

$$\text{Joint compliance} = \frac{\theta}{\tau} = \frac{1}{JS^2 + BS + K}$$

where J=moment of inertia, B=angular viscosity coefficient, K=angular stiffness, s=complex radian frequency, θ=joint angle, and τ=torque. For the totally denervated case, such a simple second order model would be a good first choice. Torques applied to the joint would be resisted by the passive inertia of the joint and the visco-elastic stiffness of the joint, muscles and connective tissue.

In the intact system, however, neural control loops exist which cannot as yet be precisely specified, and are superimposed on the mechanical system of muscles, bones, and other tissues. This introduces the distinct possibility that the system being modeled is changing with time, and this would be reflected in the model parameters. An example of this is the observed change in joint compliance seen with changes in muscle tone (Agarwal and Gottlieb, 1977). Nevertheless, it still may be possible to apply a second order linear model, with the understanding that the stiffness and viscosity so

determined are not directly comparable to the underlying biomechanical system in the denervated case, but rather represent some type of "equivalent" viscosity and stiffness.

Another problem with the simple second order model arises in the situation when voluntary intervention is present. In this case, additional torques, which are difficult to measure, are applied about the joint, after an appropriate reaction time. Reflex contractions which precede a voluntary response may also produce significant torques about the joint. Finally, it must be noted that the properties of muscle vary with the length and velocity of contraction. Thus the B and K values would again be expected to change. This implies possible amplitude and frequency dependent nonlinearities.

METHODS

Subjects: Normal human subjects between the ages of 22 and 40 were used. All were in good health with no current medical problems. The experiments were approved by the Rush Presbyterian-St. Luke's Committee on Human Investigation and informed consent was obtained.

Ischemia: Ischemia was induced by applying a standard arm or leg sphygmomanometer cuff around the upper part of the limb and rapidly inflating it to a pressure of 150 mm of mercury. Typically, the duration of the ischemia was 25 to 35 minutes. In no case was ischemia induced in the same limb of a given subject without a minimum recovery period of seven days. Following collection of ischemic data, the cuff was rapidly deflated by "popping" the valve of the cuff and allowing it to passively deflate, and recovery data was taken.

The goal of this procedure was to investigate the role of afferent signals upon motor outputs. Consequently a detailed study of altered sensations was not done since it has previously been reported (e.g. Lewis et al, 1937; Sinclair, 1951) that under most conditions sensation is lost in the following order: touch, position, temperature, and pain. Our subjects sometimes reported numbness in the fingertips, dulled pinprick sensations, difficulty in knowing exactly where the joint was without visual feedback, and the "pins and needles" sensations following cuff removal (Lewis et al, 1937). In some subjects we elicited tendon tap responses before, during and after the ischemia, in order to verify the absence of the tendon tap response after ischemia had progressed.

Apparatus: The devices used to move the joints have been previously described for the ankle (Agarwal and Gottlieb, 1977) and the wrist (Jaeger et al, 1982). They are shown schematically in figure 1. Briefly, a torque motor was used to apply step or sinusoidal torque waveforms to the joint. The

resulting angular position and stretch-evoked electromyographic activity was monitored. All experiments were performed under the control of a real time computing system.

Sinusoidal Oscillation: The methods used and the parameter estimation techniques have been described by Agarwal and Gottlieb (1977a, 1977b) and Gottlieb and Agarwal (1978). Briefly, sinusoidal torques were applied to the wrist or ankle, and the resulting joint angles were measured. The ratio of the peak-to-peak joint angle to peak-to-peak torque amplitudes are computed for each of several frequencies. The second order fit is computed to minimize the following error criterion:

$$\text{Error} = \sum_f \left[\log \left(\frac{\text{Model compliance}}{\text{Measured compliance}} \right) \right]^2$$

In the wrist experiments a slight tapering of input amplitudes was introduced to prevent excessive excursion of the wrist at the lowest frequencies. Data collected in this way is qualitatively similar to data collected without the taper, except that the wrist appears slightly stiffer at low frequencies than it actually is.

Step Torque Perturbation: The motor applied one-second duration steps of torque to flex or extend the joint. These steps were typically randomized in amplitude and interstep interval. Subjects were instructed to either do nothing when the step was percieved, or react to restore the joint to its' starting position. Responses were grouped by perturbation amplitude and averaged.

RESULTS

Normal Compliance Measurements: Compliance measurements were made using sinusoidal torque inputs over a given frequency range (0.5 to 10.0 hz for the wrist, 3.0 to 12.0 hz for the ankle). These measurements typically showed an underdamped second order response. Compliance curves for the wrist were constructed at three different input amplitudes using five different bias torques, and are given in figure 2. The parameters J, B, and K for these curves are shown in table 1. The compliance measurement was very sensitive to input amplitude, and showed large decreases in K, the stiffness, as larger amplitude sinusoidal torques were used. The compliance measurements were also sensitive to bias torque, showing increases in K with increasing bias torques. Less pronounced but consistent changes were seen in B, the viscosity term. The

moment of inertia was unchanged. Similar results were obtained for the ankle (Agarwal and Gottlieb, 1977a,b). These compliance measurements confirm that the biomechanical properties of the wrist joint determined in this manner depend strongly on input amplitude, but remain reasonably approximated by a second order system.

Ischemic Compliance Measurement: Compliance measurements were made as described in the previous section while ischemia was developing in the limb. Figure 3 shows three control wrist compliance curves and three compliance curves from ischemic data. Fitting a second order linear model to the data suggests that both K and B are decreasing to about 65 percent of their control values, as shown in table 2. As ischemia develops, EMG activity normally evoked by the sinusoidal stretching was reduced or abolished. There is some problem in interpretation of these results, as it takes six minutes to collect the data required to generate one compliance curve. During onset and recovery of ischemia, six minutes is a significant period of time. The system changes over this interval, and this must be taken into account when interpreting the data.

In more recent experiments, we have observed differences in results if we perform the experiments by first collecting data before ischemia, allowing the limb to rest during ischemia up to the point when the tendon tap response is abolished, and then taking only one or two sets of ischemia data. Compliance curves for the ankle for one such experiment is given in figure 4, with the second order fit parameters given in table 3.

In a second order linear model, the decrease in K and B as the stretch-evoked EMG is abolished suggests that at very low frequencies, a given amplitude of torque will produce a greater deflection of the joint, and the resonant frequency will decrease.

Transient Responses - Normal versus Ischemia: Transient responses to step torques under normal and ischemic conditions were compared at the wrist and the ankle. Figure 5 shows the transient response of the ankle joint to step torque and the stretch-evoked EMG as ischemia develops in the lower leg. The instruction to the subject was "do not react" to the torque. Note that the final value of joint angle and the damping increase. Phase plane trajectories are shown in figure 6. Cooke (1980) has discussed the merit of observing the phase plane trajectories of perturbations imposed on joint movements. Note the behavior of the trajectory between 110 and 150 ms as ischemia develops. In the pre-ischemic trajectories, the joint begins to return to its starting position, while in the ischemic case the joint comes to rest during this time. The additional overshoot and lack of oscillation in the ischemic case are also seen in the phase plane trajectory. The myotatic response occurred during the 40-60 ms interval. The

mechanical consequences of the myotatic response would not be apparent until about 30 ms after the EMG.

One interpretation of these results would then be that the myotatic response is in fact functional, and provides for the return of the joint about 30-50 percent towards its starting position, at a latency appropriate for additional action by subsequent responses. While this is most apparent at the ankle, it is also observed at the wrist.

For the instruction "react proportionally to the torque", data from the ankle are shown in figures 7 and 8. Trajectories are similar to their "do not react" counterparts to about 150 ms, at which time results of the later stretch-evoked EMG are apparent in figure 8, showing full return to the starting position.

At the wrist, responses for the "do not react" instruction showed saturation (that is, movement to an extreme of joint position for a prolonged period of time) upon application of all but the smallest torque steps. The "do not react" data were therefore not included in the analysis. For the react instruction, transient responses and phase plane trajectories are shown in figures 9 and 10. Note the differences in trajectories compared to the ankle. This may be due to the fact that the soleus has a very strong, isolated myotatic response, whereas the wrist has a rather weak myotatic response which begins a continuum of subsequently larger responses. The responses as ischemia progresses show an increased overshoot and a delayed return to the original position. The ischemic trajectory begins to diverge from the control trajectory at about 50 ms, a time consistent with the loss of early stretch evoked activity.

DISCUSSION

Changes in Mechanical Properties with Ischemia: In both the sinusoidal and transient response experiments, changes in mechanical properties were seen as ischemia developed, with the concomitant loss of stretch-evoked EMG activity. While correspondence of EMG activity for the two inputs used is uncertain, it is apparent that in both cases it depends on large diameter afferents. In the sinusoidal case a decrease in K is observed with ischemia. Similar behavior is also seen in the "do not react" transient case, where a constant torque produces larger steady state values of joint angle as ischemia progresses. It appears that muscle becomes more compliant (less stiff), presumably due to reduced tone, when the spindle feedback is disrupted by ischemia.

Changes in B suggest that viscous damping also decreases during ischemia. This would suggest that greater velocities of joint rotation would be reached during the torque perturbation. The trajectories of the transient responses

reflect this.

Decreases in B (by itself) would also imply more oscillatory behavior in the step response but this is not observed. The changes in B are too small to show significant changes in the damping factor.

Comparing experiments in which the muscles were active throughout ischemia with those in which they were rested has led us to explore the notion that muscle properties may be modified by at least two mechanisms: one of which is associated with observable changes in synchronous EMG bursts, and the other which is not observable in the EMG.

REFERENCES

Agarwal, G.C. and Gottlieb, G.L. Oscillation of the human ankle joint in response to applied sinusoidal torque on the foot. Journal of Physiology (London) 268:151-176, 1977a.

Agarwal, G.C. and Gottlieb, G.L. Compliance of the human ankle joint. American Society of Mechanical Engineers Journal of Biomechanical Engineering 99:166-170, 1977b.

Cooke, J.D. The role of stretch reflexes during active movement. Brain Research 181:493-497, 1980.

Dijkstra, S. and VanderGon, J.J.D. An analog computer study of fast isolated movements. Kybernetic 12:102-110, 1973.

Gottlieb, G.L. and Agarwal, G.C. Dependence of human ankle joint compliance on joint angle. Journal of Biomechanics 11:177-181, 1978.

Hammond, P.H. The influence of prior instruction to the subject on an apparently involuntary neuromuscular movement. Journal of Physiology (London) 132:17P-18P, 1956.

Houk, J.C. Regulation of stiffness by skeletomotor reflexes. Annual Review of Physiology 41:99-114, 1979.

Jaeger, R.J., Gottlieb, G.L., and Agarwal, G.C. Afferent contributions to stretch-evoked myoelectric responses. Journal of Neurophysiology 48:403-418, 1982.

Lewis, T., Pickering, G.W., and Rothschild, P. Centripital paralysis arising out of arrested blood flow to the limb, including notes on a form of tingling. Heart 16:1-32, 1937.

Melvill Jones, G. and Watt, D.G.D. Observations on the control of stepping and hopping in man. Journal of Physiology (London) 219:709-727, 1971.

Merton, P.A. Speculations on the servo control of movement. in Wolstenholm, G.E.W., Ciba Foundation Symposium: The Spinal Cord, Little, Brown, Boston, pp 247-260, 1953.

Nauta, W.J.H. and Fiertag, M. The organization of the brain. Scientific American 241(3):88-111, 1979.

Sinclair, D.C. Observations on sensory paralysis produced by compression of a human limb. Journal of Neurophysiology 11:75-92, 1948.

VanDijk, J.H.M. Simulation of human arm movements controlled by peripheral feedback. Biological Cybernetics 29:175-186.

ACKNOWLEDGMENTS

This work was supported in part by grants from the National Institutes of Health and the National Science Foundation.

R.J. Jaeger is now with the Pritzker Institute of Medical Engineering, Illinois Institute of Technology, Chicago, IL 60616.

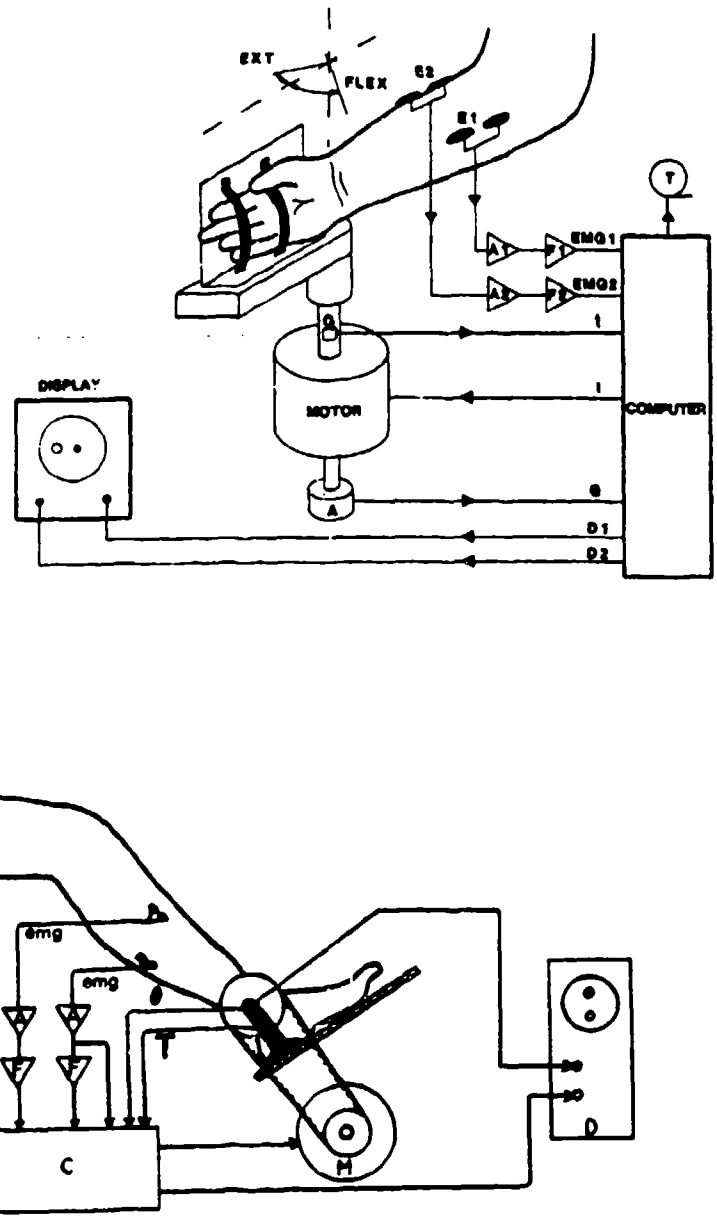


Figure 1: Schematic diagrams for the equipment used in torque perturbation experiments at the wrist and ankle.

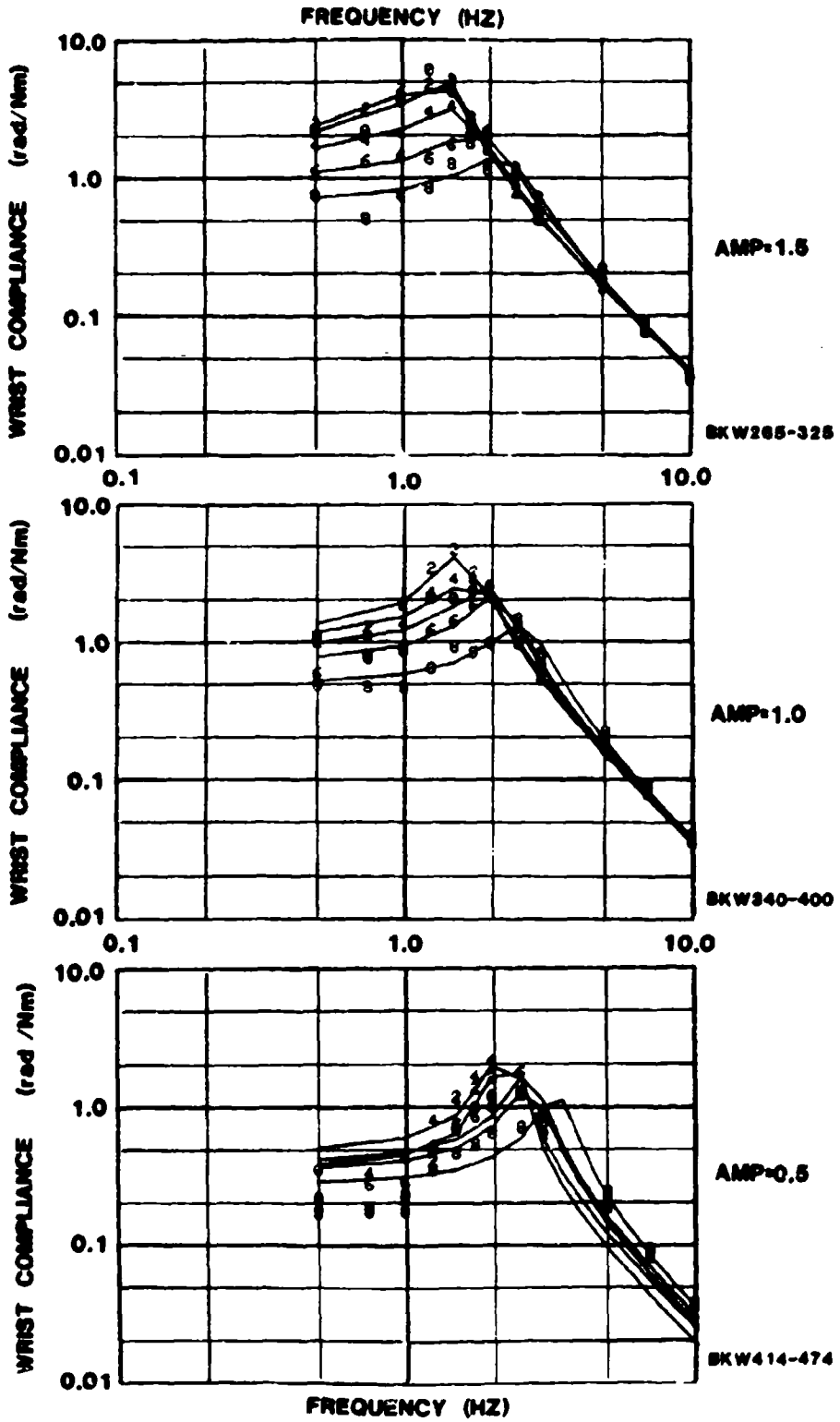


Figure 2: Wrist compliance curves at three input amplitudes. The family of 5 curves on each plot represents variations in bias torque.

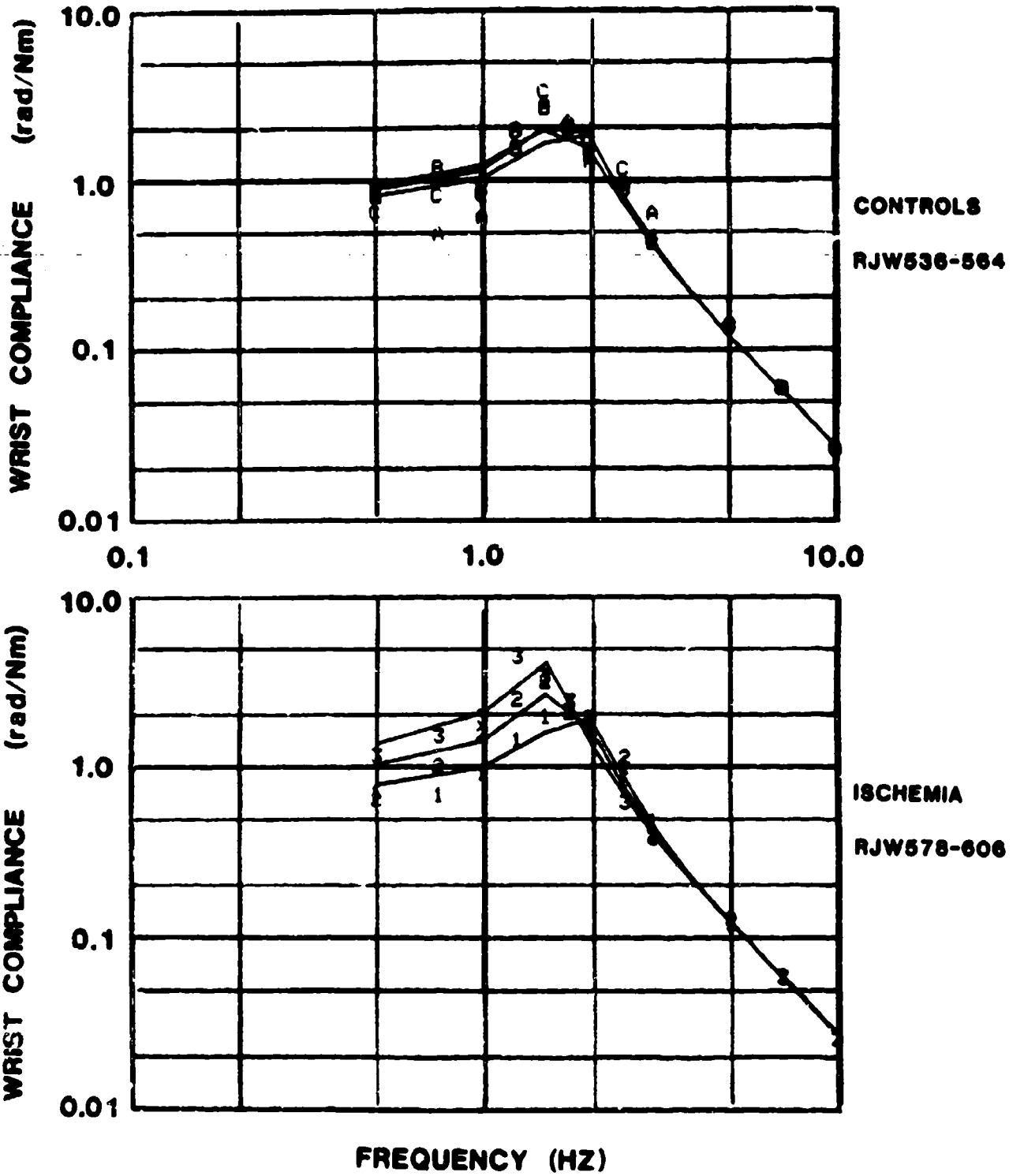


Figure 3: Wrist compliance curves under control conditions and as ischemia develops.

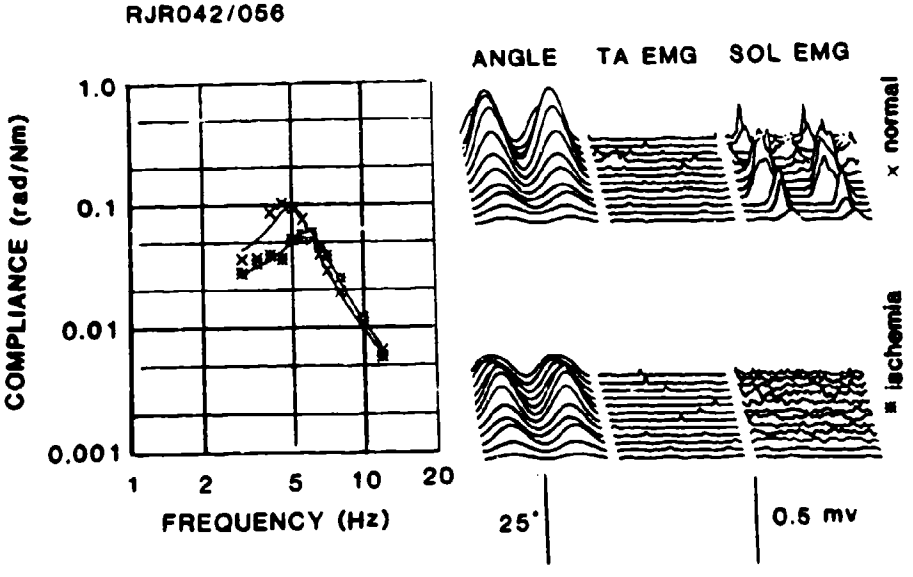


Figure 4: Compliance curves for the ankle before (x) and after ischemia (•). Joint angle and stretch-evoked EMG for tibialis anterior and soleus are also shown in the form of averages computed over two cycles. Background is three Hz, foreground is Hz. Horizontal scales are adjusted so that two cycles are displayed over the same plot range for each frequency.

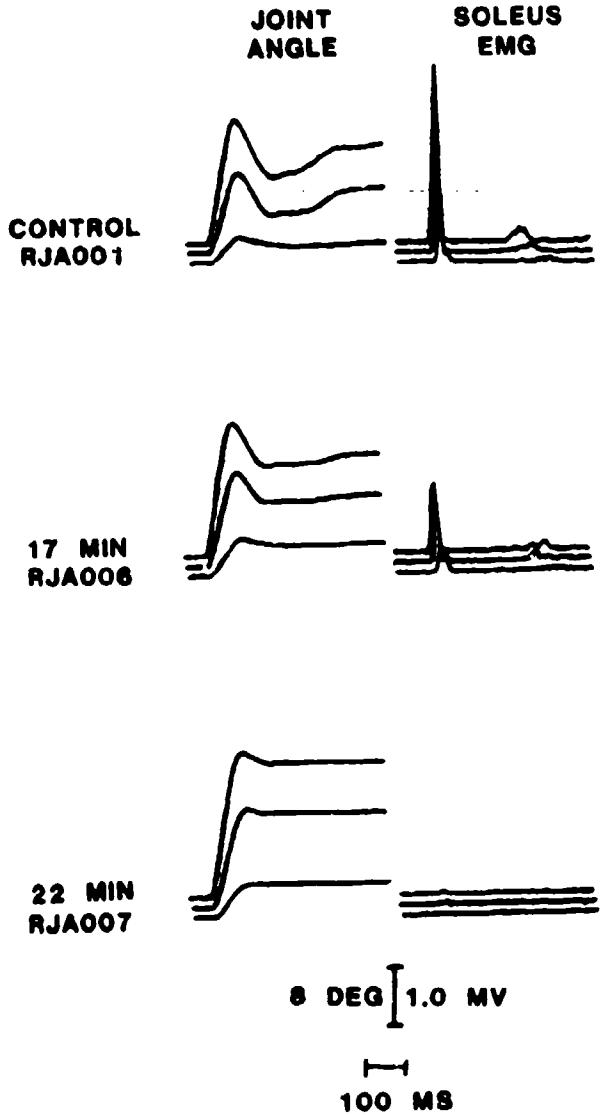


Figure 5: Transient response at ankle joint for "do not react" instruction. Soleus is stretched muscle. Each trace is the average of 10 responses. Three torque step amplitudes were used. Control, mid-ischemia, and final ischemia data are shown.

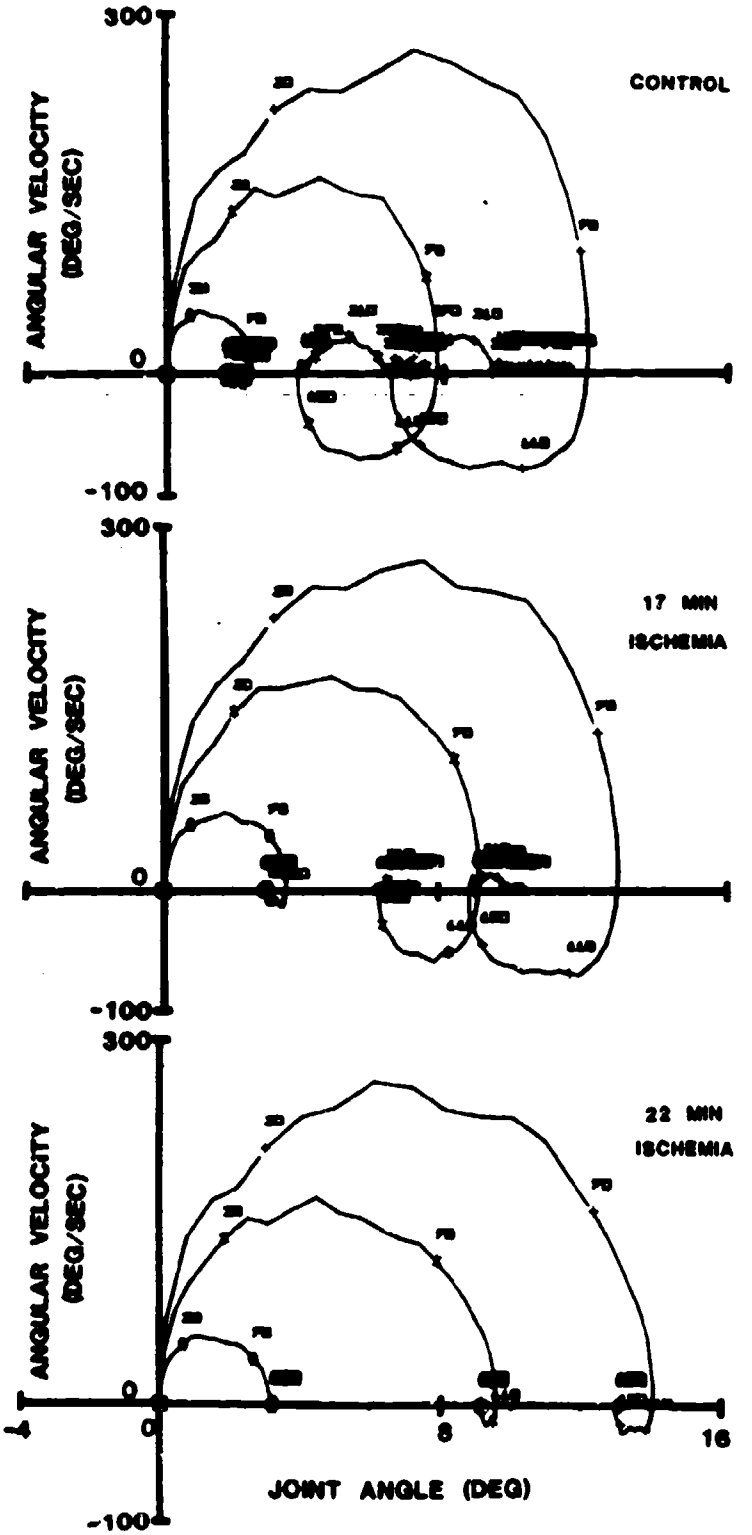


Figure 6: Phase plane trajectories for transient responses of figure 5.

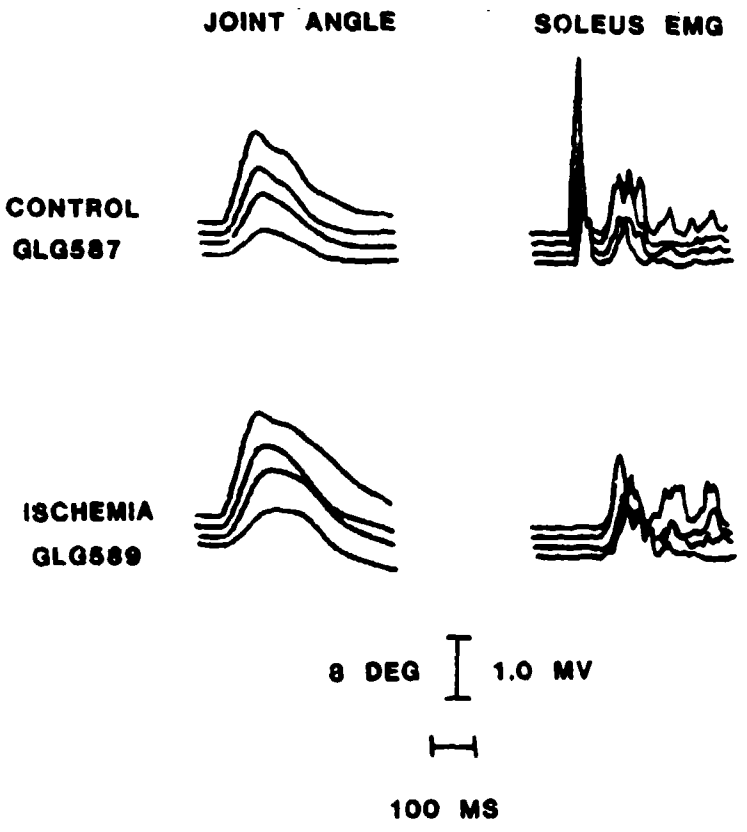


Figure 7: Transient responses at the ankle for the "react proportionally" instruction. Each trace is the average of ten responses. Four input amplitudes were used. Control and final ischemia data shown. Soleus muscle stretched.

THIS PAGE IS MISSING IN ORIGINAL DOCUMENT

Figure 8:
93

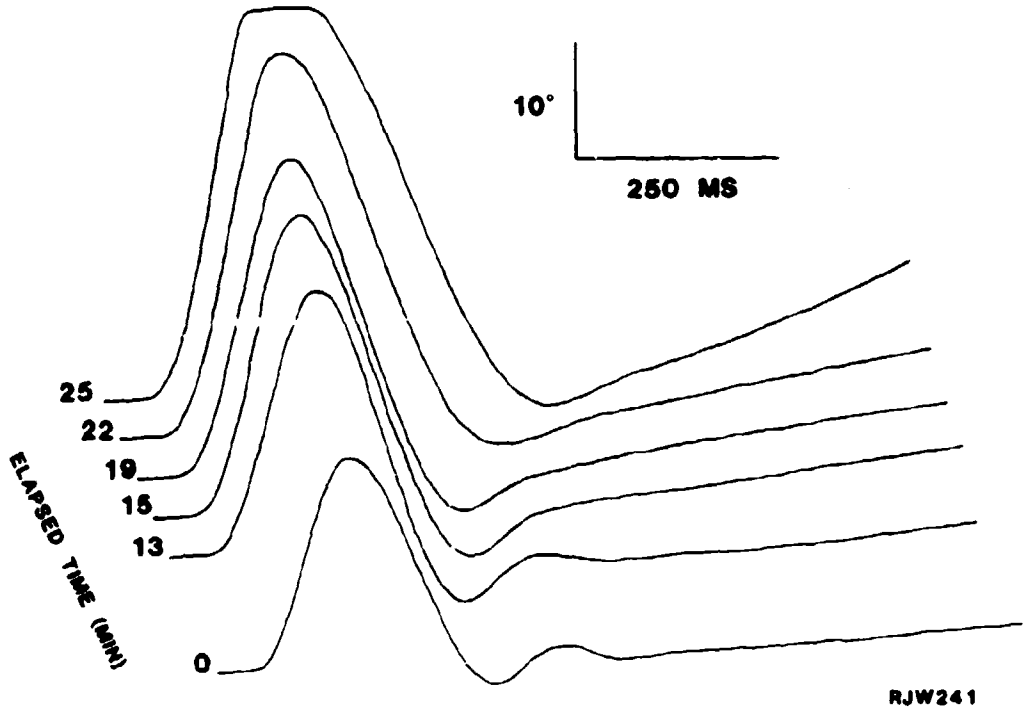


Figure 9: Transient responses at the wrist as ischemia develops. Each trace is the average of 10 responses. Step torque amplitude was constant. Time from onset of ischemia at right.

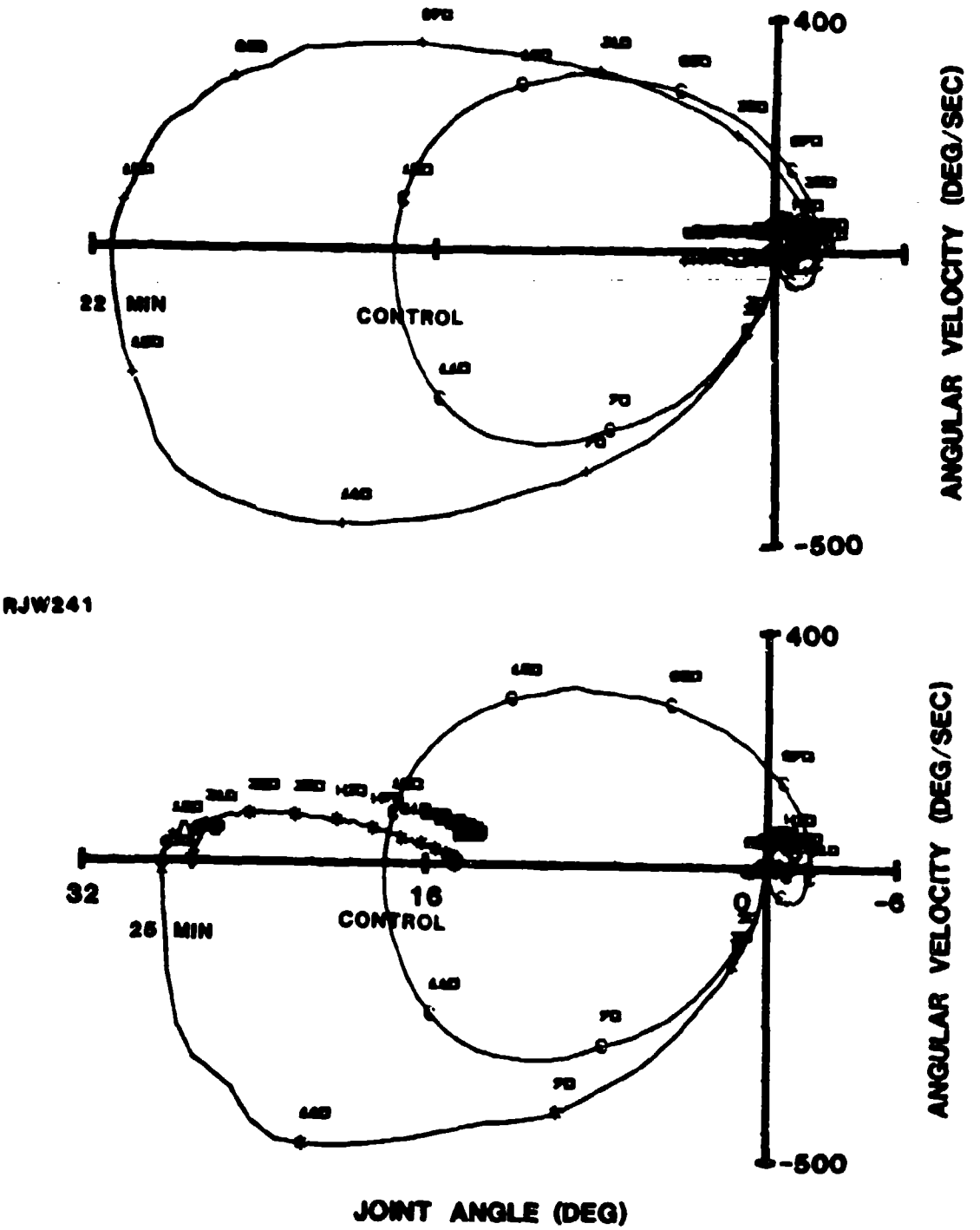


Figure 10: Phase plane trajectories for the data in figure 9.

AMP	BIAS	f (Hz)	ξ	J	B	K
0.5	0.0	2.70	0.131	0.0016	0.0074	0.4807
	0.2	2.27	0.100	0.0015	0.0030	0.3400
	0.4	2.24	0.077	0.0017	0.0039	0.3540
	0.6	2.68	0.115	0.0015	0.0058	0.4209
	0.8	3.36	0.115	0.0014	0.0065	0.6005
1.0	0.0	1.98	0.181	0.0012	0.0053	0.01857
	0.2	1.65	0.141	0.0013	0.0037	0.1365
	0.4	1.80	0.197	0.0012	0.0054	0.1559
	0.6	2.16	0.175	0.0013	0.0060	0.2308
	0.8	2.65	0.197	0.0012	0.0080	0.3389
1.5	0.0	1.41	0.179	0.0011	0.0036	0.0895
	0.2	1.35	0.189	0.0012	0.0037	0.0825
	0.4	1.58	0.246	0.0012	0.0056	0.1141
	0.6	1.94	0.255	0.0011	0.0069	0.1651
	0.8	2.33	0.251	0.0012	0.0085	0.2488

Table 1: Second order linear model parameters for normal compliance curves of figure 2. Units: J - newton·meter·second·second per rad, B - newton·meter·second per rad, K - newton·meter per rad.

RECORD	f (hZ)	ξ	J	B	K
CONTROL 1	1.86	0.182	0.0016	0.0070	0.2254
CONTROL 2	1.75	0.206	0.0016	0.0073	0.1965
CONTROL 3	1.81	0.164	0.0016	0.0060	0.2091
MEAN CONTROL	1.81	0.184	0.0016	0.0068	0.2103
0.6 MIN ISCH	1.89	0.174	0.0016	0.0068	0.2322
13-18 MIN ISCH	1.69	0.166	0.0016	0.0056	0.1787
24-30 MIN ISCH	1.49	0.150	0.0016	0.0045	0.1404

Table 2: Second order linear model parameters for the compliance curves in figure 3.

RECORD	f (hz)	ξ	J	B	k
CONTROL	5.40	0.147	0.0453	0.451	51.883
ISCHEMIA	6.00	0.188	0.0440	0.621	62.012

Table 3: Second order linear model parameters for the normal compliance curves of figure 4.

ABSTRACT

for

18th ANNUAL CONFERENCE ON MANUAL CONTROL

Title of Paper: Decision Making and the Steady State Visually Evoked EEG

Authors: Andrew M. Junker and Glen F. Wilson (AFAMRL/HEF, WPAFB, OH, 255-3325)

As part of the process of quantifying human/machine performance and interaction, it would be useful to have a noninvasive nonsubjective measure with which to assess aspects of human functioning such as attention, capacity, stress, etc. The Visually Evoked Response (VER) EEG possesses qualities that suggest it may be such a useful measure. The human does not have to consciously attend to the evoking stimulus and the response may be independent of human subjectivity. Further, it has become fairly routine to obtain EEG's due to improvements in measurement techniques and signal averaging computer technology. Of course, once the EEG is obtained, the real problem remains, one of answering the following questions: what does it indicate, is it in response to the visually evoking stimulus, is it repeatable, is there a linear portion to the response, is it consistent across subjects, etc. At the Aerospace Medical Research Laboratory we have undertaken a research program to explore the usefulness of the steady state VER and see what questions we can answer. To control the human's "internal environment" while making VER measurements, we are employing a supervisory decision-making task as developed by Pattipati and Kleinman at the University of Conneticut. The experimental configuration consists of humans performing the decision task which is displayed on a centrally located five-inch diagonal TV monitor while two fluorescent flash tubes above and below the monitor provide the visually evoking stimulus. At this time, some pilot studies have been performed. In this paper the experimental setup will be discussed, and decision-making performance results along with some of the potentially useful features observed in the EEG response will be presented.

AD P001620

OPERATOR WORKLOAD AS A FUNCTION OF THE SYSTEM STATE: AN ANALYSIS BASED UPON THE EVENT-RELATED BRAIN POTENTIAL

by
Richard T. Gill
Department of Engineering
Wright State University
Dayton, Ohio 45432

Christopher Wickens
Department of Psychology
University of Illinois
Champaign, Illinois 61801

ABSTRACT

While the processing demands of second order manual control are known to be greater than those of 1st or 0 order, the precise nature or locus of these increased demands is not well established. The purpose of this research is to determine if the demands are perceptual -related to the perception of higher derivatives of the error signal or characteristics of the system state, and thereby fluctuating with changes in these variables; or central. In the latter case, we assume the demand to be constant over time, a consequence of the increased demands of activating a more complex internal model. Event-related brain potentials -- more specifically, P300 amplitude -- were employed to assess operator workload while controlling a second order system. The ERP waveforms were categorized according to the system state at the time of the eliciting probes. Statistical analyses revealed no differences in P300 amplitude among the categories. Thus, it was concluded that the increased level of operator workload remained constant rather than fluctuating with changes in the system state. These results identify central processing rather than perception as the locus of higher order load.

INTRODUCTION

The difficulty in manually controlling higher order systems has long been recognized (1, 2, 3). The basis for this inherent difficulty may be attributed to several different sources. For example, the operator of a higher order system must act as a differentiator, generating lead in response to the perceived error signal; also there exists greater uncertainty in control because the required control input is not a one-to-one mapping of system error. Thirdly, double impulse control of second order systems requires a greater demand for accurate measurement of timing. The potential source of increased demand relevant to this research relates to the fact that the controller of higher order systems must perceive a greater number of system states and then combine them via a more complex internal model (2, 3,

4). For example, in order to optimally control a second order system, the operator must estimate both position and velocity, and then combine them via the internal model to determine the appropriate control input.

Furthermore, this increased difficulty in controlling higher order systems has also been shown, as would be expected, to result in greater operator workload (5, 6). More specifically, these investigations have suggested that the locus of this increased workload has been shown to reside in the perceptual and central processing stages of the operator and not in the response stage. The fundamental issue addressed in the present research concerns the temporal nature of this increased workload and the precise localization between perception and central processing. We hypothesize that if the source of higher workload relates to problems encountered in perceiving higher derivatives of the error signal, or particular combinations of state variables from the display, then workload should vary from moment to moment as a function of those variables. If, on the other hand, the source is attributed to the requirement to activate and maintain a more complex internal model in working memory, we anticipate that these increased demands will be more stable and constant for the duration of the tracking trial.

In this research, event related brain potentials (ERPs) were used to assess changes in operator workload. The ERP is a transient series of voltage fluctuations in the brain in response to some discrete stimulus or cognitive event. ERPs are recorded by the use of skin electrodes which are located at specific sites on the scalp. By time locking the EEG recording to a discrete eliciting stimulus event and then ensemble averaging, a very consistent pattern of positive and negative peaks or components will be observed (7). Extensive research has shown the amplitude of the P300 component to be inversely proportional to the operator's perceptual and central processing workload (5, 7, 8). P300 amplitude was used as a dependent measure of operator workload in the research described below.

EXPERIMENT

Five subjects were recruited from the university community and paid for their services. Their primary task was to minimize the RMS tracking error of a pure second order system in the presence of random noise. System error was displayed on a CRT (HP model 3010) and control inputs were made with the right hand via a spring loaded joystick. All subjects were requested to use a bang-bang or double-impulse control strategy (3) as such a strategy was essential for later analyses.

The experiment was conducted in two phases: Phase 1 - Training, and Phase 2 - Data Collection. The primary purpose of Phase 1 was to train subjects until they reached asymptotic performance. It consisted of 50 two-minute tracking trials, 25 trials per day for two consecutive days. Phase 2 was conducted on the third day during which subjects participated in 20 two-minute tracking trials. In addition to performing the primary tracking task, subjects were also required to perform a secondary task of counting discrete auditory tones or probes. The purpose of these probes was to provide the discrete event for eliciting the ERPs. They were presented, via headphones, in a Bernoulli series of low (P = .33) and high (P = .67) intensity. The

subject's task was to maintain a covert mental count of the number of low intensity probes presented and report this count at the end of each block.

ANALYSIS AND RESULTS

Each ERP waveform was selectively tagged according to the momentary state (error, error velocity, control position, control velocity) at the time of stimulus presentation. This enabled us to create selective averages of ERPs (with their associated P300 component amplitudes) according to any particular characteristics of system state. Two separate analyses were performed: one based solely upon system error and error velocity, each considered independently, and the other based upon a state-space representation.

Analysis 1:

Two dichotomies were employed to categorize ERPs, one based upon system error, and one upon system velocity. All waveforms were ordered in terms of the absolute value of system error at the time of the eliciting stimulus. Then waveforms above and below the median of this ordering were separately pooled and contrasted. A similar procedure was followed with regard to system error velocity.

A separate ANOVA was performed on the component loading scores, from the PCA, for each of the dichotomies. That is, P300 amplitude (as measured by its component loading score) elicited from probes when the absolute value of system error was above its median value was compared with the P300 amplitude when system error was below its median value. Similarly, the same comparison was made for the case when absolute system velocity was above and below the median value.

For both comparisons the results were the same, with no significant difference between high and low values. That is to say operator workload (as measured by P300 amplitude) did not vary in a systematic manner with the error variables. These results are consistent with our previous findings (5).

Analysis 2:

The rationale for the second analysis is based upon the subjects' employment of a bang-bang control strategy. Such a strategy requires the operator to maintain the maximum positive control input when attempting to eliminate a given error and then at the proper instant (as specified by an analytically determined optimal switching line) reverse the input to its maximum negative value, a relatively natural strategy often employed in 2nd order control (2).

If operator workload were to fluctuate during such a task, it is clear that it would be the greatest when the system state was in the "vicinity" of the operator's empirical switching line when a decision to control is impending. That is, workload would be higher when the operator was required to: (1) perceive the magnitude of the system error; (2) perceive the magnitude of the system velocity; (3) combine them via his internal model; and (4) decide if a control reversal was required. Alternatively, the workload would be lower when the system state was such that the control input was obvious.

For example, when system error and velocity are both of the same sign, a control reversal is obviously required since error is diverging. In short, the objective of the second set of analyses was to compare operator workload when the system state was in the vicinity of the operator's empirical switching line to that when the required control input was obvious.

An algorithm was developed that analyzed the subject's continuous control input and identified the system state at the time each control reversal was initiated. The locus of these control reversals was defined to be the "vicinity" of the subject's empirical switching line. Figure 1 is a graphical representation of the system state space on which these control reversals are plotted for a typical subject, along with the theoretically optimal switching line. After reviewing such plots for each subject, it was concluded that the general locus of control reversals for each subject was sufficiently similar that the same ERP categorization algorithm could be used for all subjects.

This algorithm is illustrated graphically in Figure 2. The rationale for each area is discussed below.

Area 1. This represents the region in the vicinity of the subject's empirical switching line. The boundaries were chosen conservatively to select probes that occurred when the system state was such that the operator would be required to sense both position and velocity, and then combine them (via his internal model) to determine the proper control input. This was assumed to be the area defining a state of greatest cognitive demands.

Area 2. This region represents a conservative selection of probes when the state was such that the subject's control input was obvious; that is, when the cursor was moving away from the target line. The response in this region should be a relatively automatic reversal.

Area 3. This represents a system state similar to Area 1, but in which the velocity, relative to the position, is too low to warrant a control reversal. This is, the cursor is moving towards the target line, but not fast enough to warrant a control reversal to prevent overshoot.

Area 4. This region is comprised of the remainder of the system state space which fails to meet the requirements of any of the preceding areas.

Areas 1 and 2 were defined in a conservative manner in order to maximize the potential differences between the two, thereby increasing the sensitivity of the subsequent analyses. Area 3 was selected to provide a condition in which workload, if workload does indeed fluctuate, would lie between that of Areas 1 and 2.

This sorting algorithm was applied to the single trial ERP data of each subject. Initially, the parameters were chosen very liberally, such that Area 4 was non-existent. The resultant ERPs from the three areas were then averaged across subjects. As before, an ANOVA was performed on the component loading scores for P300; no significant differences were observed. In order to increase the sensitivity of the test, subsequent analyses were done in

which the boundaries of Area 4 were systematically expanded by decreasing the parameter A_3 (see figure 2). That is, Areas 1 and 2 were defined in an increasingly conservative manner. The results remained unchanged, however, as no differences emerged with this manipulation.

CONCLUSIONS

The results of these two separate analyses were consistent. No matter whether ERP data were categorized via the magnitude of each state variable, relative to the mean, or whether the categorization was based on the location of the system state relative to the empirical switching line, differences in P300 amplitude between categorical levels failed to be observed. In fact, there were not even nonsignificant trends in this direction. These results provide strong support for the hypothesis that the increased resource demands of controlling second order systems remain relatively constant.

They are, of course, consistent with the hypothesis that it is the continuous demands of activating the complex internal model, rather than the momentary demands associated with changing perceptual variables that influence workload. There is, however, a second possibility that cannot be discounted altogether. It is possible that resource demand of control did, in fact, wax and wane, but that the "cost" associated with switching the allocation of resources to the auditory task was greater than the benefits derived. That is, it was more efficient to allocate a given amount of resources continuously to the tracking task, even though they would not all be used 100% of the time, than it was to vary the allocation of resources based on the instantaneous task demands. (It should be noted in this regard that, while P300 was consistantly attenuated, subjects did not ignore the tone counting. Their counts remained accurate.) It is impossible to determine from the current data which of the two hypotheses is correct. To do so would require running an additional condition in which greater priorities on the auditory task are stressed. When payoffs or costs thereby impell the operator to devote greater resources to the ERP task, we would now expect differential allocation (and therefore P300 amplitude) to the extent that the second hypothesis was in effect. If the first hypothesis were true, we would predict an overall depression in tracking performance and an increase in P300 amplitude, as the internal model was maintained with less fidelity. However, P300 would still fail to discriminate between system states.

ACKNOWLEDGEMENTS

This research was supported by a grant from the Air Force Office of Scientific Research Life Sciences Program F49620-79-C-0233. Dr. Alfred Fregly was the Technical Monitor.

REFERENCES

1. Birmingham, H.P., and Taylor, F.V.A. A human engineering approach to the design of man-operated continuous control systems. U.S. Naval Research Laboratory Report 4333. Washington, D.C., 1954.
2. Pew, R.W. Human perceptual-motor performance. In B. Kantowitz (Ed.), Human Information Processing: Tutorials in Performance and Cognition, Wiley and Sons, 1974.
3. Sheridan, T.B., and Ferrell, W.R. Man-Machine Systems: Information, Control, and Decision Models of Human Performance. Cambridge, MA, 1974.
4. Kelley, C. Manual and Automatic Control. NY: Wiley, 1965.
5. Wickens, C.D., Gill, R., Kramer, A., Ross, W., and Donchin, E. The cognitive demands of second order manual control: Applications of the event-related brain potential. Proceedings of the 17th Annual NASA Conference on Manual Control, NASA TM, 1981.
6. Wickens, C.D. and Derrick, W. The processing demands of higher order manual control: Application of additive factors methodology. University of Illinois, Technical Report EPL-81-1/ONR-81-1, 1981.
7. Donchin, E. Event-related brain potentials: A tool in the study of human information processing. In H. Begleiter (Ed.), Evoked Potentials and Behavior. New York: Plenum Press, 1979, pp. 13-75.
8. Donchin, E., Kramer, A., and Wickens, C. Probing the cognitive infrastructure with event-related brain potentials. Technical Report CPL82-2/AFOSR82-1, 1982.
9. Donchin, E. and Heffley, E. Multivariate analysis of event-related potential data: A tutorial review. In D. Otto (Ed.), Multidisciplinary Perspectives in Event-Related Brain Potential Research. Washington, D.C. U.S. Government Printing Office, EPA-600/9-77-043, 1979, pp. 555-572.

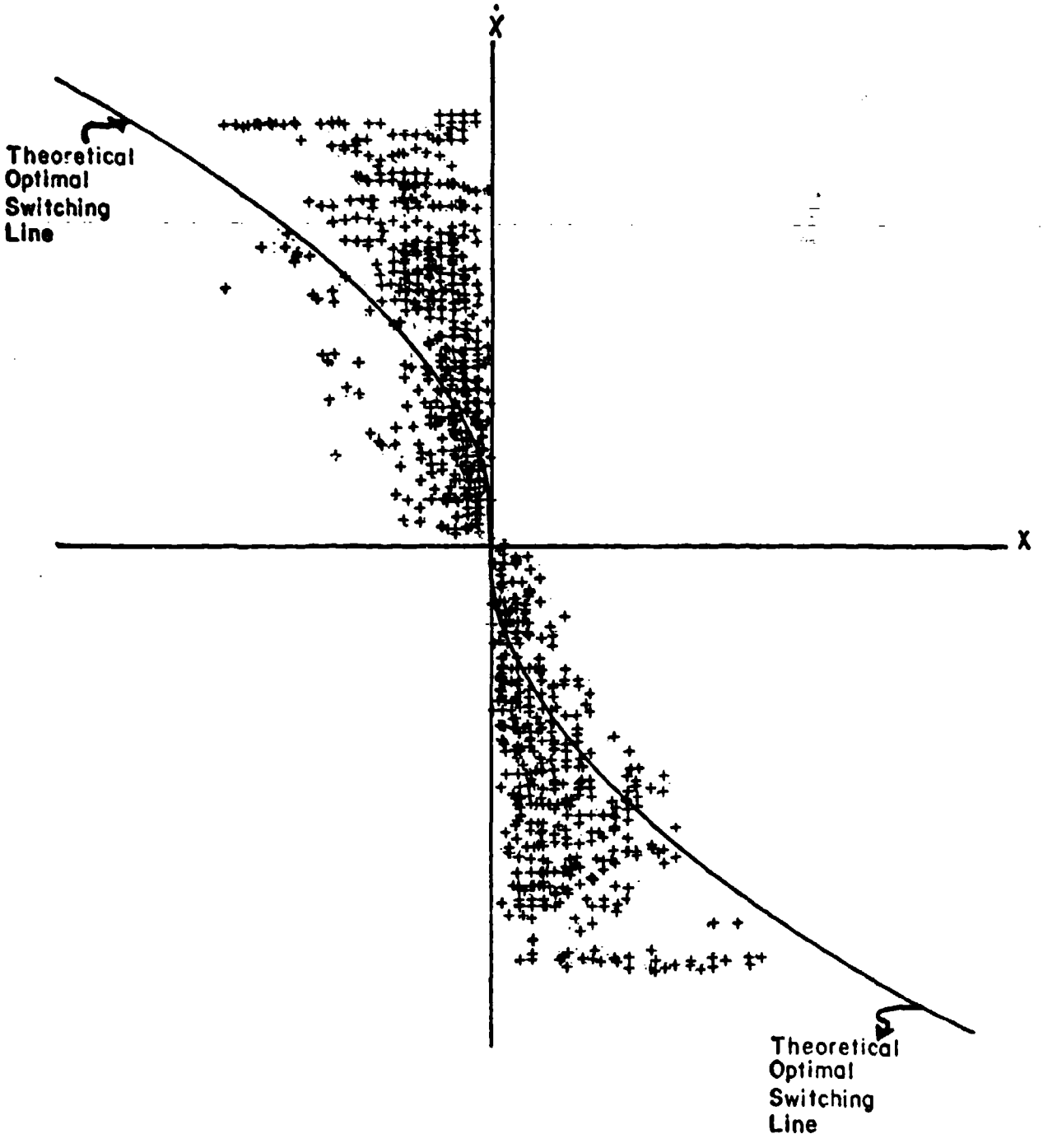


Figure 1: System state space -- typical plot of control reversals and theoretical switching line

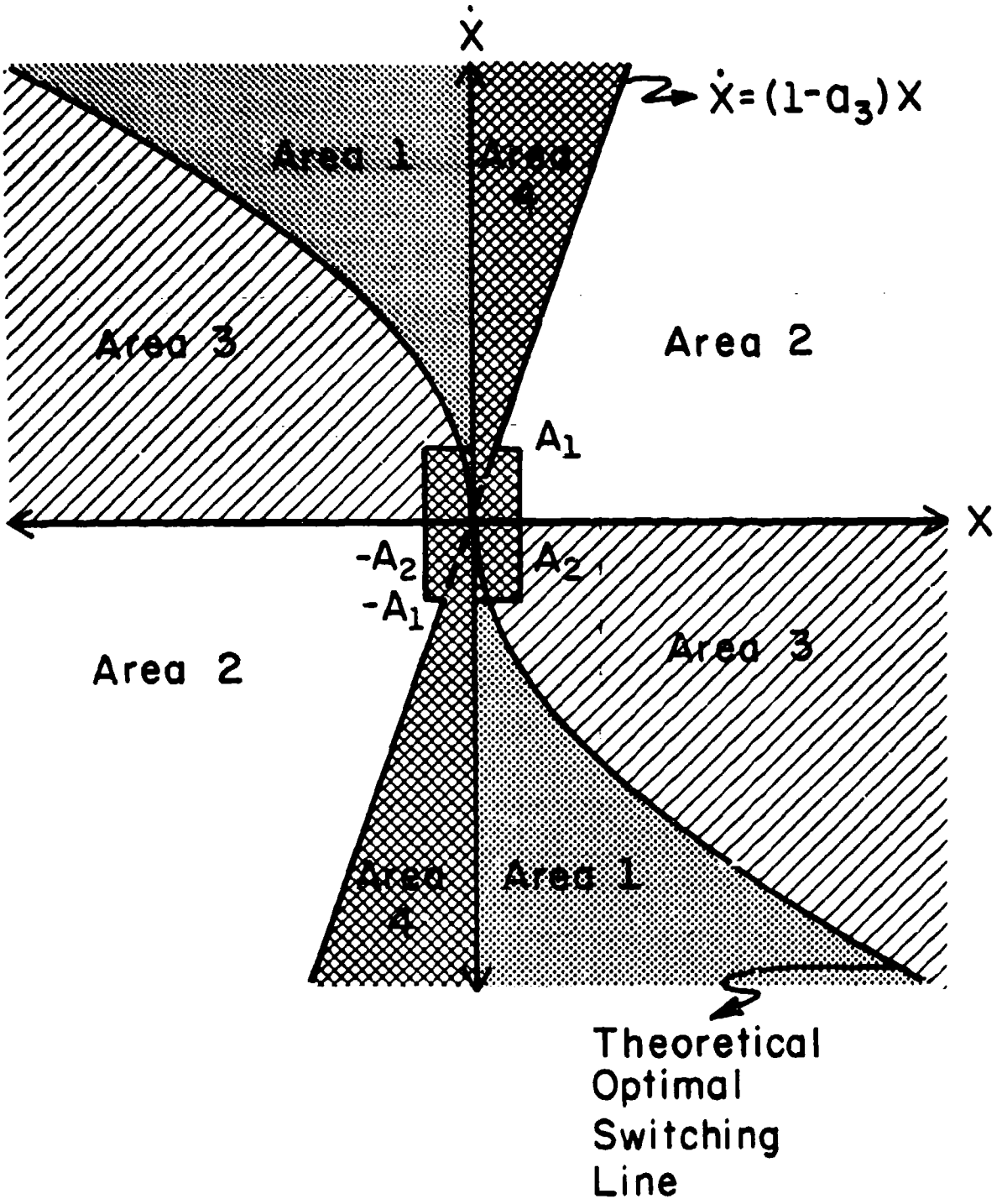


Figure 2: Algorithm for categorizing ERP probes based on location relative to empirical switching line.

AD P001621

ASSESSMENT OF MANUAL, PROCESS CONTROL STRATEGIES

by
Joe De Maio
Air Force Human Resources Laboratory
Operations Training Division
Williams AFB, AZ 85224

ABSTRACT

The present research is directed at developing a methodology for examining individuals' performance as process controllers. The process of immediate interest is maneuvering flight. In this task the pilot performs as an energy management decision maker, making moment by moment decisions regarding present and future aircraft energy states. A micro-computer-based task has been developed in which a subject "flies" a simulated "vehicle" through a course in a sequence of discrete moves. The range of movement available on each move is determined by generic, aircraft flight equations. Performance, as indexed by number of moves required to complete the course, has been found to be strongly related to flying ability. Additional performance measures, based on the decisions made on each move, have been developed which permit examination of control strategies underlying task performance. Such measures will permit analysis of cognitive functions leading to successful control performance and of the effects of environmental and individual factors on these decision making functions.

INTRODUCTION

A number of vehicle control tasks involve maintaining some predetermined steady state. Driving an automobile on a straight road, steering a ship on a constant heading and flying an aircraft at constant heading and climb rate are examples of such tasks. In these types of tasks the operator's function is primarily one of detecting error, in the form of deviations from the desired motion path, and correcting that error. For the most part the operator acts as a single closed-loop controller whose performance is determined primarily by his sensitivity to error and his control lag. (Jagacinski, 1977).

In a number of other tasks the operator's role is much more complex. He does not seek simply to maintain a steady state, but rather his goal is to achieve a particular vehicle state through execution of a sequential process of velocity changes. In these tasks the operator's actions are directed toward the goal state, which differs from his current state and from intermediate states. Intermediate vehicle states are important only insofar as they affect the ultimate vehicle state. Task sequences can be relatively short and uncomplicated, as in parallel parking a car, or quite long and complicated, as in the case of driving home in rush hour traffic in time to see the news on TV. In these instances the operator's performance is

determined both by his interface with the vehicle and by his performance as an information processor and decision maker.

Describing performance on these task sequences is complicated by the sheer number of control actions involved and by the fact that the optimum action a given point in the sequence is determined by earlier events both internal to the operator-vehicle system and external to it. For example the decision on whether to turn right at an intersection in order to get home by news time can depend on whether the vehicle will reach the intersection while the signal is green (internal, speed control), how long the red will last and how well traffic is moving on the two streets (external). The operator steers the vehicle according to a goal-directed scenario which he updates continuously in response to changing conditions. The correctness of any paticular control action in the sequence is determined by two factors: the correctness of the scenario and the appropriateness of the control action to the scenario.

At a macro level vehicular motion scenarios are very similar to scripts. A script is a set of related facts which is organized hierarchically and sequentially (Smith, Adams and Schorr, 1978). In the motion scenario facts might be events or control actions which are organized by the importance or probability of occurrences at particular points in the process. Schvaneveldt et al (in press) have shown experienced fighter pilots to organize major events and actions related to particular maneuvers in this way. Facts in these scripts are aircraft states, pilot actions, the effects of these actions, controls and instruments, and the actions of other combatants. These facts are organized primarily along a dimension of temporal criticality and secondarilty along a dimension of relative energy state.

At a finer level of detail, involving individual control inputs, vehicle control can be thought of as a similar sequential process of goal directed decisions and control actions. While the second by second control process may involve a sequence of decision events, it is less well suited to a scripts type analysis since decisions occur rapidly and may be based on information which is not organized semantically. This micro control process is similar to closed-loop tracking, but unlike tracking it consists of a sequence of actions directed at establishing a particular state rather than at simply minimizing error. In this control process several different control actions might be effective in achieving the desired state. The operator's decision about which actions to employ may depend on many factors, such as the degree and direction of motion required, the probability of success and the cost of failure associated with a particular control action sequence. In the real time motion control task it is difficult to isolate the micro-decision making process from perceptual and psychomotor factors. If, for instance, a pilot ends up with excessive airspeed on final, it is not possible to know whether the error arose from the pilot's failure to perceive the airspeed buildup or from an inability to make the appropriate control inputs or whether he saw what was happening and decided that everything was just fine.

One approach to isolating the micro-decision making process is to present subjects a discrete-time vehicle control task. In this task the subject "steers" a computerized vehicle through a course in a sequence of

moves. On each move the motion envelop is determined by the vehicle state at the end of the preceding move through a set of driving equations. The task is self-paced, and the subject types the move he desires on a keyboard. De Maio (1981) compared the performance of Air Force pilots and that of non-pilots on a discrete-time task in which subjects controlled vehicular movement by adjusting the length of two perpendicular driving vectors. As indexed by the number of moves required to complete the course, the pilot group performed better than the non-pilot group. The pilots also made longer moves on the average, although the average total track length was the same for both groups.

In the present research a discrete-time vehicle control task is used in which the driving equations more closely approximate those of an actual vehicle. Performance on individual moves is examined to determine the decision making process leading to superior task performance.

METHOD

Subjects

Twenty-three subjects were used. Eight of these were students in Undergraduate Pilot Training (UPT) at Williams AFB. Four had recently completed UPT. Eleven were civilian employees at the Human Resources Laboratory, who had no flying experience.

Apparatus

The Flight Decision-making Assessment Task (FDAT) is a discrete-time vehicle control task programmed on a Terak 8510/A micro-computer. The FDAT involves "steering" a vehicle through a winding course in a series of moves. The subjects' task was to traverse the course in as few moves as possible. The subject makes a move by entering two numbers. The first number entered is the amount of heading change desired, in degrees. Following this input the vehicle's velocity is decremented in proportion to the amount of heading change by an "induced drag" function. A thrust/drag function then determines the amount the subject may increase or decrease his velocity. The subject completes the move by entering the velocity desired, that is the length of the move. The ground track is then displayed on the course, and the new heading and velocity and associated maximum heading change and maximum induced drag penalty are shown.

The FDAT driving functions are generic flight equations for an aircraft constrained to level flight and zero angle of attack and scaled to fit on the 320 x 240 Terak graphics screen. The three functions which determine vehicle performance are: maximum heading change, induced drag velocity penalty, minimum/maximum velocity change.

A complete listing of the vehicle perrformance limits is given in Table 1. The "stall" or minimum velocity is 7 units, and the maximum velocity is 28 units. Velocity units are pixels. The course lengths are roughly 680 pixels. The maximum allowable heading change reached a minimum of 13^{o} at stall and 23^{o} at maximum velocity and a maximum of 48^{o} at the 14 unit corner velocity.

Procedure

The subject sat before the computer terminal with the experimenter seated at his left. The experimenter explained that the purpose of the research was to examine the decision making process associate with aircraft control. The experimenter then made four demonstration moves showing turning, acceleration and the effect of trying to exit the course boundaries. The subject then completed two courses, always performing the easier course first. The experimenter was present throughout the session. The subject was allowed to ask questions, but questions about strategies were not answered.

RESULTS

Overall Performance

Overall FDAT performance was given by the number of moves required to traverse each course. For the entire sample of subjects number of moves varied from 38 to 75 for Course 1 and from 40 to 87 for Course 2. One subject (the experimenter) was highly practiced at the FDAT. Discounting this subject the range was 44 to 75 for Course 1 and 50 to 87 for Course 2. There were no significant differences in overall performance between groups.

Internal Performance Analysis

Task requirements and the vehicles handling qualities placed a high premium on maintaining a velocity equal to or greater than 14 units. Due to the vehicles' low thrust/drag ratio it was not possible to sustain speed during intense cornering. As a result the subject needed to position himself on the course in such a manner as to permit acceleration and to time turns to minimumize heading changes and resultant loss of velocity. Speed control and timing errors were easily identifiable by visual inspection of individual moves and move sequences. In order to quantify speed control, velocity control and timing performance four categories of moves were defined:

Velocity control
- Extending moves - 5^0 turn accelerating
- Stagnating moves - $\geq$2 units deceleration or velocity $\leq$ 10 units

Turning control
- Sustained turn moves - 5^0-30^0 in constant dirction, roughly constant velocity
- Positioning moves - 5^0-30^0 only one move
  $\leq 5^0$ turn with 0-1 unit deceleration

Velocity control performance was given by the ratio of the number of extending moves to the number of stagnating moves. A measure of turning control performance was the ratio of the number of sustained turn moves to the number of positioning moves. A measure of timing was the total amount

of heading change regardless of direction, since on some, but not all, turns, mistiming necessitated corrective turns which increased the total heading change for the course.

In order to determine the contribution of speed control performance, turning control performance and timing performance to overall performance a multiple regression of the two ratios and the total heading change on number of moves for Course 1 and for Course 2 separately.

For Course 1 all three internal performance measures contributed significantly to overall performance (see Table 2). The primary driver of performance was speed control. Timing, as measured by total heading change, also contributed substantially to overall performance. The ratio of sustained turn to positioning moves measures the smoothness of heading control and made only a small contribution to overall performance.

The results of the regression analysis for Course 2 differed substantially from those for Course 1 (see Table2). Speed control performance accounted for over 85% of the variance in overall performance. Neither of the turn timing measures contributed to overall performance. The differing results for the two courses probably stems from a design flaw in Course 2. Course 2 contained 3 turns (out of 6) in which it was impossible or nearly impossible to avoid complete velocity stagnation. As a result speed control was more highly correlated with both overall performance and with turn timing performance on Course 2 than on Course 1 (see Table 3). In addition a much greater proportion of moves was related to speed control on Course 2 than on Course 1 (see Table 4).

DISCUSSION

The present work is based on the idea that the process of steering a vehicle from an initial energy-position state to a goal energy-position state can be described as a sequence of vehicle control decisions. Success in achieving the desired state will be a function of the quality of the control decisions made during the process sequence.

The Flight Decision-making Assessment Task used in this work is a discrete-time vehicle control task in which subjects are required to make a sequence of decisions regarding three aspects of vehicle control: velocity control, turning control and timing of turns. By examining internal performance indicators which tap these decision making functions, it is possible to quantify these decision functions and also to assess the relative importance of the decision making processes to overall task performance. The present results show that overall performance can be described by a linear combination of speed control, turning control and turn timing performance. When the FDAT course was well designed, as in Course 1, the contribution of speed control in the regression accounted for 37% of the variance in overall performance. The contribution of turn timing, as indexed by total turn, accounted for 27% of the variance in overall performance. Turning smoothness, as indexed by the ratio of sustained turn to positioning moves, had the smallest contribution to overall performance, accounting for 15% of the overall performance variance. However care must

be used in designing discrete-time tasks to insure that there is sufficient independence in the effects of the decision processes to permit analysis of individual decision components. On Course 2 only speed control affected overall performance. This was probably due to the high proportion of stagnated moves on Course 2.

The ability to investigate individual decision process can be use to gain greater insight into the cognitive demands of vehicle control tasks. By varying task requirements, vehicle control dynamics or course design, it may be possible to change the relative importance of different decision components in determining overall task performance or the difficulty of making these decisions. An additional area of study, which has not been addressed inthe present work is decision making speed. The discrete-time methodolgy can provide a powerful tool for studying this problem by permitting the unconfounding of effects arising from requirements for speeded perception and speeded responding, and those arising from the decision making process.

Velocity	Maximum Heading Change, Degrees	Maximum Induced Drag Penalty	Velocity Change Minimum	Maximum
7	13	0	0	1
8	19	-1	-1	2
9	24	-2	-1	3
10	27	-3	-1	4
11	30	-4	-2	4
12	33	-5	-2	4
13	36	-6	-2	4
14	48	-7	-2	4
15	44	-7	-2	4
16	42	-8	-2	4
17	39	-8	-2	4
18	37	-8	-2	4
19	35	-8	-3	3
20	33	-9	-3	3
21	32	-9	-3	3
22	30	-9	-3	2
23	29	-9	-3	2
24	28	-10	-4	1
25	27	-10	-4	1
26	26	-10	-4	1
27	25	-10	-4	0

Table 1

Vehicle performance limits at each velocity. Performance is optimal around the corner velocity of 14.

Factor	Beta-weight Course 1	Beta-weight Course 2
Ext/Stag	-.46 *	-.93 *
SUS/POS	-.27 *	NS
TOT TURN	-.38 *	NS
R	.89*	.93*

* $p < .001$

Table 2

Multiple regression analysis internal performance measures on overall performance

	OVERALL	EXT/STAG	SUS/POS	TOT TURN
OVERALL	-			
EXT/STAG	-.79 (-.93)	-		
SUS/POS	-.55 (-.57)	.39 (.46)	-	
TOT TURN	.72 (.69)	-.59 (-.77)	-.24 (-.41)	-

TABLE 3

Individual correlations between internal performance variables and overall performance. Course 2 shown in ().

	SUSTAINED TURN	POSITIONING	EXTENDING	STAGNATING
Course 1	.22 (.17)	.14 (.10)	.32 (.11)	.33 (.20)
Course 2	.26 (.26)	.03 (.02)	.21 (.08)	.50 (.19)

TABLE 4

Proportion of moves in each performance category.
SD in ().

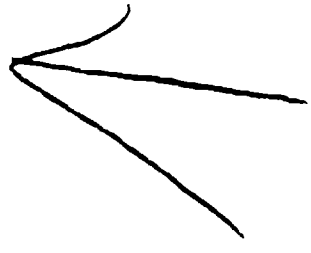
REFERENCES

De Maio, J. "Vehicle Control Decision Processes," presented at 89th annual convention of the American Psychological Association, Los Angeles, CA. 1981

Jagacinski, R. J. "A Qualitative Look at Feedback Control Theory as a Style of Describing Behavior," Human Factors, 19(4), 331-348, 1977.

Schuaneveldt, R. W., Goldsmith, T. E., Dorso, F. T., Maxwell, K., Acosta, H. M., and Tucker, R. G. Structures of Memory For Critical Flight Information, AFHRL-TP-81-46 (in process.)

Smith, E. E., Adams, X. X, and Schorr, Z. Z. "Fact Retrieval and the Paradox of Interference," Cognitive Psychology, 10, 438-464, 1978.



DEVELOPMENT OF PERFORMANCE MEASURES FOR SELECTION OF CREWS FOR FLIGHT TRAINING

EDWARD M. CONNELLY
PERFORMANCE MEASUREMENT ASSOCIATES, INC.

(PAPER NOT PROVIDED)

AD P001622

A FUZZY MODEL OF RATHER HEAVY WORKLOAD

by

Neville Moray and K. Waterton

Department of Industrial Engineering
University of Toronto
Toronto, Canada M5S 1A4

ABSTRACT

An experiment is described in which operators performed a tracking task whose difficulty was varied on two dimensions, bandwidth and order of control. RMS error was measured and subjective judgement of workload obtained on a 10-point scale. Using rms error as a measure of subjective effort, the workload was successfully predicted using fuzzy measures of bandwidth, order of control and effort.

INTRODUCTION

One aspect of mental workload which has been of interest is the relation between objective "task difficulty," performance measures, and subjective judgements (Moray, 1979). The original intention of the experiment described was to use physiological measures as well, and explore the correlations between the different measures of workload. The chosen physiological measure (power at 0.1 Hz in the heart rate spectrum) did not correlate with any of the other measures and will not be further discussed. Instead, after describing the experiment, we will report on the need to take into account the amount of effort put into a task, since that approach enabled some anomalous data to be successfully explained. Intuitively it seems clear that if a person is confronted with an objectively difficult task but does not try hard, then it may be experienced as subjectively less loading than one might expect.

METHOD

A single axis compensatory tracking task was used, in which the operator was required to keep a vertical bar between two reference bars which represented the permissible error. These bars were $\pm$ 0.5σ from the mean of the forcing function, which was Gaussian bandlimited noise, (μ=0, σ=1). Three bandwidths were used, 0.33, 0.67 and 1.0 Hz, provided by a Hewlett Packard Pseudorandom noise generator set at 50 Hz and lowpass filtered with a Kemo programmable filter. Three orders of control were used, K, K/s and K/s^2. The three bandwidths and three orders of control were combined into 9 conditions, and 9 operators performed them in a Latin Square design. They received two days' practice before the data were collected, and the analysis was based on two days worth of data, on each day of which they

performed all nine conditions. There were no significant differences between the two days of data, and both will be analysed. The experiment was run through an EAI analogue computer controlled by a PDP11/45 which provided the control of the filter and collected and scored the data. After each condition the operator was asked to rate the subjective load imposed by the experimental condition on a ten point scale.

RESULTS

The results are shown in summary in Figures 1 and 2. ANOVAS showed that the main effects of treatments were the only significant source of variance. As was expected, increasing the bandwidth and increasing the order of control each made the task subjectively more difficult. Increasing the order of control also increased the rms error, as expected. But there was an anomalous effect in that increasing the bandwidth decreased the rms error. The equipment and the data were carefully scrutinized to see whether this was an artefact, but in the end it appeared to be a real finding. Talking to operators and performing the task led to the conclusion that 0.33 Hz was so low that the operators probably did not try very hard, and that perhaps their response to the increasingly difficult bandwidths was to more than compensate for the difficulty, so that performance overcompensated for the task demands. It was this that suggested trying to estimate subjective effort and take it into account. In the absence of objective measures of effort it was decided to take the opportunity to see whether fuzzy set theory could be used to measure workload.

We define a fuzzy set of situations which are subjectively heavily loading by a linear transform on the subjective judgement scale,

$$\mu_D = 0.1\ (R) \qquad 0 \leq R \leq 10.0$$

where R is the score given by the operators on the ten-point scale. That is, the higher the subjective rating, the stronger the candidacy for membership of a "heavily loading situation."

The imposed load is due to the two task parameters, bandwidth and order of control. We propose that "high bandwidths" and "high orders of control" both impose heavy loads. The membership function of "high orders of control" can be plausibly related to known empirical facts about manual control. Controlled elements with transfer functions K and K/s are easy to control, K/s^2 is very hard, and higher orders impossible. These considerations suggest a membership function as shown in Figure 3. The exact values are at present a matter for arbitrary choice. Without trying to optimize the fit to the data, we set $\mu_s(K/s) = 0.2$, and $\mu_s(K/s^2) = 0.8$.

The membership function of "high bandwidths" we also relate to empirical data. The human operator transfer function is well known, and we propose that a possible reason for high bandwidths being difficult is the increasing phase lag associated with bandwidth. We therefore plot the Bode phase curve inverted and treat this as the membership function for "bandwidths which are high." A phase lag of -180° is certainly high in this

sense, and frequencies below the breakpoint have zero membership of high bandwidths. (Figure 4.)

Since bandwidth and order of control are independent sources of difficulty, their combined effect is the fuzzy union of their memberships, and we have

$$\mu_T = \max \{\mu_s(s); \mu_\omega(\omega)\}$$

where μ_T is the membership of "tasks which are difficult."

The problem of estimating effort is not so direct. We try to deduce it here from the rms error. To reduce error requires effort. But the operators in this experiment were allowed an explicit acceptable error band. Clearly any error within that band is not a member of "errors which are large." On the other hand, since the forcing function was Gaussian, we might reasonably say that any error which is greater than $\pm 2\sigma$ is certainly large. We assume that making a large error means that the operator is not putting a lot of effort into the task. Hence we have the membership function for "errors which are large" shown in Figure 5.

We now define the membership of "putting a lot of effort into the task" as the complement of the membership function of "errors which are large," and we have

$$\mu_E = (1 - \mu_E)$$

We now make the assumption that in order for a task to seem subjectively loading it must both be objectively difficult and also be given a lot of effort. The appropriate relation is the fuzzy intersection of task difficulty and effort, and we have finally,

$$\mu_D = \min \{ \max(\mu_s; \mu_\omega); (1-\mu_E)\}$$

which allows us to calculate the predicted subjective task difficulty. In Figure 6 we show the relation between the task difficulty predicted from this fuzzy set model, together with the observed subjective ratings rescaled to the range 0 to 1. The squares are from the first and the circles from the second days' data. Eight out of the nine individual operators showed good linear relationships. It should be emphasized that no attempt has been made to optimize the membership functions by direct scaling. Rather it was thought preferable to try to establish the membership functions independently of subjective judgement, so as to have a more analytic model. We believe that the results suggest that fuzzy measurement has a promising application in the field of mental workload measurement. Further experiments are currently underway at Toronto, including measures of the effect of combining the subjective difficulty of two tasks. Interesting preliminary results are available. For example, combining two tasks of moderate difficulty produces a difficult or very difficult task: comining two difficult tasks produces a very difficult task, and the impression is that neither averaging nor addition takes place for subjective difficulty, but rather something corresponding to a shift operator which moves the membership function towards the higher end due to combination.

REFERENCES

Moray, N. 1979. Mental Workload: Theory and Measurement. Plenum, New York.

Zadeh, L. 1965. Fuzzy sets. Information and Control, 8, 338-353.

ACKNOWLEDGEMENTS

We wish to thank the Department of Psychology, Stirling University, Scotland, for facilities to perform the experiments, and Professor I.B. Turksen of the Department of Industrial Engineering, University of Toronto, for helpful discussions on fuzzy sets.

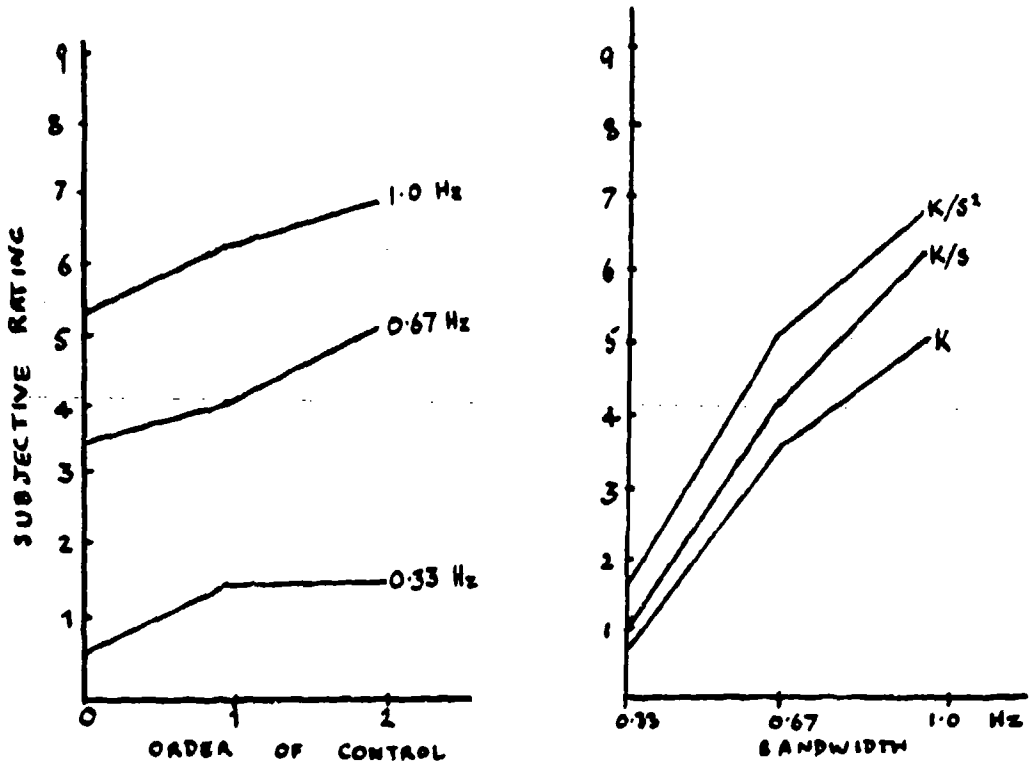


FIGURE 1

SUBJECTIVE DIFFICULTY

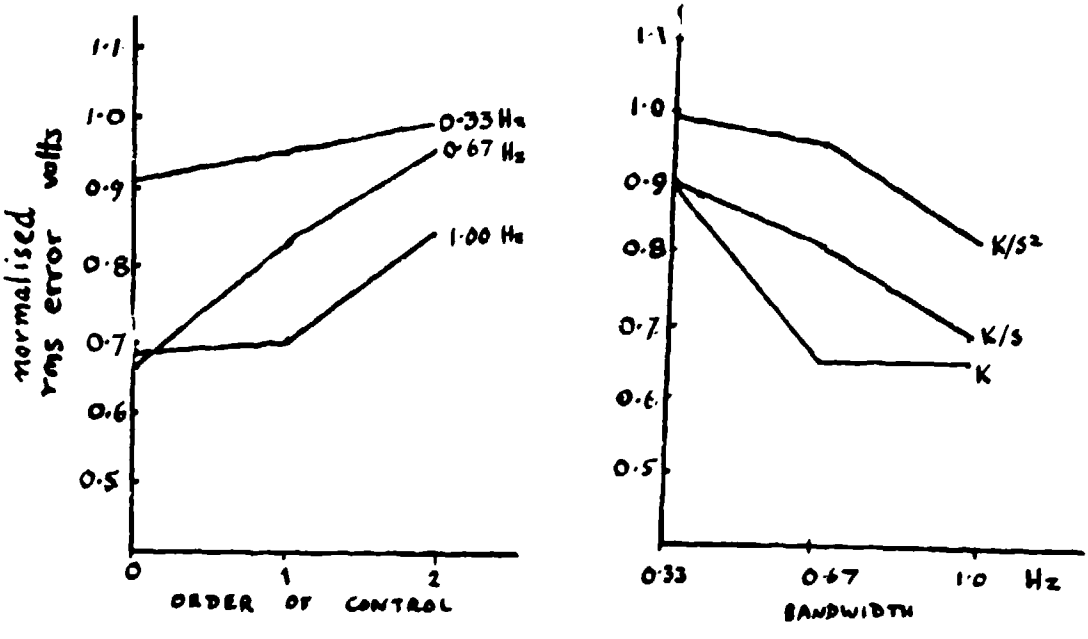


FIGURE 2

PERFORMANCE MEASURE: rms ERROR

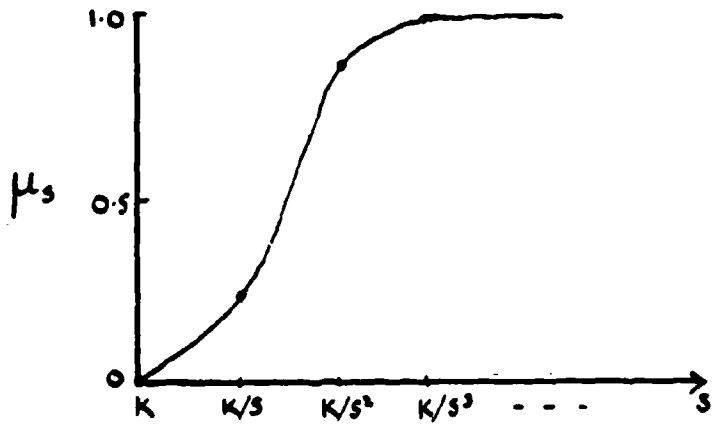


FIGURE 3

Membership of "high order of control"

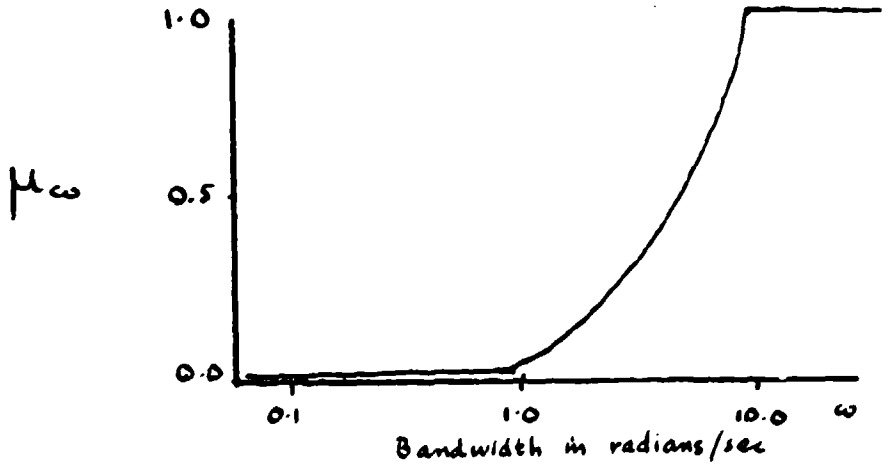


FIGURE 4

Membership of "bandwidths which are high"

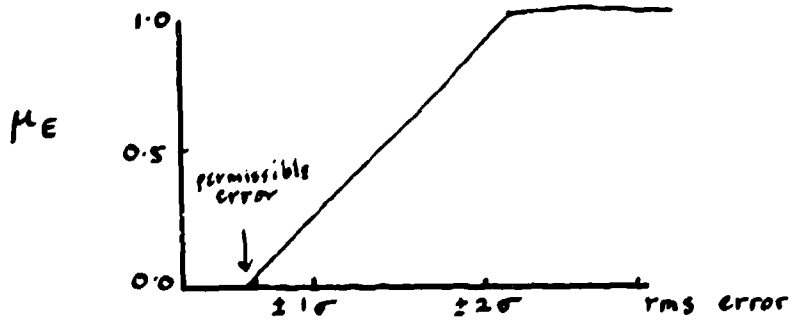


FIGURE 5

Membership of "errors which are large"

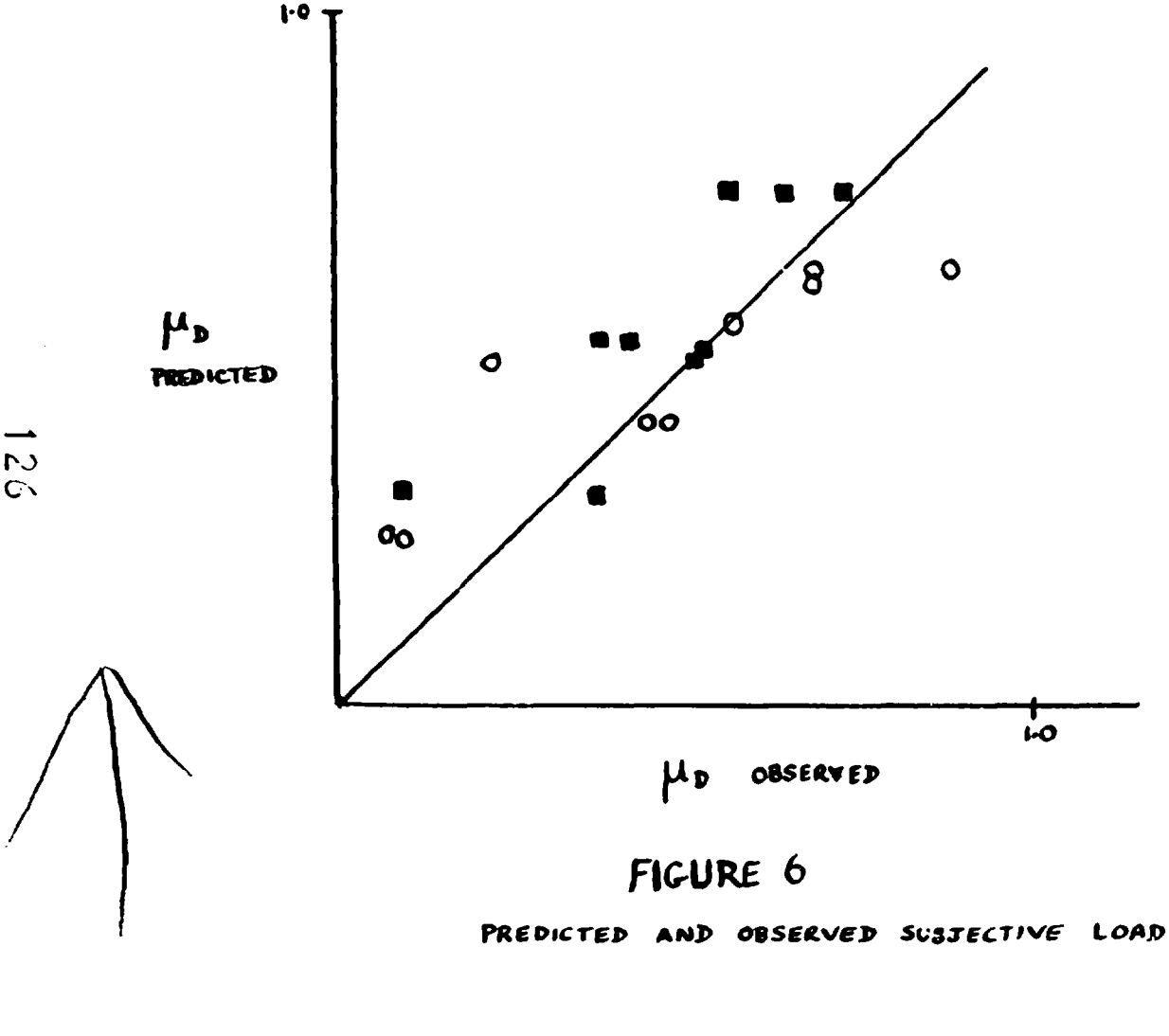


FIGURE 6

PREDICTED AND OBSERVED SUBJECTIVE LOAD

AD P001623

RATING CONSISTENCY AND COMPONENT SALIENCE IN SUBJECTIVE WORKLOAD ESTIMATION

Jan R. Hauser* Mary E. Childress* Sandra G. Hart
San Jose State University NASA/Ames Research Center

ABSTRACT

Twelve general aviation pilots participated in a two-day experiment performing four tasks intended to load on different cognitive, perceptual, and motor dimensions. The tasks were varied in apparent difficulty level so that each pilot performed a total of sixteen tasks counter-balanced for task and level. Subjective ratings of factors contributing to workload were made immediately following each level of each task using a 15 bipolar adjective scale. Results indicated that the subjective perception of workload was not related to actual performance measures; however, the subjective ratings were generally consistent with the demands made by the levels of each task. Although only two of the rating scale items, own Performance and Task Difficulty, demonstrated significant within-task differences for all four tasks, the majority of rating scales showed within-task differences for those tasks that imposed higher cognitive demands. Strong relationships were found between Overall Workload, Stress Level, and Task Difficulty ratings on all tasks.

INTRODUCTION

In the past year there has been a great deal of interest in the subject of defining, measuring and reducing pilot workload as a result of the President's task force on crew complement. In their attempt to determine how best to certify an aircraft with respect to minimum crew size, it became clear that a critical factor was workload.

Although there is some broad general agreement on definitions of workload, there is little agreement on what the specific factors are that contribute to workload either singly or in combination, nor are there any standard sets of procedures or techniques that can be used to measure workload. Recent research (Hart, Childress & Hauser, 1982) suggests that a number of independent dimensions of workload can be identified and, further, that differing combinations of these factors are relevant to different individuals in experiencing workload.

---

* Supported by NASA Grant NCC 2-034 to the San Jose State University Foundation.

The focus of the present study was to determine what aspects of different task situations cause an individual's experience of workload and related factors to vary. Subjective rating scales provide one method of evaluating workload (Bird, 1981; Hicks & Wierwille, 1979; Jex & Clement, 1979; Williges & Wierwille, 1979), particularly when combined and compared with objective performance measures. The present study built on previous research (Bird, 1981) which examined the effect of different factors on the the individual's experience of workload using a single-axis compensatory tracking task with six levels of difficulty. Subjects rated each level of the tracking task using a 15 bipolar-adjective scale that addressed a variety of factors assumed to be related to the subjective experience of workload. Factors related to the difficulty levels of the task were: skill required, task difficulty, controllability, and task demands. The increase in subject effort (stick input) was shown to be related to evaluations of skill required, task complexity, task difficulty, and task demands.

In the present experiment four different tasks (including several levels of the tracking task used by Bird), each with four levels of difficulty, were presented to the subjects, who were required to rate each level of each task immediately after its performance using a modified version of the earlier scale. Although the tasks chosen for use in this experiment (compensatory tracking, the Sternberg task, auditory monitoring, and time estimation) are commonly employed as secondary measures of workload, in this instance they were used only as primary tasks in order to obtain ratings that would stand alone and thus be useful for future research comparisons where these tasks will function as secondary tasks.

METHOD

Subjects

Twelve general aviation pilots, ten males and two females, 21 to 43 years old, served as paid volunteers. All subjects were right handed.

Apparatus

All experimental sessions were conducted in a small, sound-attenuated experimental chamber. The subject was seated in a comfortable chair approximately 100 cm away from the Cathode Ray Tube (CRT) screen. The control stick used for the tracking task was located on the right arm of the chair. The button press used for the time estimation task was centrally mounted on the top of the control stick. Press buttons used for both the Sternberg memory task and the auditory monitoring task were mounted on the left arm of the chair, with all functions clearly labeled. The analog control for rating evaluations was

also mounted on the left arm of the chair. Data were recorded and task presentations were programmed on a Digital Equipment Corporation PDP-12 computer.

Subjective Rating Scale. The multi-dimensional rating scale consisted of 15 descriptive items that addressed specific factors that may contribute to an individual's perception of workload; cognitive, perceptual, and psychomotor dimensions as well as feelings of stress, fatigue, motivation, time pressure, etc. The specific scales were dichotomized to two extreme opposite levels: Overall Workload (High/Low), Performance (Great/Lousy), Training (Too Much/Too Little), Attention Level (Constant/None), Motivation (Excited/Indifferent), Stress Level (Extremely Tense/Relaxed), Physical State (Wide Awake/Exhausted), Time Pressure (Very Rushed/None), Task Difficulty (Very Hard/Very Easy), Task Complexity (Very Complex/Very Simple), Activity Level (Very Busy/Idle), Physical Effort (High/Low), Mental Effort (High/Low), Sensory Effort (High/Low).

The scales were displayed singly in the above sequence on a CRT screen. The specific scale descriptor was positioned at the bottom of the screen, with the scale itself represented by a vertical line (11.0 cm) marked with "ticks" at both ends and in the middle. The dichotomous definitions were placed adjacent to the identifying ticks at the top and bottom of the scale. Subjects assigned a subjective rating by positioning a cursor along the vertical line. They were required to "zero" the cursor by positioning it at the base of the vertical line before making each rating. Subjects ratings were computed as a percentage of the distance from the bottom stationary tick mark (0) to the top stationary tick mark (100).

Tracking Task. Four levels of difficulty for a single-axis compensatory tracking task were presented on a monitor. The display consisted of a vertical line (3.7 cm) which moved quasi-randomly in a lateral direction. The subjects' task was to keep this cursor centered between two stationary vertical lines (1.5 cm) by movement of a control stick.

The output of a random number generator provided a rectangular distribution of frequencies simulating white noise with a bandpass of 62 rad/sec. The values were passed through a second-order filter having a natural frequency of 1.0 or 2.0 rad/sec. The resulting forcing functions were presented with standard deviations of either 32 pixils (12% of the screen width) or 64 pixils (24% of the screen width) to reproduce four of the six experimental conditions used by Bird (1981). The resulting task difficulty levels thus were varied with respect to frequency and amplitude so as to determine the independent and combined effects of these two factors on the experience of workload.

Sternberg Memory Task. The second set of four tasks was designed to investigate the effect of memory load and presentation rate on the experience of workload. The widely used Sternberg task (Sternberg, 1966) was programmed so as to create four levels of difficulty. This task involves memorization of one or more alphabet characters, the "positive set", with subsequent presentation of a series of characters requiring true or false responses to those which were or were not members of the positive set.

Two levels of difficulty were manipulated by the use of either a single letter or five letters as members of the positive set. There were two levels of rate of presentation with the inter-stimulus interval set at either 1.5 or 3.0 seconds. Letters were displayed for 1.0 seconds in all experimental conditions. A randomly generated set of alphabetic characters was produced for each of the four levels of this task, with the stipulation that the positive set character(s) appeared 50 per cent of the time. The letters (0.75 cm high, 0.50 cm wide) were presented on a CRT screen. Subjects responded by button press (yes or no) to the stimuli depending on whether or not the displayed character was a member of the positive set.

Auditory Monitoring Task. An auditory monitoring task was included in the test battery to further investigate the impact of memory load and presentation rate with a non-verbal task. The complex counting task developed by Kennedy (1968; 1971) was selected because it has been shown that this task is associated with very stable within- and between-subject performance for different levels of difficulty.

The present version of the task was presented with two levels of memory load, similar to those used by Kennedy. Two levels of rate of presentation were also employed, departing from Kennedy's use of one presentation rate, to investigate the effect of time pressure variation at low levels of task demand. The memory load required either a response to all occurrences of one designated tone (no load) or to every 2nd or 4th occurrence of such a tone (memory load). Three tones (low/448 Hz, medium/1152 Hz, high/1988 Hz) were sequenced so that the high tone occurred 8 times per minute, the middle tone 12 times per minute, and the low tone 16 times per minute for the fast version and 4, 6, and 8 times per minute for the slow version. All tones in all conditions were activated for 0.4 seconds. The tones were produced by a Wavetek Voltage Control Generator with voltage fed by the PDP-12 D-to-A converter, the output gated on and off by a logic signal from the PDP-12 computer and presented to the subjects through a four-inch intercom speaker. Subject responses were made by labeled button press (hi, med, lo).

Time Estimation Task. Time estimation tasks were included in the test battery because they represent tasks that are stimulus-deficient, invoking a simple infrequent response but

necessitating continued attention by the subject to produce consistent and accurate responses (Hart, McPherson, & Loomis, 1978).

Two levels of method of time estimation were used: production or verbal estimation. Two estimation techniques were varied under each method: either counting aloud, or with no attempt at any type of rhythmic monitoring of the passage of time. Although estimation and production accuracy have been shown to be similar with both techniques, subjective reports indicate that the two techniques differ markedly with respect to stress, attention required, uncertainty, and perceived accuracy. Again, these four levels of task demand were selected to determine the effect of relatively subtle, but none-the-less important, primarily cognitive parameters on the experience of workload.

The intervals used ranged in duration from 5-14 seconds and were presented in a random order with each interval appearing twice in every session. The beginning of each trial was signaled by a message on a CRT screen, e.g., "Production _ seconds", or "Begin interval" for the production or verbal estimation responses respectively. The subject responded with a button press which initiated either the production of the specified interval or the presentation of the interval to be estimated. A second button press signaled completion of a produced interval or indicated that the subject had noticed that the interval had ended and thus could make the estimate. This procedure was designed to hold constant both the number and type of overt responses and visual information.

Procedure

Subjects were given a briefing at the start of the first day of experimental sessions. Descriptions of the tasks were given, as were instructions for the performance of the tasks and the subsequent ratings. The importance of maintaining an equally high standard of performance across all tasks was stressed. All subjects were given a copy of the definitions provided for each of the rating scale items and were encouraged to refer to it and to exercise careful consideration during the rating procedure.

The order of presentation of tasks and levels of tasks was counterbalanced across the 12 subjects. The entire experiment was run over a one and a half day period. Subjects performed three or four of the laboratory tasks per 45-minute session, interspersed with sessions that involved flying a motion-based general aviation trainer (GAT) simulator. The GAT simulator study will be reported elsewhere. Breaks were given between sessions.

For all tasks, subjects practiced for one minute immediately prior to a four-minute experimental session. The time esti-

mation task length varied as a result of subject responses, but typically lasted 4-5 minutes. After completion of each experimental task, subjects rated their subjective experience on each of the 15 bipolar items.

At the end of the experiment, subjects rated their aggregate experience across all of the tasks on the same 15 bipolar rating scales so that a comparison could be made between the aggregate rating and the mean ratings across all levels and tasks.

RESULTS AND DISCUSSION

The primary focus of the data analysis was on the ratings obtained within and between tasks. In addition, appropriate measures of effort and performance were analyzed for each task and compared to the subjective ratings. All data analyses on the performance measures of the four experimental tasks and the related rating scales were computed using one- and two-way Analyses of Variance (ANOVAS) with repeated measures.

Between-Task Ratings

Fifteen one-way ANOVAS were performed using the means of the subjects' ratings across the four levels within each task in order to determine if the scales discriminated between the four types of tasks. There were significant differences for all scales except Training and Physical State (see Figure 1), indicating that the ratings did reflect the different demands imposed by different task type. A Newman-Keuls test for ordered means showed that the significant effects for the ratings of Overall Workload, Attention Level, Stress Level, and Activity Level were due to higher ratings on the tracking and Sternberg tasks than for the auditory monitoring and time estimation tasks.

The significantly better evaluations of own Performance on the Sternberg and auditory monitoring tasks probably reflected the more quantitative nature of these two tasks. Subjects were in effect able to monitor their own success or failure on a simple right or wrong basis, whereas no such evaluation could be made for the tracking or time estimation tasks as the perception of success was much more nebulous for these tasks. It should be remembered that no feedback was given to subjects about their performance.

Time estimation was perceived as the least interesting task the four, but there were no differences on the rated Motivation between the tracking, Sternberg, and auditory monitoring tasks. As the same result was found on the Sensory Effort scale, it appears intuitively reasonable that the low perceptual demands of time estimation resulted in a lack of interest.

The tracking tasks were significantly more difficult than

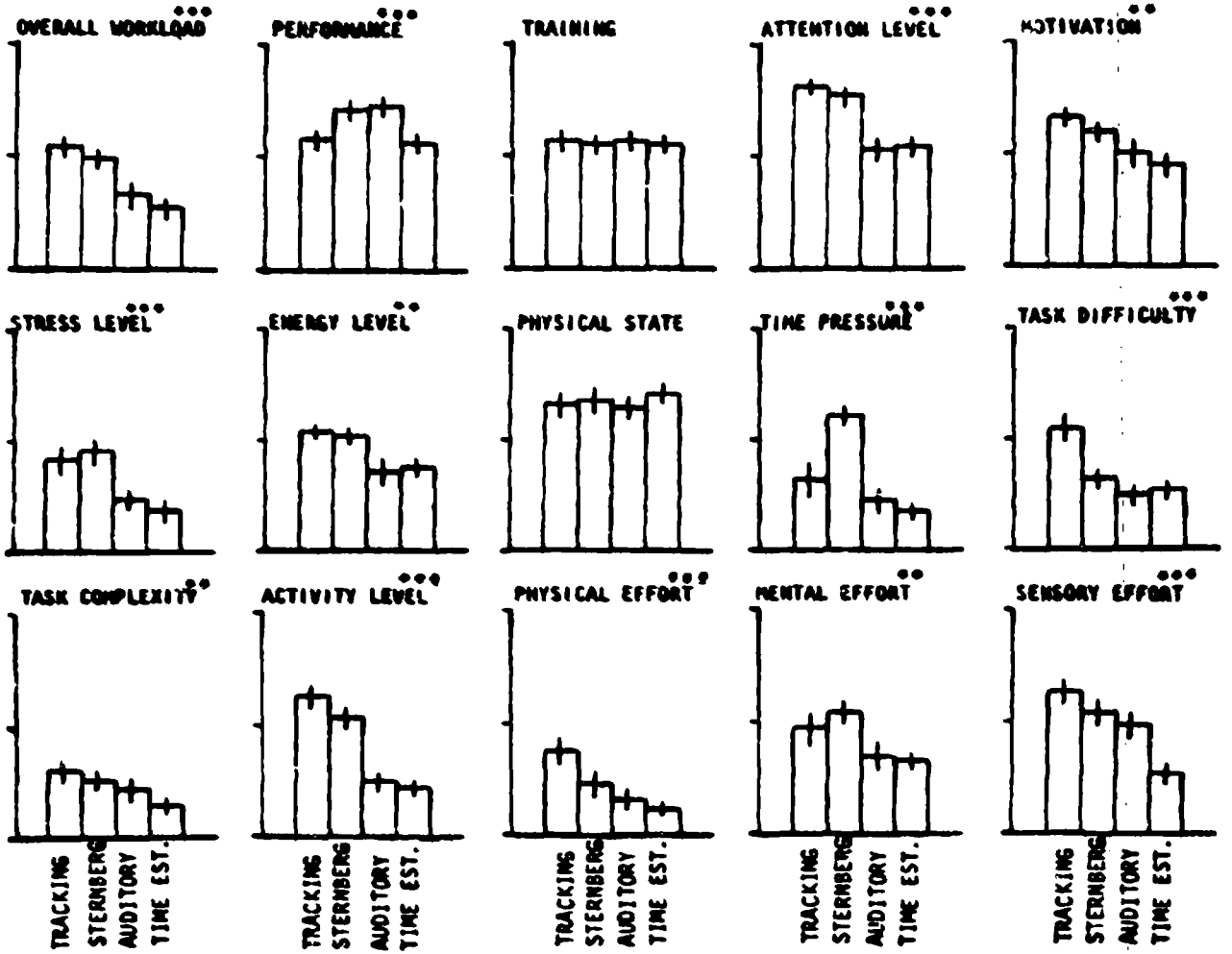


Figure 1. Rating scale means and standard errors calculated over the four levels of each task with significant between task differences identified ** $p<.01$, *** $p<.001$.

any of the other three tasks (which were rated as equivalently difficult), whereas Physical Effort ratings, although low, were significantly greater for the tracking and Sternberg tasks, reflecting the high response demands, than for either the time estimation or auditory monitoring tasks.

Differences between tasks on the Energy Level, Task Complexity, and Mental Effort scales were small in general, and the Training and Physical State scale ratings were nearly identical across tasks. All subjects rated training as at least adequate (at the 50 per cent level) or slightly more than adequate. The subjects' Physical State or level of fatigue, was apparently not differentially affected by task type. Overall Workload was rated at about the mid-range for the tracking and Sternberg tasks, and 15 to 20 per cent lower than that for the auditory monitoring and time estimation tasks. The effort ratings were rarely above the mid-range with the exception of the Sensory Effort ratings on the tracking and Sternberg tasks, indicating that in general subjects felt that the demands of the tasks were not very high.

Within-Task Analyses

Tracking Task: Objective Measures. Two measures of performance were examined: (1) RMS (root mean squares) tracking error over RMS input (i.e., how much the error was reduced by the subjects' action) and (2) RMS of the stick samples (i.e., how much stick movement was made by the subject). An increase in the bandwidth caused an increase in tracking error RMS, which might be expected considering the relative amount of error the subject was required to reduce. Increasing the standard deviation (from 32 to 64) produced an increase in the stick output which is also intuitively reasonable as the travel distance of the cursor was doubled.

The more difficult levels of standard deviation ($F(1,11)=19.12$, $p<.001$) and bandwidth ($F(1,11)=273.53$, $p<.0001$) conditions resulted in significant decrements of the degree to which subjects succeeded in reducing signal errors, however, the standard deviation accounted for only 2% of the variance whereas the bandwidth accounted for 79% of the variance. For the bandwidth of 1.0 rad/sec, the error was reduced by about 50%, but the bandwidth of 2.0 rad/sec showed an error reduction of only 15%. The error reduction was greater for the standard deviation of 32 than for 64 only when the bandwidth was 1.0 rad/sec.

Stick activity was also significantly greater for the more difficult levels of both variables; standard deviation $F(1,11)=117.01$, $p<.0001$, and bandwidth $F(1,11)=33.60$, $p<.0001$. In this case the standard deviation accounted for 40% of the variance and bandwidth for 16% of the variance. Doubling the standard deviation did in fact double the stick activity,

whereas increasing the bandwidth from 1.0 to 2.0 rad/sec increased stick activity by about 50%.

Tracking Task: Rating Scales. It was anticipated that the increasing difficulty levels of the task would be reflected in subjects' rating of the workload involved. In fact, only five of the rating scale items did produce significant differences in the expected direction, as shown in Figure 2. The percentages of variance accounted for were generally small, so although differences were perceived they were apparently in a limited range. The increase in bandwidth affected Performance and Activity Level ratings; subjects evaluated their performance as less successful when the bandwidth was 2.0 rad/sec, and higher Activity Level ratings reflected increased stick activity associated with an increase in the bandwidth from 1.0 to 2.0 rad/sec. The increased standard deviation produced lower ratings on the Motivation scale and higher ratings on the Task Difficulty scale. The interaction of bandwidth and standard deviation affected the Stress Level scale with the largest difference again attributable to the standard deviation variable.

In general, the Attention Level required for the tracking task was thought to be very high, and the Sensory Effort and Activity Level were both perceived as being moderately high. Task Complexity, Physical Effort, and Time Pressure were all rated as relatively low.

Sternberg Task: Objective Measures. Two measures of performance were examined; (1) reaction time for correct responses, and (2) per cent of correct responses. There was a significant increase in reaction time ($F(1,11)=185.68$, $p<.0001$) as the memory set size was increased from 1 to 5 characters as expected, and reaction times were slower, but not significantly so, for the slower rate of presentation. There was no significant variation in the per cent of correct responses for either rate of presentation or memory set size, although there was a decrease in per cent correct for the faster rate of presentation when the memory set size was 1.

Sternberg Task: Rating Scales. The rate of presentation did produce a significant effect on 11 of the 15 bipolar rating scales as shown in Figure 3. In every case the faster rate produced higher ratings, except for the Performance scale, where performance was rated as poorer for the faster rate. The ten scales that showed higher ratings for the faster rate of presentation were: Overall Workload, Stress Level, Energy Level, Time Pressure, Task Difficulty, Task Complexity, Activity Level, Physical Effort, Mental Effort, and Sensory Effort. Ratings on these scales did accurately reflect the greater demands made by the faster rate as subjects did have to respond twice as often for the faster rate of presentation with 40 button presses per minute compared to 20 for the slower rate. Rather than perceiving the increased rate as fatiguing, subjects apparently felt

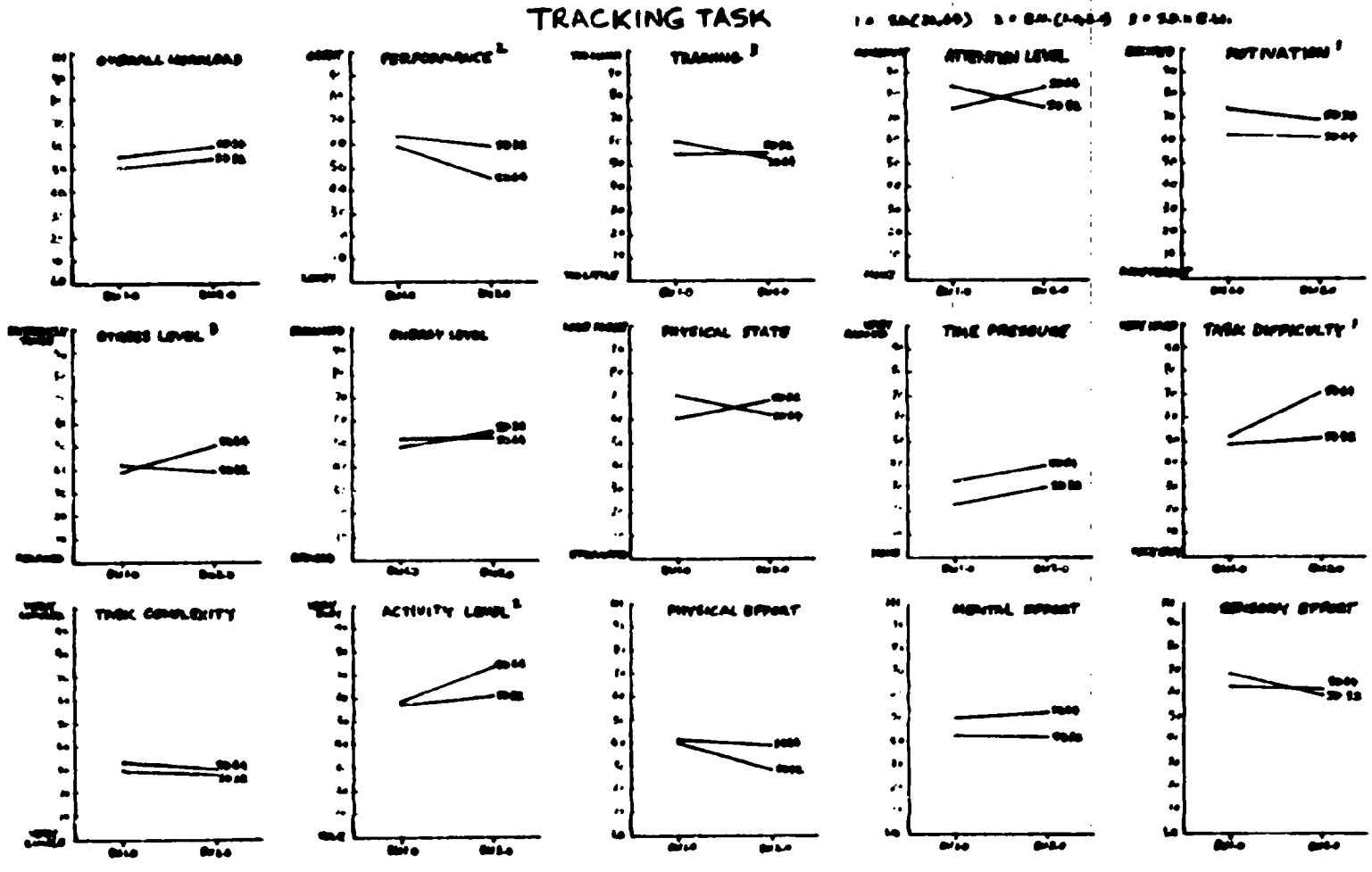


Figure 2. Mean ratings and significant effects at or beyond the .05 level for standard deviation (1), bandwidth (2), and the interaction of standard deviation and bandwidth (3).

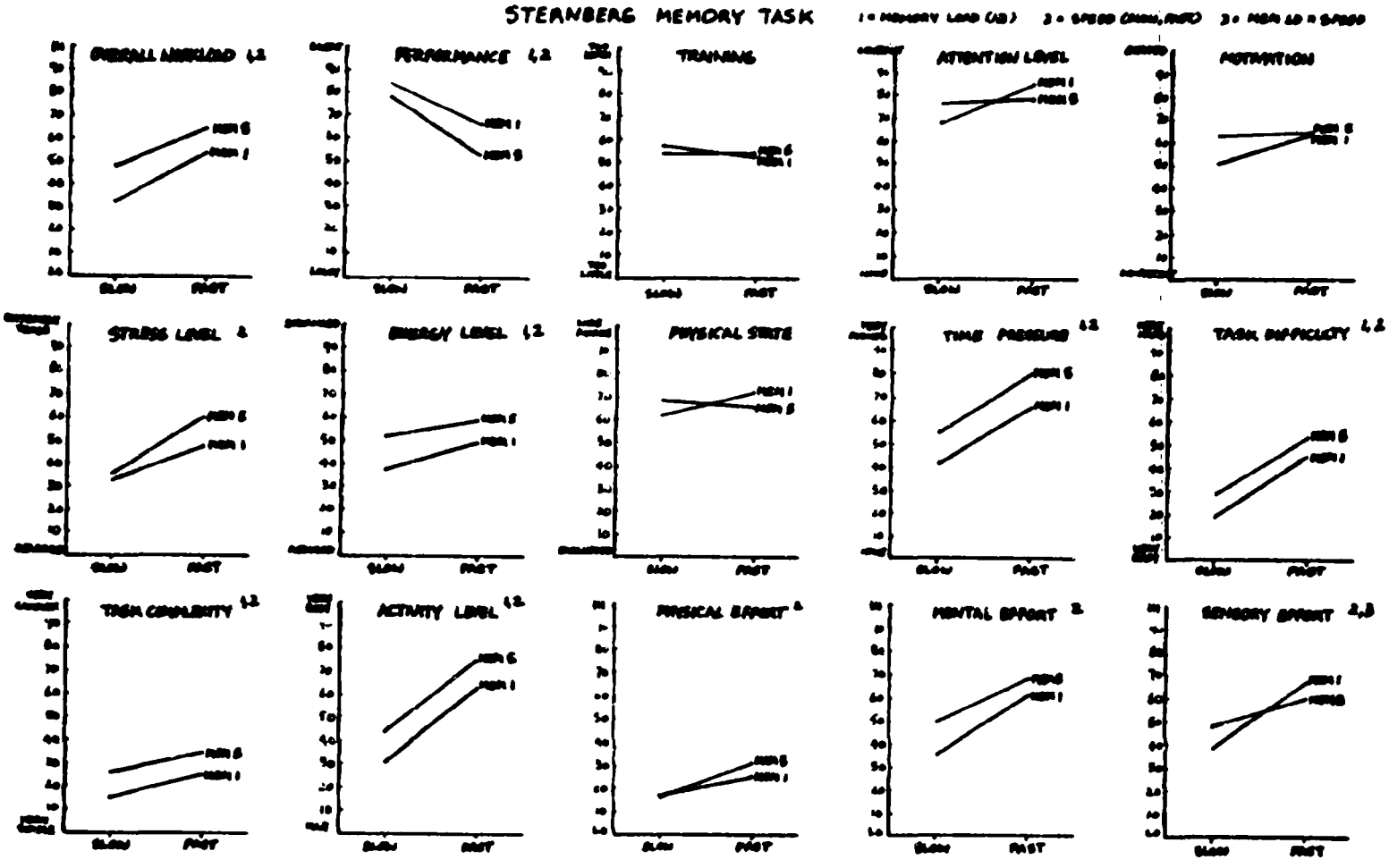


Figure 3. Mean ratings and significant effects at or beyond the .05 level for memory load (1), rate of presentation (2), and the interaction of memory load and rate of presentation (3).

that their energy level was enhanced.

The number of characters in the positive memory set resulted in significant differences on seven of the 15 bipolar scales: Overall Workload, Performance, Energy Level, Time Pressure, Task Difficulty, Task Complexity, and Activity Level, and again the greater demand was reflected in higher ratings except for Performance, where the rating was lower for the memory set of 5. As was shown in the objective measure analysis, successful performance was significantly lower for the memory set of 5 as compared to that of 1. As was the case for increased rate of presentation, increased memory set size also enhanced subjects' energy levels, although the difference was not as great.

It is interesting to note here that the increased rate of presentation rather than the increased size of the memory set affected the ratings of stress and effort. In general subjects perceived greater differences between the levels of difficulty for this task than for the tracking task. The percentages of variance accounted for on the rating scale analyses were larger than those discovered for the tracking task. The ratings on the Attention Level scale were generally high, as they were for the tracking task. Similarly the ratings on Task Complexity and Physical Effort were low overall.

Auditory Monitoring Task: Objective Measures. The performance measures for this task were the number of correct responses and number of wrong or missed responses. The task demands were relatively low for all levels of the task; the number of responses required per minute were either 2 or 8 depending on the condition, whereas the number of responses required for the Sternberg memory task were either 20 or 40 per minute by comparison. Thus, performance was virtually error-free and no significant differences were found as a function of memory load or presentation rate.

Auditory Monitoring Task: Rating Scales. The rate of presentation showed a significant effect on three of the scales: Overall Workload, Attention Level, and Energy Level. The faster rate was reflected in higher ratings on these scales. Although there was no significant effect for the presentation rate on the Performance scale, subjects perceived their performance as better with the faster rate of presentation (although in fact there was no difference in performance measures). It is possible that the level of boredom imposed by the slower rate affected the subjective assessment of performance, i.e., subjects felt they did better when they had more to do.

The effect of memory load was reflected by significant differences on eight of the scales: Overall Workload, Performance, Training, Motivation, Task Difficulty, Task Complexity, Mental Effort, and Sensory Effort. In all cases the higher load condition generally resulted in higher ratings with the excep-

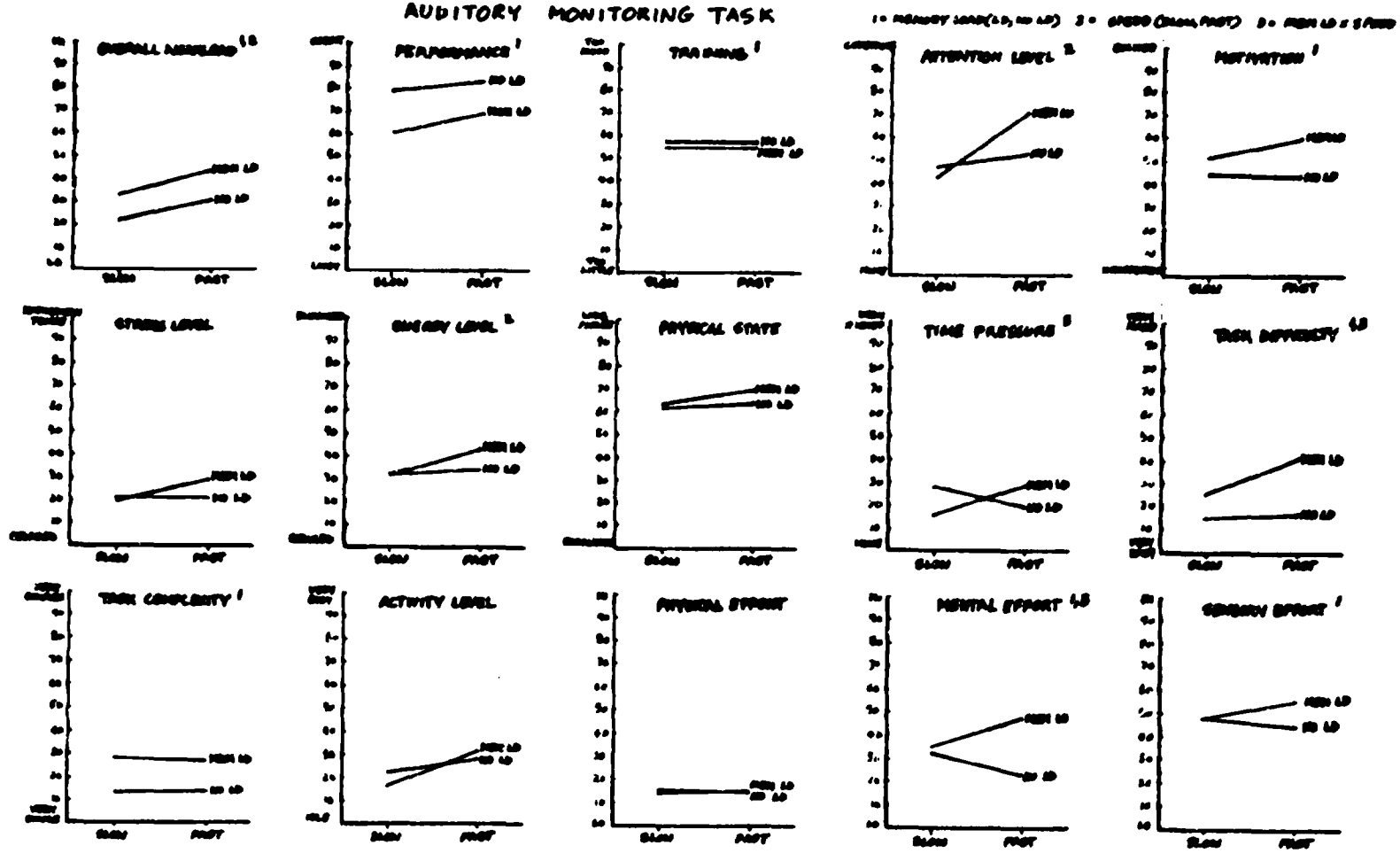


Figure 4. Mean ratings and significant effects at or beyond the .05 level for memory load (1), rate of presentation (2), and the interaction of memory load and rate of presentation.(3).

tion of Performance and Training, where the higher load resulted in a lower rating. The actual differences for the ratings on the Training scale were very slight, but may have reflected a desire for a longer practice period under the memory load condition. The one-minute practice session allowed only two responses for each of the memory load conditions and may not have been sufficient. The differences due to memory load for both Mental and Sensory Effort were perceived only under the faster rate of presentation.

Significant interactions between memory load and rate of presentation were found for Time Pressure, Task Difficulty, and Mental Effort. For both Task Difficulty and Mental Effort, subjects perceived no differences due to memory load with the slow rate of presentation. For the faster rate, the higher memory load was felt to be more difficult, and require more thought, although in both cases the ratings were still below the 50 per cent level.

Ratings were generally low on the majority of scales, particularly on those of Stress Level, Time Pressure, Task Difficulty, and Physical Effort (Figure 4). These ratings accurately reflected the low demands made by this task and the variance accounted for was again very small in all instances.

Time Estimation: Objective Measures. The ratio of estimated or produced duration to the standard duration was used as the performance measure for time estimation. The means and standard deviations were computed across all 20 estimates or productions made by each subject under each experimental condition. The mean ratios of estimated to standard durations were significantly different ($F(1,11)=9.03$ $p<.01$) from the mean ratios of produced to standard durations: the latter were typically greater than 1.0 and the former typically less than 1.0. Since a long production and a short verbal estimate both represent underestimation of elapsed time (Bindra & Waksberg, 1956), the difference is due to the same underlying process. The technique employed did not significantly affect average estimation or production accuracy.

The variability of produced and verbally estimated durations was similar; however, in this case the effect of technique resulted in a highly significant difference ($F(1,11)=57.53$ $p<.0001$). Estimates or productions made with the counting technique were much less variable than estimates made without counting. This phenomenon occurred both within and between subjects and is in agreement with earlier results (Bird & Hart, 1980; Hart, McPherson, & Loomis, 1978).

Time Estimation: Rating Scales. In general the time estimation task produced very low ratings on Overall Workload, Stress Level, Time Pressure, Task Complexity, Activity Level, Physical Effort, and Sensory Effort. There was no signif.cant

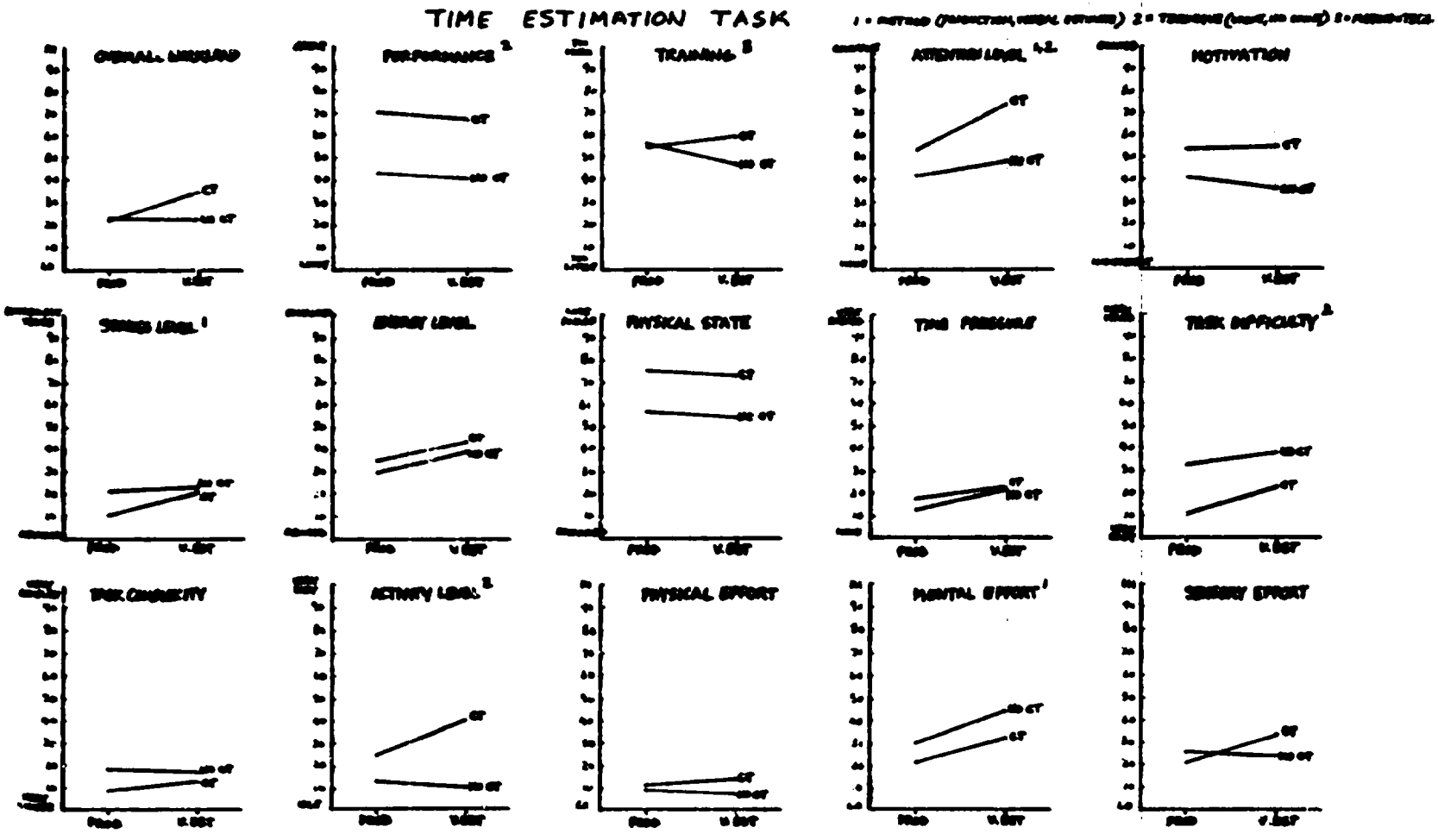


Figure 5. Mean ratings and significant effects at or beyond the .05 level for method (1), technique (2), and the interaction of method and technique (3).

effect for Overall Workload on the time estimation task, but for the verbal estimation method, where subjects had to count aloud and then make the estimate over the intercom, there was a much higher rating than for any other level, reflecting the fact that there was more to do in this condition than in any other. There were significant effects for method of time estimation for three of the rating scales: Attention Level, Stress Level, and Mental Effort (Figure 5). In every case the ratings for verbal estimation were higher than for production, probably reflecting a feeling of reduced control. The duration of the time interval is controlled by the experimenter for verbal estimation, whereas the subject has control of the elapsed duration for production.

Technique significantly affected four of the rating scales: Performance, Attention Level, Task Difficulty, and Activity Level. Subjects perceived their performance as being much better with counting. In fact, their estimates were equally inaccurate regardless of the technique, but they were more consistent when able to count aloud. Counting was perceived as requiring more attention and more activity, as it in fact did, but no counting was rated as being more difficult than counting, because counting provides a concrete and rhythmic basis upon which to estimate the passage of time. Although the Motivation scale did not show a significant effect for the technique variable, subjects were generally more highly motivated when they were able to count.

Relationships Between Subjective and Objective Measures

Examination of relationships among objective measures of performance and subjective ratings and the inter-relationships among subjective ratings was performed using Pearson product moment correlation coefficients. Significance levels were determined using a one-tailed test with 10 degrees of freedom.

The mean rating scale intercorrelations across the four levels within each task were calculated to determine the relationships within tasks (Table 1). There were no significant relationships between any of the subjective ratings and the appropriate objective measures on any of the four tasks. As noted by Moray (1982), this finding is not at all unusual. It may be that subjects make allowances for the demands of the task when making ratings. If the demands are such that they are perceived as being beyond reasonable requirements then the subject may allow himself or herself more errors but still consider his or her performance or effort as adequate. Furthermore, the effort involved in meeting task demands may be increased without any significant change in the perception of workload, providing the effort required remains within the capabilities of the operator. Such variables as time pressure, motivation, fatigue, physical state, etc., all may impact the operator's ability to meet the task demands.

Table 1
Mean rating scale intercorrelations across the four levels within each of the four tasks; $p<.05$*, $p<.02$**, $p<.01$***.

	OVERALL WORKLOAD	PERFORMANCE	TRAINING	ATTENTION LEVEL	MOTIVATION	STRESS LEVEL	ENERGY LEVEL	PHYSICAL STATE	TIME PRESSURE	TASK DIFFICULTY	TASK COMPLEXITY	ACTIVITY LEVEL	PHYSICAL EFFORT	MENTAL EFFORT	SENSORY EFFORT
OVERALL WORKLOAD	1.00	-.22	.30	.45	.55	.73***	.62	-.19	.60	.74***	.54	.63*	.55	.67**	.64*
PERFORMANCE		1.00	-.13	.07	.03	-.30	.02	.07	-.21	-.30	-.44	-.12	.06	-.16	-.09
TRAINING			1.00	.03	.03	.35	.02	.15	-.00	.10	.26	.04	.11	.32	.18
ATTENTION LEVEL				1.00	.45	.43	.44	.09	-.01	.36	.05	.47	.35	.35	.44
MOTIVATION					1.00	.38	.45	.31	.25	.23	.23	.38	.16	.32	.45
STRESS LEVEL						1.00	.46	-.33	.38	.74***	.42	.64*	.61*	.68**	.53
ENERGY LEVEL							1.00	-.01	.23	.47	.36	.40	.21	.34	.41
PHYSICAL STATE								1.00	-.11	-.25	.04	-.14	-.05	-.09	-.21
TIME PRESSURE									1.00	.44	.38	.43	.10	.47	.28
TASK DIFFICULTY										1.00	.21	.78***	.52	.62*	.64*
TASK COMPLEXITY											1.00	.21	.07	.31	.10
ACTIVITY LEVEL												1.00	.63*	.50	.66**
PHYSICAL EFFORT													1.00	.32	.65
MENTAL EFFORT														1.00	.61
SENSORY EFFORT															1.00

COMPENSATORY TRACKING TASK

OVERALL WORKLOAD	1.00	-.14	.17	.32	.56	.65*	.30	.21	.44	.62*	.45	.64*	.51	.47	.54
PERFORMANCE		1.00	-.16	.36	.09	-.27	-.15	.12	-.20	-.22	-.24	-.09	-.16	-.01	.05
TRAINING			1.00	-.08	.21	.17	-.11	.15	.10	.01	.43	.21	.42	.14	.24
ATTENTION LEVEL				1.00	.34	.31	.41	.31	.28	.06	-.07	.33	-.13	.37	.15
MOTIVATION					1.00	.23	.45	.43	.38	.09	.02	.33	.17	.28	.30
STRESS LEVEL						1.00	.39	-.20	.63*	.73***	.50	.37	.58	.58*	.61
ENERGY LEVEL							1.00	.10	.42	.35	-.10	.26	-.04	.19	.05
PHYSICAL STATE								1.00	-.30	-.20	-.11	.13	-.16	.10	.12
TIME PRESSURE									1.00	.45	.09	.33	.29	.46	.42
TASK DIFFICULTY										1.00	.57	.39	.47	.46	.49
TASK COMPLEXITY											1.00	.29	.60*	.25	.24
ACTIVITY LEVEL												1.00	.38	.31	.47
PHYSICAL EFFORT													1.00	.07	.50
MENTAL EFFORT														1.00	.51
SENSORY EFFORT															1.00

STERNBERG MEMORY TASK

	OVERALL WORKLOAD	PERFORMANCE	TRAINING	ATTENTION LEVEL	MOTIVATION	STRESS LEVEL	ENERGY LEVEL	PHYSICAL STATE	TIME PRESSURE	TASK DIFFICULTY	TASK COMPLEXITY	ACTIVITY LEVEL	PHYSICAL EFFORT	MENTAL EFFORT	SENSORY EFFORT
OVERALL WORKLOAD	1.00	-.16	.29	.50	.61	.78**	.76***	.11	.78**	.76***	.78***	.55	.64*	.88*	.62*
PERFORMANCE		1.00	.33	.06	.11	-.14	-.10	-.07	-.37	-.32	-.29	.04	.00	.02	.12
TRAINING			1.00	.05	.64	.37	.36	.30	.11	-.06	.27	.18	.45	.31	.06
ATTENTION LEVEL				1.00	.45	.45	.67**	.30	.43	.42	.33	.57	.25	.55	.34
MOTIVATION					1.00	.47	.52	.27	.39	.29	.31	.32	.38	.54	.40
STRESS LEVEL						1.00	.69**	-.01	.67**	.67**	.70**	.41	.63*	.54	.52
ENERGY LEVEL							1.00	.27	.57	.56	.66**	.55	.99*	.57	.67**
PHYSICAL STATE								1.00	.00	.01	.05	.07	-.02	.19	.15
TIME PRESSURE									1.00	.68**	.72***	.52	.61*	.49	.42
TASK DIFFICULTY										1.00	.75***	.51	.59*	.44	.50
TASK COMPLEXITY											1.00	.55	.65*	.50	.34
ACTIVITY LEVEL												1.00	.46	.42	.53
PHYSICAL EFFORT													1.00	.45	.51
MENTAL EFFORT														1.00	.59*
SENSORY EFFORT															1.00

AUDITORY MONITORING TASK

OVERALL WORKLOAD	1.00	.00	.15	.18	.26	.66*	.43	-.09	.36	.32	.38	.40	.52	.17	.34
PERFORMANCE		1.00	.13	.32	.04	-.27	-.12	.14	-.48	-.21	-.28	.11	.10	-.08	.16
TRAINING			1.00	.23	.15	.29	.06	.16	-.10	.37	.26	.29	.54	.29	.52
ATTENTION LEVEL				1.00	.46	.10	.21	.29	-.02	.07	.05	.13	.16	.25	.29
MOTIVATION					1.00	.09	.46	.46	.19	-.02	-.10	-.15	.13	-.09	.18
STRESS LEVEL						1.00	.31	-.41	.52	.65*	.27	.30	.48	.39	.18
ENERGY LEVEL							1.00	.10	.20	.23	.38	.30	.33	.24	.31
PHYSICAL STATE								1.00	-.12	-.37	-.28	-.03	-.08	-.14	.24
TIME PRESSURE									1.00	.33	.26	.07	.38	.29	.10
TASK DIFFICULTY										1.00	.68**	.41	.62	.75***	.33
TASK COMPLEXITY											1.00	.38	.46	.62*	.29
ACTIVITY LEVEL												1.00	.52	.55	.29
PHYSICAL EFFORT													1.00	.45	.38
MENTAL EFFORT														1.00	.27
SENSORY EFFORT															1.00

TIME ESTIMATION TASK

Although no relationships were found between subjective ratings and objective performance measures, there were significant relationships among the rating scales. There was a positive correlation between ratings for Overall Workload and Stress Level for every task, and for Overall Workload and Task Difficulty for the tracking, Sternberg, and auditory monitoring tasks. Stress Level correlated positively with Task Difficulty on all four tasks, however, the ratings for perceived stress and difficulty were both higher on the tracking and Sternberg tasks. Overall Workload correlated significantly with Activity Level on the tracking and Sternberg tasks. Both Activity Level and Overall Workload were rated much higher on these two tasks than for either the auditory monitoring or time estimation tasks.

Subjective Performance ratings did not correlate with any other rating nor with any objective measure of performance. Subjects rarely rated their own performance as being below the 50 per cent level; they always felt that their performance was at least adequate. There were virtually no correlations between either Attention Level or Motivation and any other scale, indicating either that these two factors are unrelated to the other rated dimensions of workload or that there was little variation within these two factors. The latter explanation is most likely as all tasks required relatively continuous attention, there were no distractors, and the task levels were probably equivalently interesting (or uninteresting).

Time Pressure correlated with Stress Level on both the tasks for which the rate of presentation was manipulated, the Sternberg and auditory monitoring tasks. Mental Effort ratings showed no direct relationship with Overall Workload, but correlated with Stress Level for the tracking and Sternberg tasks.

The large number of correlations for the auditory monitoring task are somewhat surprising because the task demands were so low under all conditions, and there were absolutely no differences in performance. This combination of a monitoring and memory load task may have produced relatively complex interrelationships for the rating scale items examined here, but no performance differences could be seen because the subjects were so successful in meeting all task demands that their performance was error free. This illustrates the importance of evaluating workload-related factors independently from, and in addition to, performance, in order to gain a clear understanding of the operator's experience.

Relationship of Task and Experiment Rating Scales

The rating scale correlations for all sixteen task conditions were averaged in order to correlate the combined ratings for each of the tasks to those made at the end of the entire experiment. Six of the scales were significantly correlated: Performance (r=.645, p<.05), Training (r=.819, p<.01), Attention

Level (r=.585, p<.05), Motivation (r=.787, p<.01), Physical State (r=.767, p<.01), and Task Complexity (r=.592, p<.05). Ratings on these scales were apparently less affected by the particular task demands and were consistent across the length of the experiment. Evidently subjects perceived less variability between tasks and task conditions when making ratings on these scales. Subjects generally rated their performance as being moderately high, training was perceived as being neither too little nor too much, the attention level required was always perceived as being above the mid-point, motivation was also relatively stable at about the mid-point or slightly above, physical state remained in a range of 60 to 65%, and task complexity was rated relatively low on all tasks.

The lowest correlations were for the Activity Level and Task Difficulty scales where large differences were perceived between tasks; thus, the final rating of the entire experiment was probably an averaged perception of task demands which would not reflect the range originally recorded. There is at least some face validity in the results of this analysis; tasks did vary considerably in activity requirements and in difficulty levels. The low correlations for the three effort scales also reflect the considerably different demands imposed by the different tasks.

CONCLUSIONS

The rating scales used in this study were not selected for the purpose of producing a definitive subjective workload scale, but rather as exploratory probes for the purpose of identifying the factors which are consistently perceived as related to workload regardless of the type of task or condition imposed, and those factors which are perceived as task or condition dependent.

The two tasks that manipulated both rate of presentation and memory load, the Sternberg and auditory monitoring tasks, demonstrated clear within-task rating scale differences. This was particularly striking because there were no performance differences for the auditory monitoring task, and only weak effects on performance for the Sternberg task, providing further evidence that workload and performance may not covary. However, the different levels of task demands for each task affected the rating scale variability in different ways. The results are consistent with the fact that the rate of presentation and number of responses required were greater for the Sternberg task than for the auditory monitoring task. Both the rate of presentation and memory load affected the perception of Overall Workload for both tasks, but for the Sternberg task the rate of presentation also significantly affected ratings on 10 additional scales, compared to only 2 others for the auditory monitoring task. Memory load manipulation affected 7 of the scales for the Sternberg task and 8 for the auditory monitoring task with four

scales common to both; Overall Workload, Performance, Task Difficulty, and Task Complexity.

The tracking and time estimation tasks showed no within-task differences on the Overall Workload rating scale, suggesting that the differential perception of overall workload may be dependent on more discrete variations of task demands such as those imposed by the Sternberg and auditory monitoring tasks. The number and type or responses required for time estimation were held relatively constant over all levels of the time estimation task and were not clearly discriminable across the four levels of the tracking task. This assumption is supported by the fact that a total of only 5 of the rating scales were affected by variations in task demands within the tracking task, and only 6 of the scales for time estimation, whereas for the Sternberg and auditory monitoring tasks, 11 and 12 of the rating scales respectively demonstrated within task-differences on one or both task dimensions.

There were only two scales that consistently demonstrated within-task differences: Performance and Task Difficulty. Ratings on these scales were generally consistent with the demands made by the levels of the task, although the Performance scale reflected evaluations that apparently combined both the feeling of success and some element of discomfort or insecurity imposed by a particular condition. This was most clearly demonstrated by the time estimation task where the no counting condition resulted in much lower ratings of perceived adequacy of performance than did the counting condition, in spite of the fact that there were no actual differences in performance.

Mental Effort ratings were significantly affected by the within-task demands of all tasks except the tracking task, probably because the demands of the latter task were primarily limited to simple hand-eye coordination. Stress Level ratings were affected only by the tracking and Sternberg tasks, and Time Pressure and Energy Level ratings were affected only by those tasks which actually manipulated the rate of presentation--the Sternberg and auditory monitoring tasks, with the Energy Level generally enhanced by increased demands. However, the actual ratings on the Energy Level scale were higher for the two tasks that required the greatest measurable output--the Sternberg and tracking tasks. Similarly, the scales that reflected task demands were also rated higher than for the time estimation or auditory monitoring tasks.

Ratings of Overall Workload always covaried with Stress Level ratings, and Stress Level always covaried with Task Difficulty. Overall Workload covaried directly with Task Difficulty for all tasks except time estimation. These findings are generally concordant with relationships reported by McKenzie (1979), Steininger (1977), and others. However, Attention Level, which is often related to workload, was not perceived as be-

ing related to Overall Workload or any other rating scale item used in this study. This could have occurred either as a consequence of the definition provided for attention level, or because of the short duration of the experimental tasks (Enstrom & Rouse, 1977).

Further research is planned to determine the effects of a number of other simple flight-related activities on perceived workload and other related factors. Variation in levels and types of mental workload and other subtle but relevant-to-flight factors will be a major focus. The information thus obtained will then be applied to the analysis of different levels of such tasks when used as secondary measures of workload for primary flying tasks. The ultimate goal is to derive a metric sensitive to the combined effects of the many component tasks involved in flying on pilot workload.

REFERENCES

Bindra, D., & Waksberg, H. Methods and terminology in studies of time estimation. Psychological Bulletin, 1956, 53, 155-159.

Bird, K. L. Subjective rating scales as a workload assessment technique. Proceedings of the 17th Annual Conference on Manual Control, Los Angeles, June 1981.

Bird, K. L., & Hart, S. G. Effect of counting and tracking on verbal and production methods of time estimation. Proceedings of the 16th Annual Conference on Manual Control, Cambridge, May 1980.

Enstrom, K. D., & Rouse, W. B. Real time determination of how a human has allocated his attention between control and monitoring tasks. IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-7, No. 3, 153-161, March 1977.

Hart, S. G., Childress, M. E., & Hauser, J. R. Individual definitions of the term "workload". Proceedings of the Psychology in the DOD Symposium, U.S. Air Force Academy, Colorado Springs, April 1982.

Hart, S. G., McPherson, D., & Loomis, L. Time estimation as a secondary task to measure workload: Summary of research. Proceedings of the 14th Annual Conference on Manual Control, Los Angeles, April 1978.

Hicks, T. G., & Wierwille, W. W. Comparison of five mental workload assessment procedures in a moving-base simulator. Human Factors, 1979, 21, 129-144.

Jex, H., & Clement, W. F. Defining and measuring perceptual motor workload in manual control tasks. In N. Moray (Ed.), Mental workload: Theory and measurement, New York: Plenum, 1979.

Kennedy, R. S. A comparison of performance on visual and auditory monitoring tasks. Human Factors, 1971, 13, 93-97.

Kennedy, R. S. A sixty-minute vigilance task with 100 scoreable responses. Bureau of Medicine & Surgery Doc. No. MF12.524.002-5002.11, Naval Aerospace Medical Institute, Pensacola, Florida, July 1968.

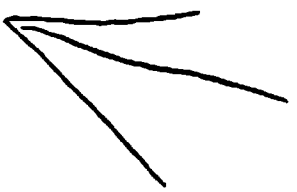
McKenzie, R. E. Concepts of stress. In B. O. Hartman & R. E. McKenzie (Eds.), Survey of methods to assess workload. NATO/AGARDograph, AG-246, London, 1979.

Moray, N. Subjective mental workload. Human Factors, 1982, 24, 25-40.

Steininger, K. Subjective ratings of flying qualities and pilot workload in the operation of a short haul jet transport aircraft. Proceedings of AGARD Conference on Studies on Pilot Workload, AGARD-CP-217, April 1977.

Sternberg, S. High-speed scanning in human memory. Science, 1966, 153, 652-654.

Williges, R. C., & Wierwille, W. W. Behavioral measures of aircrew mental workload. Human Factors, 1979, 21, 549-574.



AD P001624

THE SENSITIVITY OF TWENTY MEASURES OF PILOT MENTAL WORKLOAD IN A SIMULATED ILS TASK

by

Walter W. Wierwille and Sidney A. Connor
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

ABSTRACT

Twenty workload estimation techniques were compared in terms of their sensitivity to changes in pilot loading in an ILS task. The techniques included opinion measures, spare mental capacity measures, physiological measures, eye behavior measures, and primary task measures. Loading was treated as an independent variable and had three levels: low, medium, and high. The load levels were obtained by a combined manipulation of windgust disturbance level and simulated aircraft pitch stability. Six instrumented-rated pilots flew a moving-base general aviation simulator in four sessions lasting approximately three hours each. Measures were taken between the outer and middle markers.

Two opinion measures, one spare mental capacity measure, one physiological measure, and one primary task measure demonstrated sensitivity to loading in this experiment. These measures were: Cooper-Harper ratings, WCI/TE ratings, time estimation standard deviation, pulse rate mean, and control movements per unit time. The Cooper-Harper ratings, WCI/TE ratings, and control movements demonstrated sensitivity to all levels of load, whereas the time estimation measure and pulse rate mean showed sensitivity to some load levels.

The results ~~of this experiment~~ demonstrate that sensitivities of workload estimation techniques vary widely, and that only a few techniques appear to be sensitive in this type of ILS task, which emphasizes psychomotor behavior.

INTRODUCTION

One of the major problems in mental workload estimation is the lack of available information on the sensitivity of various workload estimation techniques [1,2]. When a researcher or human factors engineer needs to assess workload in a given experimental situation, it is not clear which technique or techniques should be used [3]. The danger is that insensitive techniques may be used. If so, experimental results will show no differences in workload when in fact there are differences.

Sensitivity in regard to workload estimation can be defined as the relative ability of a given workload estimation technique to discriminate statistically significant differences in operator loading. High sensitivity requires

discriminable changes in the score means as a function of load level and low variation of the scores about the means. When sensitivity is defined in this way, it becomes subject to experimental determination. Based on experiments that emphasize specific operator behaviors, it should be possible to predict which given techniques are sensitive.

An experiment directed at evaluating the sensitivity of workload estimation techniques in a psychomotor task has been completed and is reported briefly in this paper. An ILS piloting task was used for the evaluation. (For a more detailed discription of the experiment and results, see reference [4]).

EXPERIMENT

Subjects

Six male instrument-rated pilots served as subjects in this experiment. The flight time of the subjects ranged from 500 to 2700 hours with a mean of 1300 hours.

Apparatus

The primary apparatus in this experiment was a modified fli task simulator (Singer Link, Inc., General Aviation Trainer, GAT-1B). T imulator had three degrees of freedom of motion (roll, pitch, and yaw). ansulucent blinders were used to cover the windows of the simulator to re e outside distractions and cues and to aid in the control of cockpit illumination.

Several modifications to the flight simulator were made for the experiment. These modific.tions permitted primary task load manipulation, secondary task operations, re:-nse measurement, and scoring. Primary task load manipulation was accomplished by changing aircraft pitch stability and random wind-gust disturbance level simultaneously. Three load conditions were developed: low, medium, and high, as shown in Table 1. Table 2 provides a list of the workload measurement techniques selected for inclusion in the present study.

Experimental Design

A complete 3 x 20 within-subject design was used for the sensitivity analysis. Load was the factor with three levels. Measurement technique (Table 2) was the factor with twenty levels.

Workload measures from different techniques were taken simultaneously on some of the data collection runs. Only those measures which were not likely to affect each other were taken simultaneously. Table 3 shows the scheme used for combining different measurement techniques for data collection. The combination of measurement techniques shown in the table was, to an extent, based on previous investigations of workload. Hicks and Wierwille's [3] study supported the combination in condition 2. The two rating scales were administered

in separate measurement conditions to prevent the ratings on one scale from biasing the ratings on the other scale. The secondary task measures were divided among several conditions because of potential intrusion and interference. Vocal measures were recorded from the two secondary tasks which required a verbal response as per Schiflett and Loikith's [5] recommendation.

It should be noted that primary task measures were recorded on all subjects and on all data collection flights for the intrusion analysis. However, only data from measurement condition 1 were used for the sensitivity analysis of the primary task measures.

General Procedure

After receiving instructions, subjects flew nine familiarization flights in the simulator. These flights were similar, but not the same as, the data collection flights. All subjects flew the familiarization flights in the same order. Steady crosswinds were introduced for each run, and subjects were given heading corrections.

After the familiarization session, the subjects participated in three data collection sessions. The familiarization session and each data collection session were held on a different day.

Each data collection session consisted of two sets of a warm-up practice flight and three data collection flights. The practice flight was the same as the first data collection flight. Since the data collection flights were counterbalanced, equal amounts of practice were provided for the low, medium, and high load conditions. The data collection flights also contained steady crosswind conditions, for which the subject was given heading corrections. The purpose of introducing steady crosswinds was to disguise the load conditions, thereby requiring subjects to fly each flight as a separate entity.

Flight Task Procedures

The flight task in this experiment was an ILS approach in the Singer Link GAT-1B aircraft simulator. Prior to the beginning of a flight, the simulated aircraft was positioned on the ground 5 miles outbound from the outer marker on the 108 degree radial, heading into the wind. When ready to begin, the experimenter informed the subject of the wind direction and speed, and gave him a heading correction for the crosswind. When contacted by the experimenter, the subject took off and climbed to 2000 feet. The subject then flew directly to the outer marker by following the localizer at 100 miles per hour until the glide slope was intercepted. Upon interception of the glide slope, the subject reduced airspeed to 80 miles per hour and proceeded down the glide slope while following the localizer to a landing. Data were recorded between the outer and middle markers. For the opinion measures, subjects gave ratings for the flight segment between the outer and middle markers immediately after landing and parking the simulated aircraft.

RESULTS

The computed scores for each technique were first converted to Z-scores (normalized scores) so that technique measure units would not affect the sensitivity analysis. Subsequently, an overall analysis of variance was performed on the scores. Since Z-scores were used, a technique main effect was not possible. A significant main effect of load was found, F (2,10) = 5.34, $p < 0.0001$, and a significant load by technique interaction was found, F (38,190) = 2.76, $p \leq 0.05$.

The load by technique interaction indicated that the measurement techniques were differentially sensitive to load. Therefore, individual ANOVAs were used to isolate the sensitive techniques.

The individual ANOVAs indicated that five of the twenty measures were sensitive. They were the Cooper-Harper scale F (2,10) = 16.39, $p = 0.0007$; the Workload-Compensation-Interference/Technical Effectiveness (SCI/TE) scale, F (2,10) = 31.15, $p < 0.0001$; the time estimation standard deviation, F (2,10) = 5.69, $p = 0.022$; the pulse rate mean, F (2,10) = 8.89, $p = 0.006$; and the control movements measure, F (2,10) = 33.34 $p < 0.0001$. The normalized means for each technique are plotted in Figures 1 through 5 as a function of load.

Newman-Keuls comparisons were then performed on the normalized means of the sensitive measures. The comparisions included low vs. medium, medium vs. high, and low vs. high load conditions. Results indicated that all differences were significant at $p < 0.05$, except for pulse-rate mean (low vs. medium and medium vs. high) and time estimation standard deviation (low vs. high).

A logical classification of techniques based on demonstrated sensitivity was generated from an examination of the Newman-Keuls comparisons, as shown in Table 4. Techniques which demonstrated sensitivity to all pairs of load conditions (i.e., low vs. medium, medium vs. high, and low vs. high) were included in class I. These measures are preferred over other techniques which demonstrated only partial sensitivity, or no sensitivity in the present study. Techniques which showed sensitivity to some differences in load conditions (but not all) were included in class II. These measures are less preferred than class I techniques, but are more preferred than class III techniques. Class III techniques did not demonstrate sensitivity to load in the present study. This class includes all techniques except those in class I and class II.

One possible reason that only five of the twenty workload assessment techniques demonstrated sensitivity in the present study is that the other techniques simply required a greater number of subjects to show a significant effect of load. It is possible to estimate the sample size required to detect a reliable load effect for a given workload assessment technique at specified levels of significance and power. These calculations were performed for techniques which did not demonstrate sensitivity in the present study, to provide an indication of the practical costs of achieving statistical significance. The procedure used for estimating the sample size required for finding sensi-

tivity is described by Bowker and Lieberman [6]. Sample sizes were estimated for a significance level of 0.05 and for a power of approximately 0.80. The results of these estimates are presented in Table 5.

CONCLUSIONS

This study has shown that five measures of workload estimation were sensitive indicators of load in a piloting task that is predominantly psychomotor in nature. Another fifteen measures, believed to be "good" measures of workload, showed no reliable effect. The main conclusion that must be drawn from the study is that few measures are sensitive to psychomotor load.

Of the five techniques demonstrating sensitivity, only three exhibited monotonic score increases with load as well as statistically reliable differences between all pairs of load levels. Consequently, only the three meet all criteria for sensitivity to psychomotor load. These class I techniques are the ones that are recommended for measurement of psychomotor load:

Cooper/Harper ratings,
WCI/TE ratings, and
Control movements per second.

The other two techniques showed sensitivity to psychomotor load, but did not discriminate between all pairs of load levels. These class II techniques are:

Time estimation standard deviation, and
Pulse rate mean.

These measures would be helpful in evaluating psychomotor load, but they should not be relied on exclusively. At least one class I technique should also be used in conjunction with these measures.

It is worth noting that only two opinion measures were taken in the present experiment, and both proved sensitive. This suggests that well-designed rating scales are among the best of techniques for evaluating psychomotor load. In regard to the primary task measures, the control movements measure alone was sensitive. However, this measure is also the only primary task measure which reflected "strategy" of the pilot. Consequently, one could speculate that selecting a primary task measure that reflects strategy will most likely result in good sensitivity.

Fifteen (techniques) measures showed no reliable change as a function of load. When these fifteen measures were subjected to a power analysis to determine sample size, the number of subjects required ranged from 12 to well over 100 (Table 5). One can only conclude that at best the fifteen measures, as taken, are much less sensitive to psychomotor load than the five appearing in Classes I and II. Of course, there is always the possibility that the measures would be sensitivite to loading along other dimensions of human performance, such as psychomotor tasks of a different nature, or mediational or cognitive tasks, for example.

In general, the results of the experiment show that there are wide variations in the sensivity of workload estimation measures. Great care must be taken in selecting measures for a given experiment. Otherwise, it is possible that no changes in workload will be found, when indeed there are changes.

REFERENCES

1. Wierwille, W. W. and Williges, R. C. Survey and analysis of operator workload assessment techniques. Blacksburg, Virginia: Systemetrics, Inc. Report No. S-78-101, September, 1978.

2. Wierwille, W. W. and Williges, B. H. An annotated bibliography on operator mental workload assessment. Patuxent River, Maryland: Naval Air Test Center Report No. SY-27R-80, March, 1980.

3. Hicks, T. G. and Wierwille, W. W. Comparison of five mental workload assessment procedures in a moving base driving simulator. Human Factors, 1979, 21, 129-143.

4. Connor, S. A. and Wierwille, W. W. Comparative evaluation of twenty pilot workload assessment measures using a psychomotor task in a moving base simulator. Moffett Field, CA: NASA-Ames Research Center, (Forthcoming report).

5. Schiflett, S. G. and Loikith, G. J. Voice stress as a measure of operator workload. Patuxent River, Maryland: Naval Air Test Center, Technical Memorandom TM 79-3 SY, December 31, 1979.

6. Bowker, A. H. and Lieberman, G. J. Engineering statistics. New Jersey: Prentice-Hall, Inc., 1959.

ACKNOWLEDGEMENTS

The authors wish to thank Mrs. Sandra Hart, NASA-Ames Research Center, for helpful technical suggestions. This work was sponsored under NASA grant NAG2-17.

TABLE 1

Primary Task Load Conditions

	LOAD CONDITION		
	Low	Medium	High
RANDOM GUST LEVEL	Low	Medium	High
Estimated Std. Dev. (mph)	0	2.7	5.9
PITCH STABILITY	High	Medium	Low
a. Control input to pitch rate output equivalent gain (degrees/s per % of control range)	0.522	3.560	7.83
b. Control input to pitch rate output equivalent time constant(s)	0.097	0.660	1.45

TABLE 2

Workload Assessment Techniques Which Were Tested in the Present Experiment

---

OPINION

1. Cooper-Harper Scale
2. WCI/TE Scale

SPARE MENTAL CAPACITY

3. Digit Shadowing (% errors)
4. Memory Scanning (Mean time)
5. Mental Arithmetic (% errors)
6. Time Estimation Mean (Seconds)
7. Time Estimation Standard Deviation (Seconds)
8. Time Estimation Absolute Error (Seconds)
9. Time Estimation RMS error (Seconds)

PHYSIOLOGICAL

10. Pulse Rate Mean (Pulses per minute)
11. Pulse Rate Variability (Pulses per minute)
12. Respiration Rate (Breath cycles per minute)
13. Pupil Diameter (Normalized units)
14. Voice Pattern (Digit Shadowing Task)
15. Voice Pattern (Mental Arithmetic Task)

EYE BEHAVIOR

16. Eye Transition Frequency (Transitions per minute)
17. Eye Blink Frequency (Blinks per minute)

PRIMARY TASK

18. Localizer RMS Angular Position Error (Degrees)
19. Glide Slope RMS Angular Position Error (Degrees)
20. Control Movements per second
    (Aileron + Elevator + Rudder)

TABLE 3

Combination of Measurement Techniques
for Data Collection

Measurement Condition	Measurement Techniques
1.	Cooper-Harper Scale Pupil Diameter Eye Transition Frequency Eye Blink Frequency Localizer RMS Error Glide Slope RMS Error Control Movements
2.	WCI/TE Scale Pulse Rate Mean Pulse Rate Variability Respiration Rate
3.	Digit Shadowing Voice Pattern
4.	Memory Scanning
5.	Mental Arithmetic Voice Pattern
6.	Time Estimation (Mean) (Std. Dev.) (Abs. Error) (RMS Error)

TABLE 4

Logical Classification of Techniques
Based on Demonstrated Sensitivity

Class I: Complete Sensitivity Demonstrated
- Cooper-Harper Scale
- WCI/TE Scale
- Control Movements/Unit Time

Class II: Some Sensitivity Demonstrated
- Time Estimation Standard Deviation*
- Pulse Rate Mean**

Class III: Sensitivity Not Demonstrated
- All Other Techniques (See Table 5)

*Double valued function
**Limited sensitivity

TABLE 5

Estimated Sample Sizes Required for Achieving a Significant Load Effect for Techniques not Demonstrating Sensitivity

Technique	Estimated Sample Size
SPARE MENTAL CAPACITY	
Digit Shadowing	18
Memory Scanning	>100
Mental Arithmetic	25
Time Estimation (Mean)	53
Time Estimation (Abs. Error)	>100
Time Estimation (RMS Error)	53
PHYSIOLOGICAL	
Pulse Rate Variability	45
Respiration Rate	15
Pupil Diameter	>100
Speech Pattern (D. Shadow.)	28
Speech Pattern (M. Arith.)	>100
EYE BEHAVIOR	
Eye Transition Frequency	42
Eye Blink Frequency	25
PRIMARY TASK	
Localizer RMS Error	12
Glide Slope RMS Error	41

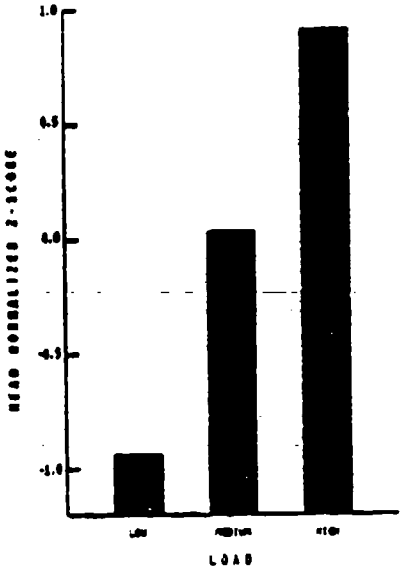


Figure 1. Mean normalized scores for the Cooper-Harper rating scale measure plotted as a function of load.

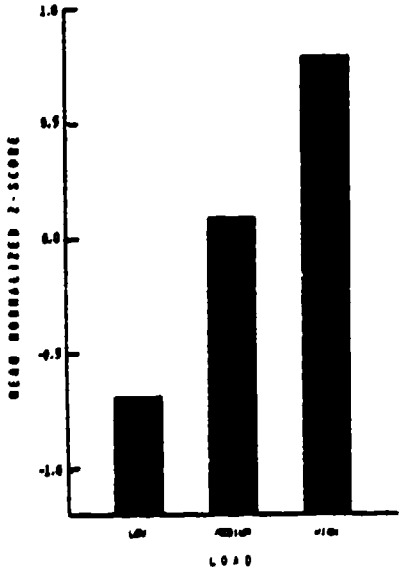


Figure 2. Mean normalized scores for the WCI/TE rating scale measure plotted as a function of load.

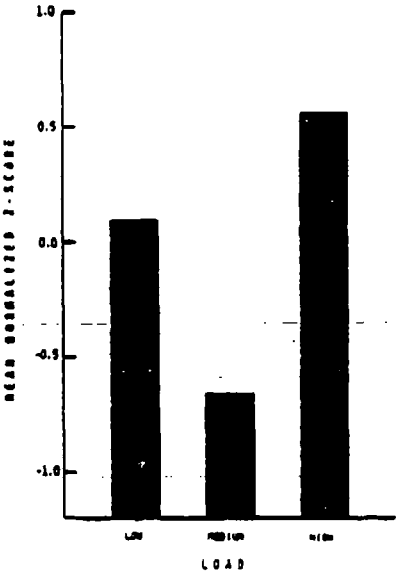


Figure 3. Mean normalized scores for the time estimation standard deviation measure plotted as a function of load.

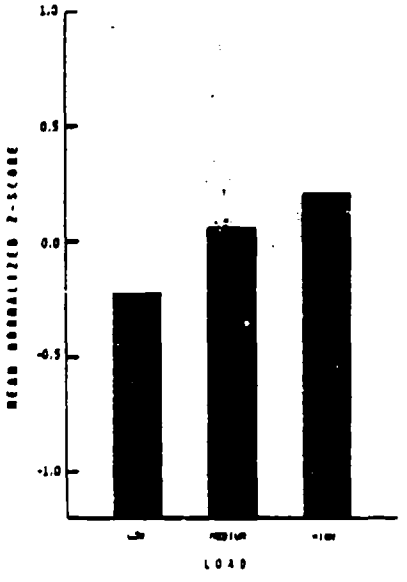


Figure 4. Mean normalized scores for the pulse rate mean measure plotted as a function of load.

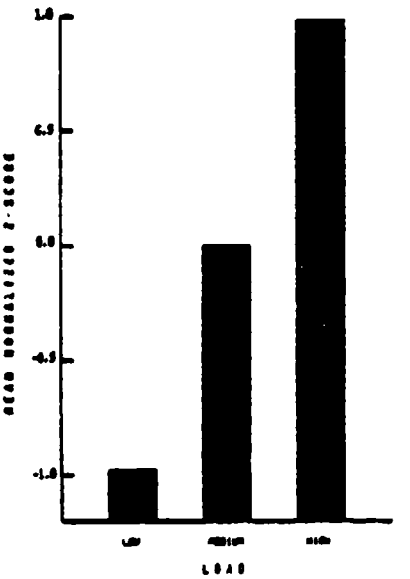


Figure 5. Mean normalized scores for the control movements measure plotted as a function of load.

SESSION 3: SIMULATION AND MODEL BASED ANALYSIS

Moderator: Andrew M. Junker, Air Force Aerospace Medical Research Lab

3

ABSTRACT

for

18th ANNUAL CONFERENCE ON MANUAL CONTROL

Title of Paper: Development of a G-Seat Roll-Axis Drive Algorithm

Authors: Edward A. Martin (ASD/ENETS, W-PAFB, OH, 255-4408) and Grant R. McMillan (AFAMRL/HEF, W-PAFB, OH, 255-3325)

The paper reports the design and evaluation of a g-seat drive algorithm developed in preparation for a comparison of g-seat versus whole-body roll-axis cuing. Force transducers were used to measure pressures on the buttocks during roll motion in the Advanced Low Cost G-Cuing System (ALCOGS) and the Roll-Axis Tracking Simulator (RATS). The tests were conducted under a range of amplitude and frequency conditions. The results indicated that pressures in the ALCOGS are a function of seat pan roll amplitude. In the RATS (a whole-body motion simulator) pressures varied as a function of roll amplitude and roll acceleration. Using these data, a drive equation was developed for the ALCOGS that produces buttocks pressures similar to the RATS. This drive equation was then evaluated in a pilot study which compared human performance on a disturbance regulation task under a static condition and under motion conditions in the two simulators.

Compensation For Time Delays In Flight Simulator Visual Display Systems

D. FRANCIS CRANE
NASA, Ames Research Center
Moffett Field, California

There is a trend toward the use of computer-generated imagery (CGI) systems to generate flight simulator out-of-the-window visual scenes. CGI visual systems offer important advantages including large field-of-view, large gaming area, multiple-observer viewpoint, and moving targets. CGI systems construct a visual display from a description of the scene stored in a computer. The image construction time, though short (~100 msec), introduces a delay into the pilot-aircraft system. Several authors (Gum and Albery 1977;Larson and Terry 1975) have reported simulation problems traced to time delays in visual system cueing. The multi- million dollar simulator evaluated by Decker (1980) was rated unsatisfactory for training pilots to perform precision flight tasks - at least in part because of CGI delays.

A piloted aircraft can be viewed as a closed-loop man-machine control system. When a pilot is performing a precision maneuver in a simulator, a delay in the visual display of aircraft response to pilot-control input decreases the stability of the pilot-aircraft system. The less stable system is more difficult to control precisely. Pilot dynamic response and performance change as the pilot attempts to compensate for the decrease in system stability. The changes in pilot dynamic response and performance bias the simulation results by influencing the pilot's rating of the handling qualities of the simulated aircraft. (Crane 1980).

An approach to display-delay compensation based on conventional control system design principles was evaluated in an earlier study (Crane 1981). In that study, the pilot's task was to maintain precise attitude control in simulated turbulent atmospheric conditions - a small CRT displayed attitude information to the pilot. The compensation was effective in that delay-induced changes in pilot performance and dynamic response were substantially reduced (Crane 1981). The display-delay compensation approach was re-evaluated in a simulation wherein helicopter pilots performed a precision attitude control task using only visual cues from a CGI display. The design and preliminary results of this just completed study are described.

REFERENCES:

Crane, D. F. (1980), Time Delays in Flight Simulator Visual Displays, Proceedings 1980 Summer Computer Simulation Conference. Seattle, Washington.

Crane, D. F., Flight Simulator Visual-Display Delay Compensation, Proceedings 1981 Winter Simulation Conference Atlanta, Georgia.

Decker, J. F. (1980), Qualification Operational Test and Evaluation of the FB-111A Weapon System Trainer Visual System Attachment. USAF Strategic Air Command, Final Report, Project 77-SAC-331.

Gum, D. R. and W. B. Albery (1977), Time-Delay Problems Encountered in Integrating the Advanced Simulator for Undergraduate Pilot Training, Journal of Aircraft, Vol. 14, pp. 327-332.

Larson, D. F. and D. Terry (1975), ASUPT Visual Integration Technical Report, Singer-Simulation Products Division, Technical Report ASUPT 82, Binghamton, New York.

AD P001625

AN AUTOMOBILE AND SIMULATOR COMPARISON OF DRIVER BEHAVIOR IN A COLLISION AVOIDANCE MANOEUVRE

Eric N. Solowka
Atlantis Flight Research Inc.*
3924 Chesswood Dr., Downsview
Canada

Lloyd D. Reid
University of Toronto
Institute for Aerospace Studies
4925 Dufferin St., Downsview
Canada

ABSTRACT

A collision avoidance manoeuvre in which drivers were required to successfully steer around a suddenly appearing obstruction in the vehicle's path was investigated. Both an instrumented automobile and a matching fixed base, computer generated image simulator were employed. Driver performance and the simulator's validity in representing an extreme lateral manoeuvre are examined.

INTRODUCTION

Variations of the linear model of a human controller (McRuer, Kendel, 1973) have been applied with some success to various driving scenarios. Using a fixed base simulator it has been demostrated (Reid, Graf, Billing, 1980a), that for a simple driving task in which drivers were required to maintain lane position while manoeuvring along a serpentine roadway, that driver behavior can be represented in the average sense using a linear model. The technique was extended (Reid, Graf, Billing 1980b) to model driving response for an extreme lateral control task. That is, one in which subjects were required to respond to a suddenly appearing obstacle in the vehicle's path by steering around it. Data obtained from a fixed base simulator was again used to validate the model with relative success. Although good fits to the experimental data were obtained, the resultant model parameters were not always intuitively satisfying sometimes having values that either seemed to have the wrong magnitude or sign. Another experiment in which the collision avoidance was modelled (Maeda, Irie, Hidaka and Nishimura) involved automobile data for model verification and also yielded good model fits to the data.

The current investigation of linear modelling techniques involved an experiment to generate data in both an automobile and a matching fixed-base, computer generated image driving simulator. The availability of data for both vehicles allows the validity of the simulator in eliciting driver response for the collision avoidance manoeuvre to be

* This work was undertaken at the Institute for Aerospace Studies and supported by General Motors Inc.

determined. If the simulator is valid in its representation of the automobile for this type of manoeuvre, then it may be unequivocally used as a tool to validate collision avoidance driver models. However, the nature of the task with extreme lateral motions, and the absence of acceleration cues and limited visual fidelity of the simulator render suspect its ability to elicite the same driver behavior as the automobile. The present paper discusses the results of the experiment, addressing simulator validity for the collision avoidance manoeuvre and driver behavior in performing the manoeuvre.

EXPERIMENT

The collision avoidance scenario was as follows. The task was to maintain lane position via the steering wheel while driving under cruise control at 50 kilometers per hour down either lane of a two lane divided highway with no other cars present. The shoulder of the lane in which the subject was driving was lined with poles at regular 20 metre intervals. The poles could be remotely triggered to fall in front of an approaching vehicle and cover 80% of the established lane thereby creating an obstruction (figure 1). A successful collision avoidance manoeuvre required the drivers to miss the obstruction by steering, and only steering (no braking), into the adjacent lane and then returning as rapidly as possible to the original lane without striking either lane's shoulder. Varying degrees of urgency, or severity of the manoeuvre were created by varying the preview distance ahead of the approaching vehicle at which the unknown obstructing pole would begin to fall. In less severe cases the falling pole was identified for the driver or even left completely down to represent an obstruction with no surprise element.

Subjects were required to drive both the automobile and simulator on alternate days throughout a one month period during which the experiment was conducted. On consecutive days however, one for the automobile and one for the simulator, the same task was driven.

The test track (figure 2) was laid out on a flat, straight one kilometer roadway with the centre 520 metres lined with 27 potentially obstructive poles. The roadway centre line was represented by a continuous 4 inch white line and the far shoulder identified by a series of cones. This scene was chosen because it could be readily simulated using the available vector display system of the simulator (figure 3). The solid centre line on the track was required to facilitate the optical lane tracking system used. The simulator's display centre line however was represented by a dashed line to provide some textural visual information.

The automobile (figure 4) employed was a compact North American sedan. Instrumention featured an on-board PDP 11/03 with 2 floppy discs for data acquisition as well as two lane trackers, one facing forward and one backward, to permit the accurate determination of automobile lane position and heading angle at all times throughout the experiment. This was

particularly critical during the actual collision avoidance manoeuvre when the vehicle made significant excursions in both lane position and heading angle from nominal lane tracking values.

The driving simulator (figure 5) is fixed base with a large screen vector refresh display. The equations of motion are linear with parameters either determined experimentally or supplied by the manufacturer. The tire model was non-linear due to the extreme nature of collision avoidance tasks. Subjects adjusted well to the simulator with no ill side effects and all concurred in a post experiment questionaire that the simulator was an accurate representation of the actual driving task.

EXPERIMENTAL PROCEDURE

Eight subjects with driving experience ranging from 6 to 11 years were selected from 12 drivers solicited at the Institute for Aerospace Studies. The drivers were all graduate students whose ages fell in the mid to late twenties range. One subject dropped out of the experiment, and due to time constraints could not be replaced, leaving seven subjects. For their particiption in the experiment subjects were paid a fee of $50.00 plus an added incentive pay of $.50 for each successful run completed and a penalty charge of $2.50 levied for each unsuccessful manoeuvre.

The procedure (summarized in Table 1) consisted of three fundamental phases: transfer of learning, training, and production runs. During the transfer of learning phase the subjects were divided into two equal groups. This grouping was based on a questionaire completed by the subjects regarding their driving experience. Group 1 commenced the experiment in the simulator doing an obstacle avoidance manoeuvre for the case where the pole location was unknown and the preview distance, defined as the distance ahead of the approaching vehicle at which the unknown pole would begin to fall, set at 52 metres . The following day, the same task was performed in the automobile. The third day, again in the simulator, the collision avoidance manoeuvre was performed with a 32 metre preview followed on the forth day by the same task performed in the automobile. Group 2 performed the same tasks, except that the order of the vehicles was reversed, ie Group 2 began the experiment with the automobile. The training phase of the experiment imposed no specific vehicle order of presentation to the subject, except that a task performed for the first time in either vehicle had to be repeated on the following day in the alternate vehicle. The subjects were trained in stages for manoeuvres in both directions and with varying degrees of severity. In ascending order of severity the different cases were:

fixed obstacle - a pole was down and remained down throughout the run. No surprise element.

designated obstacle - the falling pole was identified by placing a marker cone beside it. The preview distance which could be 1 of 2 cases (52 or 32 metre) was revealed to

the subject. Little surprise element since subject knew location of pole.

emergency case - the obstructing pole was unknown to the subject. It was always either a 52 or a 32 metre preview. Most severe of the manoeuvres although it was not a complete surprise as subjects would anticipate an event.

At the completion of the training phase all subjects were competent in handling either the automobile or the simulator for the varying tasks. The production runs required subjects to perform in a random order the fixed obstacle case, the emergency case, and another case during which no pole falling event took place. Thus the final scenario for the production runs was one of driving along a roadway, with the potential for an emergency to occur, or an obstructing hazard might be constantly present, or no event at all occurs.

ANALYSIS and RESULTS

Data was recorded at the rate of 10 Hz in the automobile and 20 Hz in the simulator for a total of 30 seconds during each run. This interval was centered around the time that the vehicle's centre of gravity (C of G) passed the obstructing pole. Sample plots for a typical 32 metre preview of the emergency type are given in figure 6. The manoeuvre was one that required going from an established path in the right lane to the left lane in order to avoid striking the obstacle. Lane position is the lateral position of vehicle's C of G with respect to the roadway centre line. Positive being to the right. Heading angle was with respect to the centre line as well, with the clockwise sense as viewed from above being positive. A positive steering wheel angle causes a positve yaw rate in the same sense as the heading angle. The vertical bar on the time axis at approximately the 13 seconds mark represents the time when the obstructing pole began its fall.

Performance parameters for the collision avoidance manoeuvre were calculated from the recorded time histories for each run. These parameters, with their definitions are listed in table 2. Table 3 contains parameter values averaged over all subjects for the emergency manoeuvres performed during production runs. Analysis of the data was done using the TTEST procedure of the Statistical Analysis System (SAS). Differences in the analysis were required to be significant to the level $p < .05$ and substantial, ie greater than experimental errors, before they were accepted.

Validation of the simulator in its ability to represent the automobile for the collision avoidance manoeuvre was sought within the contexts of transfer of learning, absolute similitude, and relative similitude.

Transfer of Learning

As an a posteriori check on the equality of driving ability between the two groups of subjects formed for the transfer of learning phase, their

performance during production runs was compared for the same types of runs employed during the transfer of learning. It was found (table 4, with condition equal to .../PRODUCTION) that although the two groups were similar with respect to their performance in the automobile, Group 2 which started training in the automobile, exhibited a more sluggish response in the simulator than did Group 1. This is evidenced by longer steering wheel rise, delay, peak, and damping times. Significant and substantial effects (table 4, with condition equal to .../TRANSFER) not disqualified by this anomality still indicate that transfer of learning did occur. Group 2 had a 20% longer response time and a 14% greater steering wheel angle peak in the automobile for the 52 metre preview case than did Group 1 which had prior training in the simulator. Group 2 exhibited a greater number of steering wheel reversals in the simulator than did Group 1 for both the 52 and 32 metre preview cases.

Absolute Similitude

Comparing the subjects performance in both the automobile and the simulator for all the production run experimental cases yielded the following results (see table 5). With respect to vehicle motion, the simulator exhibited a 20% greater maximum heading angle deviation for all cases. The distance over the center line was also greater for the simulator in all but the right emergency manoeuvres. The difference was about 65 cm which represented quite a substantial deviation (65%) from the automobile equivalent. Steering wheel activity showed differences as well. The number of reversals was 30% greater for the simulator, but only for right manoeuvres. The rise, delay and peak times were all greater for the subjects in the simulator for the 32 metre preview cases, although the simple response time, or time to react to the initial appearance of the obstructing pole was not significantly different between the automobile and simulator.

Relative Similitude

Table 6 lists the differences in performance between the varying experimental cases. The Designated vs Fixed and Emergency vs Fixed comparisons do not include time parameters (T RES...T DAMP) because in the fixed obstruction case the pole is always down and doesn't have an associated trigger time thereby rending these parameter indeterminate.

The Designated vs Emergency and 52m vs 32m preview comparisons do include all the performance parameters.

The Designated vs Fixed case exhibited a 30% greater path deviation and a 26% greater heading angle deflection for the designated obstacle case. This difference observed for the automobile was not present for the simulator. This same trend occured for the Emergency vs Fixed comparison in which the emergency obstacle case exhibited a 30% greater path deviation and a 20% greater heading angle deflection.

Comparing the designated and emergency cases yielded only a single significant difference. The response time to the falling obstacle was 35%

longer for the emergency case but only for the simulator 32 metre preview case.

The 52 vs 32 metre preview comparison showed the heading angle greater by 20% and the steering wheel peak greater by 40% for the 32 metre preview in the automobile only. However, for both the automobile and simulator, the time parameters T RES, T DELAY, T PEAK, T SPARE and T DAMP were substantially longer for the 52 metre case. The differences in T RES, T DELAY and T PEAK being quite large, ie 150%, 100% and 78% larger respectively.

DISCUSSION

Although there existed some inate differences between the assigned groups in the transfer of learning phase, a transfer of learning was exhibited. For subjects trained first in the simulator on the 52 metre preview case and then followed by the automobile, a positive transfer occured in the T RES performance parameter. That is, response was quicker to the falling pole than for the group with no previous training. Also, positing that a smaller steering wheel peak is associated with a lower work ouput and hence more desirable, a postive transfer occured for SW PEAK as well since the simulator trained group exhibited smaller steering wheel peak angles in their automobile performance. These effects however did not occur for the second part of the transfer of learning when subjects were required to perform a 32 metre preview avoidance manoeuvre. A possible explanation might be that subjects required no extra training having adapted directly to the more severe 32 meter case based only on practise of the 52 metre case in both the automobile and simulator. Another transfer effect was exhibited although it was negative . The group transfering from the automobile to the simulator demonstrated a greater number of steering wheel reversals in the simulator. Thus it would seem that caution must be exercised when switching from a high fidelity system, the automobile, to one with less fidelity, the simulator - lacking in acceleration cues, peripheral vision cues and visual textural cues.

In the context of absolute similitude there was no signficant difference in subject response time (T RES) between the car and simulator. This suggests that all the cues required or utilized by the driver in detecting the obstacle were well represented in the simulation. The simulator trajectory demonstrated greater perterbations (as determined by D OVER C and PSI MAX) than did the automobile. Since the simulator visual image and hood outline were calibrated to represent a perfect perspective view of the actual automobile's hood and roadway scene the conclusion drawn is that in the absence of acceleration cues, the visual display system with its limited field of view and textural cues was not sufficient to allow the subject to control the simulator trajectory in the same absolute manner as the automobile. The number of steering wheel reversals for both left and right manoeuvres in the automobile was the same. The simulator however showed a greater number of reversals for the right manoeuvre than the left. The right manoeuvre required that on returning to the original lane after avoiding the obstacle the subject realign the vehicle's path

using roadway cues from the left lane edge. This unfamiliarity of driving with respect to the left lane edge combined with the limited visual and no lateral acceleration cues might account for the increase in the number of steering wheel reversals.

Any parameter that shows good absolute similitude should also exhibit a good relative similitude. This was the case with the response time. T RES was approximately 1 second longer for the 52 metre case than the 32 metre case for both manoeuvre directions and both vehicles. In comparing the Designated vs Emergency cases T RES was significantly shorter for the simulator designated 32 metre preview case. No difference was exhibited between the designated and emergency cases for the automobile. The shorter response time for the designated case in the simulator may be an indication of a premature steering response based on an improper estimation of vehicle position due to the lack of sufficient depth of field information from the visual display. It is not an indication of better performance since it was not exhibited in the automobile. Good relative similitude was also exhibited between the 52 metre and 32 metre cases for steering activity, ie T DELAY, T PEAK, T SPARE and T DAMP were all greater for the 52 metre case in both vehicles.

Designated vs Fixed and Emergency vs Fixed comparisons yielded similar results, as was to be expected since the designated and emergency cases differred only in the T RES parameter which as mentioned previously cannot be calculated for the fixed obstacle case. The maximum heading angle and the distance over the centre line was less for the fixed case than for the other two cases. This however was true only for the automobile. Thus for the differing degrees of task severity, the similator responses exhibit the same degree of clearance whereas in the automobile the least severe of the three manoeuvres is performed with a greater degree of confidence and an associated smaller safety margin.

It was noted that subjects' control strategy differred between the 52 and 32 metre cases. The 32 metre case required prompt steering responses. However the 52 metre case gave a sufficient time latitude for subjects to make decisions and change their control technique. Although this was not anticipated by the experimenters it is interesting to note that D OVER C, T OVER C and AY MAX were not different between the two manoeuvres, and the longer response time for the 52 metre case placed the vehicle close to the point where the pole would be triggered if it were a 32 metre preview case . This indicates that subjects changed their control strategy for the 52 metre case in such a way that it appeared like an avoidance manoeuvre for the 32 metre case.

Summary and Conclusions

In comparing driver obstacle avoidance performance in an automobile and a fixed base, computer generated image driving simulator the following was observed:

Transfer of Learning

1) positive transfer from simulator to automobile occurred for time to respond to initially falling obstructing pole and maximum steering peak angle.
2) negative transfer from automobile to simulator occurred for the number of steering wheel reversals.

Absolute Sense

1) during the obstacle avoidance manoeuvre similator trajectories made greater excursions than did the automobile's.
2) the number of steering wheel reversals exhibited during the obstacle avoidance manoeuvre was greater in the simulator for right manoeuvres only.
3) steering wheel activity in manoeuvring to avoid an obstacle was more sluggish in simulator than automobile for 32 m preview.

4) time required to respond to initial falling of obstructing pole was not significantly different between vehicles.

Relative Sense

1) time required to respond to initial falling of obstructing pole showed good relative similitude between 52 and 32 metre preview cases. It did not exhibit relative similitude between the designated and emergency obstruction cases.
2) steering wheel acitivity characteristic times (T DELAY, T PEAK, T SPARE , T DAMP) exhibited good relative similitude between 52 and 32 metre preview cases.

Manoeuvre Severity

1) during obstacle avoidance manoeuvre simulator trajectories exhibited no difference between fixed, designated or emergency cases. Automobile trajectories for the fixed case however , showed smaller excursions than did the designated or emergency cases. No difference for automobile trajectories was noted between disignated and emergency cases.
2) time to respond to initially falling pole was longer for emergency than designated case for simulator 32 metre preview case only.
3) subjects modified their control strategy between 52 and 32 metre preview cases in such a manner that the vehicle trajectory of the 52 metre case approximated that of the 32 metre case.

Conclusions

The simulator exhibited good absolute and relative similitude as well as postive transfer of learning for the T RES parameter. This however does not constitute simulator validity for the collision avoidance manoeuvre since the reponse time parameter is not specific to the task of lateral control of the vehicle. Except for the relative validity of the steering wheel activity time parameters and subjective opinion of the drivers no strong evidence supports the validity of the fixed base simulator in

representing an automobile for a lateral collision avoidance manoeuvre. This lack of validity does not however preclude the simulator's usefulness as a tool in studying extreme cases of lateral control. Driver modelling techinques applied to simulator behavior in an attempt to fine tune or correct them for lack of simulator fidelity may facilitate the prediction of automobile performance. This is the direction of investigation currently being undertaken.

REFERENCES

Maeda T., Irie N., Hidaka K. and Nishimura H., 1977, "Performance of Driver-Vehicle System in Emergency Avoidance"., SAE International Automotive Engineering Congress and Exposition, paper 770130.

McRuer D.T., Krendel E.S., 1973, " Mathematical Models of Human Pilot Behavior", AGARDOgraph M-188.

Reid L.D., Graf W.O., Billing A.M. 1980, "The Validation of a Linear Driver Model", University of Toronto Institute for Aerospace Studies, Report No.245.

Reid L.D., Graf W.O., Billing A.M., 1980, "A Preliminary Simulator Study of the Obstacle Avoidance Manoeuvre", The Ontario Ministry of Transportation and Communications, CVOS-TR-80-07.

EXPERIMENTAL PROCEDURE

TRANSFER of LEARNING - 2 groups

Day	1	2	3	4	
GROUP 1	SIM 52	AUTO 52	SIM 32	AUTO 32	(48 runs per subject)
GROUP 2	AUTO 52	SIM 52	AUTO 32	SIM 32	(48 runs per subject)

TRAINING - Alternating between automobile and simulator on consecutive days.

- 6 days
- 52 and 32m preview, right manoeuvre (24 runs per subject)
- fixed obstacle, no surprise (32 runs per subject)
- designated obstacle, subject was informed as to location of obstacle and degree of 'surprise' (16 runs per subject)

PRODUCTION RUNS - Alternating between automobile and simulator on consecutive days.

- 4 days
- 52, 32, fixed obstacle, and blank (no obstructing event) runs in random order and both directions (54 runs per subject).

TABLE 1

Performance Parameters

D OVER C - (distance over the centre line) maximum lateral distance that vehicle C of G goes past the roadway centre line.

T OVER C - (time over centre line) time that the vehicle C of G is in the adjacent lane.

AY MAX - (lateral acceleration maxium) absolute value of the maximum lateral acceration incurred during the manoeuvre.

PSI MAX - (heading angle maximum) absolute value of the maximum heading angle incurred during a manuoeuvre.

SW MAX - (steering wheel maximum) absolute value of the maximum heading angle incurred during the manoeuvre.

SW PEAK - (steering peak) absolute magnitute of the first steering wheel response after obstructing pole in triggered.

SW RVSL - (steering wheel reversals) number of times that the steering wheel rate changes sign. Counting starts at the time when the steering wheel hits the value of 10% of the initial response peak and finishes when the RMS value falls below the pre-emergency RMS value.

T RES - (time of response) time taken by the subject to respond to falling pole. Starting time is when pole is triggered, finishing time is where the tangeant line to the 10% and 90% points of the initial steering response crosses the time axis.

T RISE (rise time) time taken for initial steering response to go from 10% to 90% of initial peak.

T DELAY - (delayed time) time taken for initial steering response to reach 50% of initial peak from time of falling pole trigger.

T PEAK - (time to peak) time taken for initial steering response to reach its peak value from time of falling pole trigger.

T SPARE - (time to spare) time between first steering response to reach its peak value from time of falling pole trigger.

T DAMP - (time to damp) time from trigger time of falling pole until steering wheel activity fall below pre-emergency RMS value.

SW RATE - (steering wheel rate) slope of tangeant through 10% and 90% points of initial steering wheel response peak.

TABLE 2

AVOIDANCE MANOEUVRE PERFORMANCE PARAMETERS

Parameter	Units	52m Preview LEFT Mean	52m Preview LEFT S.D.	52m Preview RIGHT Mean	52m Preview RIGHT S.D.	32m Preview LEFT Mean	32m Preview LEFT S.D.	32m Preview RIGHT Mean	32m Preview RIGHT S.D.
AUTOMOBILE									
D OVER C	(METRES)	0.89	0.32	1.31	0.26	0.99	0.35	1.38	0.36
T OVER C	(SECS)	1.56	0.47	1.92	0.37	1.53	0.28	1.91	0.28
AY MAX	(M/S^2)	4.76	1.11	4.69	1.06	5.26	0.75	5.07	0.96
PSI MAX	(RADS)	0.17	0.03	0.16	0.02	0.20	0.02	0.19	0.03
SW MAX	(RADS)	2.05	0.70	2.42	0.60	2.57	0.57	2.58	0.48
SW PEAK	(RADS)	1.81	0.66	1.23	0.64	2.35	0.51	1.98	0.55
SW REVSL	(-)	4.21	1.44	3.30	0.67	3.90	1.48	3.33	0.76
T RES	(SECS)	1.49	0.57	1.57	0.56	0.58	0.08	0.61	0.09
T RISE	(SECS)	0.40	0.13	0.43	0.12	0.30	0.07	0.28	0.05
T DELAY	(SECS)	1.74	0.63	1.89	0.64	0.75	0.10	0.78	0.10
T PEAK	(SECS)	2.12	0.67	2.19	0.65	1.07	0.19	1.02	0.14
T SPARE	(SECS)	2.21	0.57	2.14	0.54	1.72	0.15	1.66	0.16
T DAMP	(SECS)	6.99	0.76	6.67	0.53	5.90	0.59	5.47	0.63
SW RATE	(RADS/S)	3.84	1.91	2.22	0.95	6.42	1.91	5.91	2.26
SIMULATOR									
D OVER C	(METRES)	1.53	0.25	1.33	0.37	1.69	0.50	1.46	0.41
T OVER C	(SECS)	1.93	0.30	1.76	0.32	2.00	0.50	1.80	0.34
AY MAX	(M/S^2)	4.45	1.07	4.54	0.87	4.81	0.81	4.79	0.73
PSI MAX	(RADS)	0.21	0.04	0.20	0.04	0.23	0.04	0.22	0.03
SW MAX	(RADS)	2.18	0.76	2.23	0.58	2.43	0.68	2.43	0.54
SW PEAK	(RADS)	1.73	0.48	1.68	0.59	1.88	0.57	1.83	0.43
SW REVSL	(-)	3.95	1.11	4.04	1.53	4.47	1.47	4.00	0.94
T RES	(SECS)	1.69	0.50	1.82	0.51	0.66	0.19	0.66	0.11
T RISE	(SECS)	0.46	0.10	0.46	0.09	0.43	0.13	0.40	0.13
T DELAY	(SECS)	2.01	0.52	2.12	0.50	0.93	0.21	0.91	0.18
T PEAK	(SECS)	2.35	0.52	2.49	0.46	1.29	0.26	1.24	0.26
T SPARE	(SECS)	2.12	0.50	1.98	0.50	1.70	0.20	1.69	0.10
T DAMP	(SECS)	6.86	0.77	7.00	1.10	5.88	1.08	5.52	0.79
SW RATE	(RADS/S)	3.17	1.22	3.07	1.44	3.87	1.90	4.00	1.49

TABLE 3

TRANSFER OF LEARNING

GROUP 1 (SIMULATOR FIRST VS GROUP 2 (AUTOMOBILE FIRST

PARAMETER	SIGNIFICANT DIFFERENCE (p<.05)	APPROXIMATE DIFFERENCE MAGNITUDE & PERCENT)	CONDITION FOR WHICH DIFFERENCE OCCURRED
-	NO SIGNIFICANT DIFFERENCES BETWEEN GROUPS	-	AUTOMOBILE/PRODUCTION
T RISE	GREATER FOR CAR FIRST GROUP	.1s/25%	SIMULATOR/PRODUCTION
T DELAY	GREATER FOR CAR FIRST GROUP	.6s/35%	SIMULATOR/PRODUCTION
T PEAK	GREATER FOR CAR FIRST GROUP	.6s/30%	SIMULATOR/PRODUCTION
	* * * * * * * * * * *		
SW PEAK	GREATER FOR CAR FIRST GROUP	11°/14%	AUTOMOBILE/TRANSFER
T RES	LONGER FOR CAR FIRST GROUP	.2s/20%	AUTOMOBILE/TRANSFER
SW REVSL	GREATER FOR CAR FIRST GROUP	1/20%	SIMULATOR/TRANSFER

TABLE 4

ABSOLUTE RESULTS

AUTOMOBILE VS SIMULATOR

PARAMETER	SIGNIFICANT DIFFERENCE (p<.05)	APPROXIMATE DIFFERENCE (MAGNITUDE & PERCENT)	CONDITION FOR WHICH DIFFERENCE OCCURRED
D OVER C	GREATER IN SIMULATOR	65cm/ 65%	32 L 52 L Fixed L,R
PSI MAX	GREATER IN SIMULATOR	2°/20%	32 L,R 52 L,R Fixed L,R
SW RVSL	GREATER NUMBER IN SIMULATOR	1/30%	32 R 52 R Fixed R
T RISE	LONGER IN SIMULATOR	.13s/50%	32 L,R
T DELAY	LONGER IN SIMULATOR	.15s/20%	32 L,R
T PEAK	LONGER IN SIMULATOR	.22s/20%	32 L,R

TABLE 5

RELATIVE RESULTS

PARAMETER	SIGNIFICANT DIFFERENCE (p<.05)	APPROXIMATE DIFFERENCE (MAGNITUDE & PERCENT)	CONDITION FOR WHICH DIFFERENCE OCCURRED
DESIGNATED VS FIXED			
D OVER C	GREATER FOR DESIGNATED	28cm/30%	AUTO L,R, 32, 52
PSI MAX	GREATER FOR DESIGNATED	2.5°/26%	AUTO L,R, 32
EMERGENCY VS FIXED			
D OVER C	GREATER FOR EMERGENCY	30cm/30%	AUTO L,R, 32, 52
PSI MAX	GREATER FOR EMERGENCY	2°/20%	AUTO L,R, 32
DESIGNATED VS EMERGENCY			
T RES	GREATER FOR EMERGENCY	.17s/35%	SIM L,R, 32
52m VS 32m PREVIEW			
PSI MAX	GREATER FOR 32	2°/20%	AUTO L,R
SW PEAK	GREATER FOR 32	35°/40%	AUTO L,R
T RES	GREATER FOR 52	1s/150%	AUTO L,R, SIM L,R
T DELAY	GREATER FOR 52	1s/100%	AUTO L,R, SIM L,R
T PEAK	GREATER FOR 52	1s/78%	AUTO L,R, SIM L,R
T SPARE	GREATER FOR 52	.4s/25%	AUTO L,R, SIM L,R
T DAMP	GREATER FOR 52	1.2s/25%	AUTO L,R, SIM L,R

TABLE 6

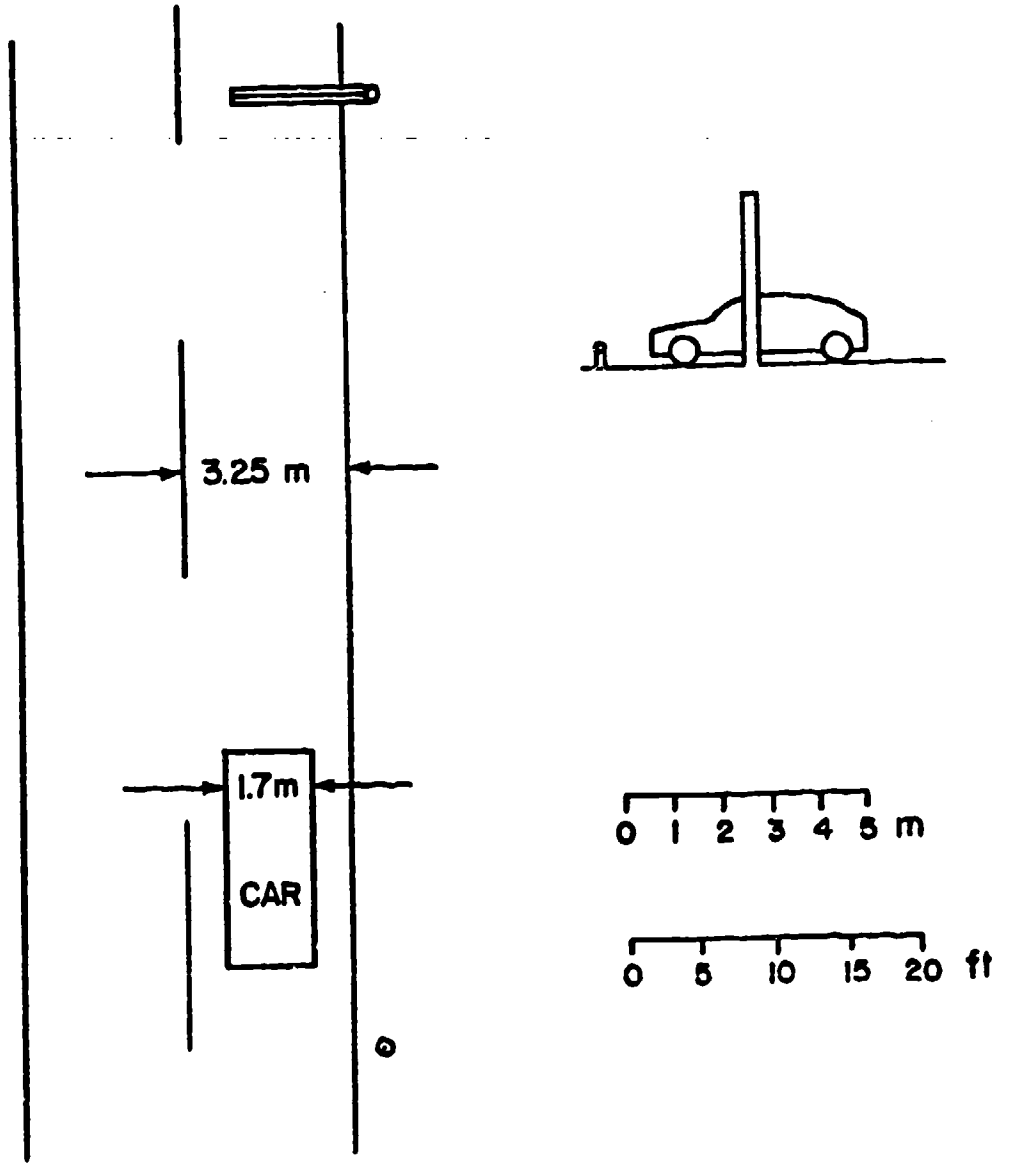


FIGURE 1 ROADWAY AND OBSTRUCTION GEOMETRY

FIGURE 2 TEST TRACK

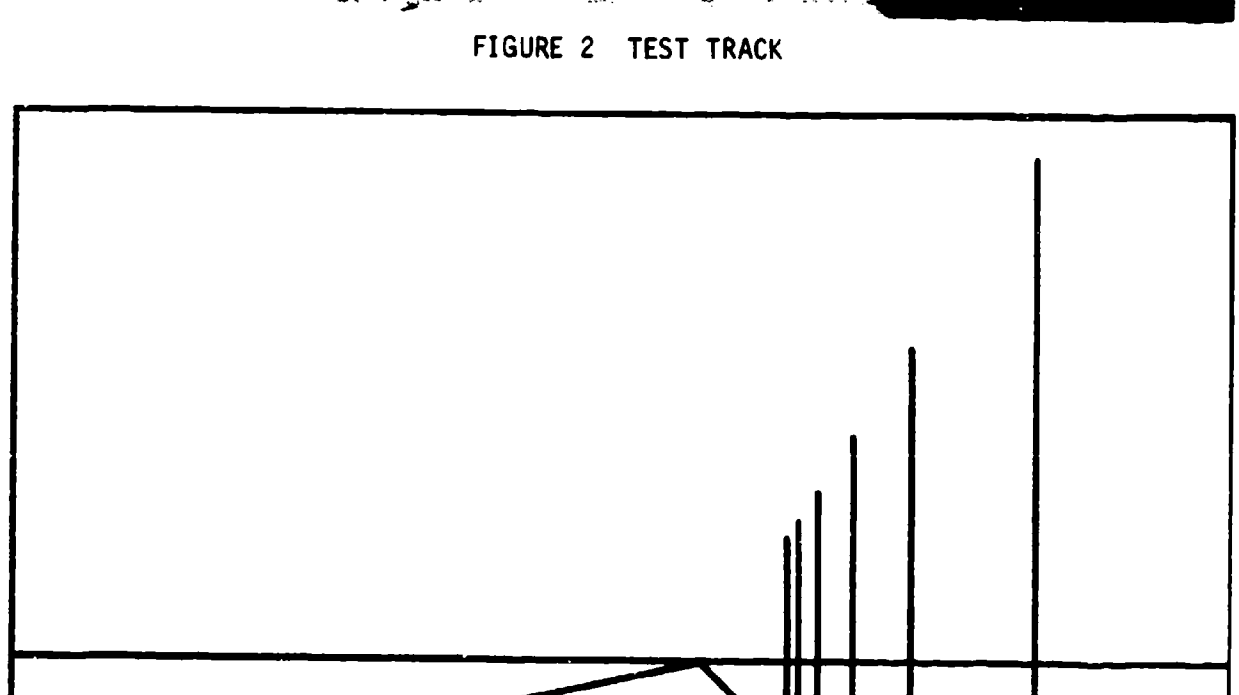


FIGURE 3 SIMULATOR DISPLAY

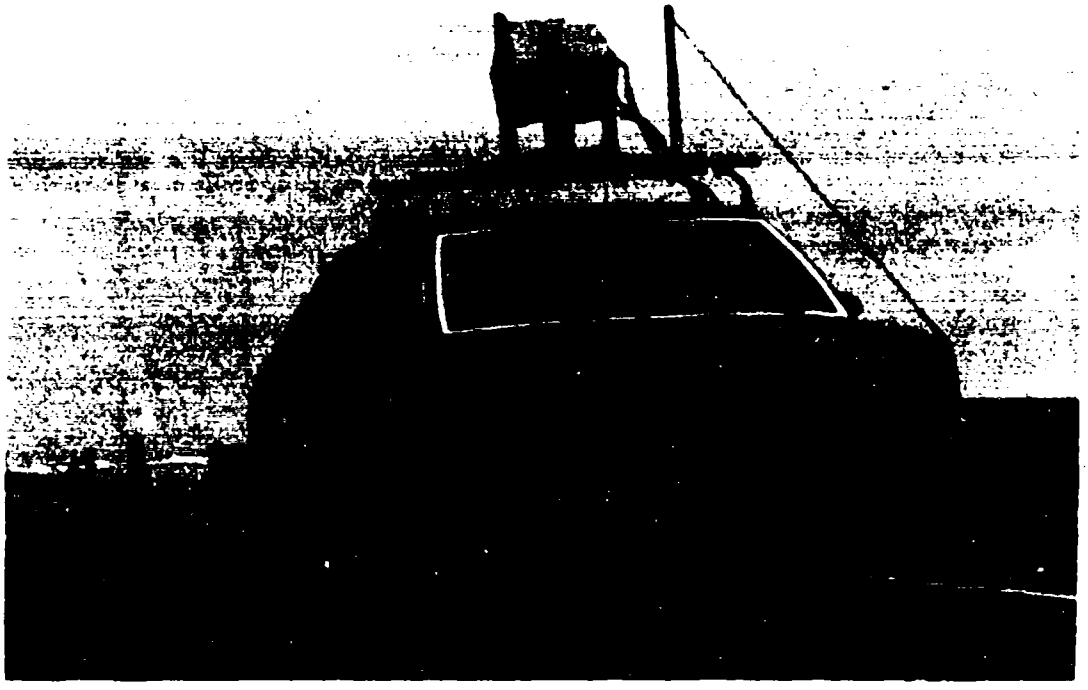
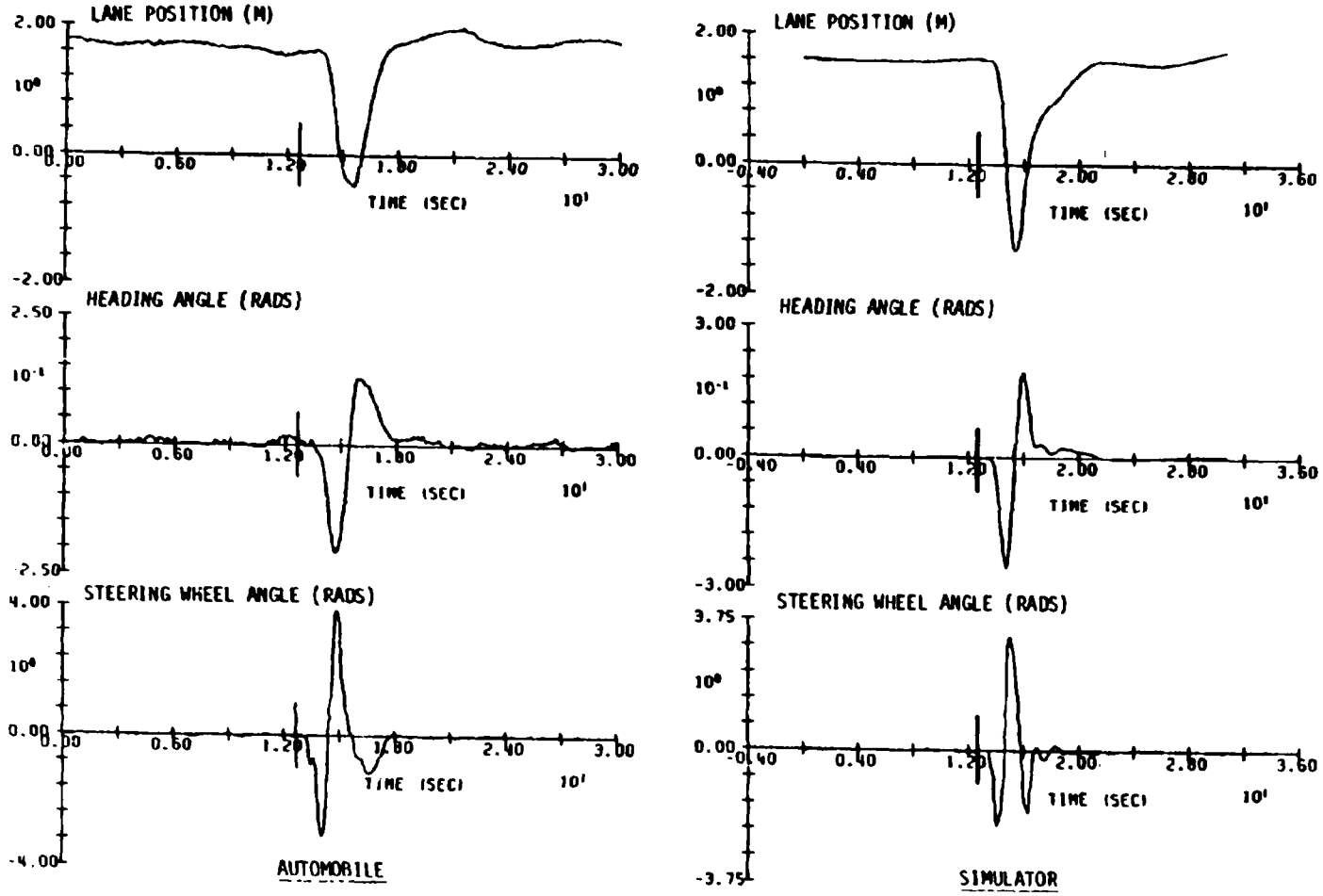


FIGURE 4 INSTRUMENTED VEHICLE DURING OBSTACLE AVOIDANCE MANOEUVRE

FIGURE 5 FIXED BASE SIMULATOR WORK STATION



TYPICAL TIME TRACES FOR 32 METRE PREVIEW CASE

FIGURE 6

AD P001626

AN OPTIMAL CONTROL MODEL ANALYSIS OF DATA FROM A SIMULATED HOVER TASK

by

Sheldon Baron
Bolt Beranek and Newman, Inc.
10 Moulton Street
Cambridge, MA 02238

Presented at the 18th Annual Conference on Manual Control
Dayton, OH
June 8-10, 1982

INTRODUCTION

The development of engineering requirements for man-in-the-loop simulation is a complex task involving numerous trade-offs between simulation fidelity and costs, accuracy and speed, etc. The principal issues confronting the developer of a simulation involve the design of the cue (motion and visual) environment so as to meet simulation objectives, and the design of the digital simulation model to fulfill the real-time requirements with adequate accuracy.

In specifying the cue environment the designer must establish the need for particular cues as well as the requisite fidelity for their presentation. The choices made here are important because the validity and utility of the resulting simulation can be critically dependent upon them and because the decisions involve major costs of the simulation. Unfortunately, these decisions are quite difficult to arrive at rationally, inasmuch as the choices depend on complex psychological as well as engineering factors. The requirements will be governed by the purpose of the simulation; training simulators have different needs than research simulators. They will also be problem dependent (e.g., the need for motion cues in the analysis of aircraft control in a gusty environment will depend on the gust response of the aircraft). Finally, the capabilities of the adaptive human controller both help and compound the problem. The human pilot may be able to compensate for simulator shortcomings and maintain system performance; however, this could result in negative transfer in a training environment or reduced acceptability of the device or an incorrect evaluation in a research simulation.

The design of the simulation model has become increasingly important and difficult as digital computers play a more central role in the simulations. The need for an adequate discrete simulation is also related closely to the cue generation problem inasmuch as the errors and, in particular, the delays introduced by the simulation will be present in the information cues utilized by the pilot.

"Past experience," open loop measurements, and subjective feedback from pilots are all helpful in developing the engineering requirements for simulators. However, for simulations in which the operator's principal task is flight control, analytic models for the pilot/vehicle (simulator) system can be very useful. Such models allow quantitative examination, in a closed-loop context, of the (inevitable) tradeoffs in simulator design, prior to commitment to prototype or full scale development. The design of elements or algorithms to compensate for simulator shortcomings can also be facilitated by modelling of this type. The models can serve as insightful ways of looking at and compressing empirical data so that it can be extrapolated to new situations. Finally, the parameters of an analytic model may prove to be sensitive measures of operator performance and adaptation.

The potential value of analytic models in the simulator design and evaluation process has been recognized in recent years and several development efforts have been undertaken. Of particular relevance to the work described here are several recent studies involving application of the Optimal Control Model (OCM) for pilot/vehicle analysis. Baron, Muralidharan and Kleinman [1] developed techniques for using the OCM to predict the effects on performance of certain simulation model design parameters, such as integration scheme, sample rate, data hold device, etc. The model was applied to a relatively simple air-to-air tracking task and showed significant sensitivity to several parameters. Model results were later compared with data from an experimental study of Ashworth, McKissick and Parrish [2] and the agreement was very encouraging (Baron and Muralidharan [3]).

In another study [4], the OCM was used to examine the closed-loop consequences in a helicopter hover task of the performance limitations associated with a computer generated image visual system and a six-degree of freedom motion system. The hover task was linearized and decoupled into separate longitudinal and lateral control tasks. Performance/workload effects of these simulation elements were analyzed by incorporating elaborated sensory perception sub-models in the OCM. The model results suggested that simulator deficiencies of a reasonable nature (by

current standards) could result in substantial performance and/or workload infidelity with respect to the task in flight. Unfortunately, there were no corresponding experimental data to confirm or deny these predictions.

Finally, we mention the brief study of Baron [5] to integrate the earlier efforts into a Multi-Cue OCM and to apply the resulting model to analyze effects of control loader dynamics and a g-seat cues on the air-to-air tracking problem investigated earlier [1]. The results of this study have been partially validated empirically but further definition of the proprioceptive model appears necessary. The Multi-Cue model described by Baron [5] provides the analytic basis for the current investigation.

This paper describes the application of the Multi-Cue OCM to the analysis of data obtained in an experimental study of simulated helicopter hover [6]. The goal was to demonstrate that the model could be used to analyze or predict the effects of simulator changes in a complex flight control task. Thus, to a degree, the investigation complements the earlier studies which were either completely analytic or involved data from relatively simple control tasks. It is worth mentioning that this study is of broader interest than the simulation context because of the opportunity to compare model results with data in what appears to be, in certain ways, the most complex steady-state control task modelled by the OCM.

Section 2 of the paper describes the particulars of applying the Multi-Cue OCM to the specific task being investigated. Model predictions are then compared with experimental data in Section 3. The last section summarizes the results and presents concluding remarks.

APPLICATION OF OCM TO SIMULATED HOVER TASK

In this section, the task considered in the experimental simulation study is summarized briefly. (A more detailed review of this study is beyond the scope of this report but may be found in Ricard, et al [6]). Then, the representation and specification of the task and pilot in the OCM framework are described.

GENERAL TASK DESCRIPTION

The experiment was designed to examine the effects on hover control of motion cues (as provided by a motion platform or g-seat) and of delays in the generation of visual cues. The simulated

flight task corresponded to maintaining a Huey Cobra helicopter in a high hover relative to a ship moving at 15 knots. The desired hover position was fifty feet from the ship in line with a diagonal deck marking and at a mean height of twenty feet above the deck. The helicopter was subjected to turbulence modelled by Dryden gust spectra appropriate to an altitude of forty feet. Motions of the deck of the ship could also be simulated and the presence or absence of such ship motion was an experimental variable.

The data collected were rms values or tracking errors (rail-rail, bow-stern and height), helicopter pitch and roll and positions of controls (collective, pitch and roll cyclic, and rudder pedal).

MODEL SPECIFICATION

To analyze the task with the OCM, we must specify system dynamics, disturbances, displays (including cues), a performance objective and the assumed pilot limitations. Each of these is discussed below.

System Dynamics and Disturbances

A linearized set of equations of motion for the Huey Cobra, in state-variable format (Equation 2.1), was provided by Sperry Support Services, NASA-LRC [7]. These equations were obtained from the full nonlinear system simulation by means of a special program designed to develop such linear models [8].

Because of the 15 knot sideslip, decoupling of lateral and longitudinal helicopter motions did not appear to be justified (this was later verified). In addition, it did not seem appropriate to ignore the helicopter rotor dynamics. Thus, the basic equations of motion involved seventeen state variables: three positions (relative to the deck) and three linear velocities (in body axes), three body angular rates and attitude angles; and five states describing the rotor dynamics. Twelve additional state variables were required to model the stability augmentation system (SAS) of the helicopter. The state equations used in this analysis are given in Baron [9].

Dryden gust spectra were used to model the turbulence environment. In particular, the vehicle body axis gust disturbances u_g, v_g and w_g, were generated by passing white noise through filters so as to yield gust spectra appropriate to the forty foot altitude and fifteen knot airspeed of the helicopter. The rms intensity levels of the gusts u_g, v_g and w_g were 5 ft/sec, 5

ft/sec and 1.42 ft/sec, respectively. The state-variable models for the gusts are given in Baron [9]; five state variables are needed to model the three gusts.

In the experimental study, ship motion in each of the pitch, roll and heave axes was modelled by the sum of four sine waves. These inputs drove g-seat bladders upon which the ship model was mounted. It is theoretically possible to duplicate these inputs exactly but this would require twelve more state variables. Furthermore, it would be necessary to conduct a very careful analysis of the amplitudes of the motion as seen through the visual display system in order to model the effects of this ship motion precisely. For this analysis a much simpler approach was used. Only heave of the deck was considered and this was modelled by passing white noise through a second order filter; the parameters of this filter were selected so that the autocorrelation of this input matched that of the heave motion input used in the simulation. The rms level of the input used for this study was set at 3.5 feet, corresponding to the five foot *maximum* heave of the landing pad reported by Ricard [6].

In the OCM analysis, ship movement was treated as a "display" disturbance; i.e., it was added to the pilot's perception of vertical error relative to the landing pad. The pilot's assumed objective was to maintain the helicopter's position relative to the mean position of the deck.

When the coupled vehicle dynamics, the rotor dynamics, the SAS and the gust and ship motion disturbances are all included, there are 36 state variables needed to model the "system" and environment (34 states for the case of no ship motion). Moreover, this does not even include any state variables to account for simulator dynamics or human sensory dynamics. This is, to our knowledge, a significantly larger problem than any previously treated with the OCM and, indeed, required modification of existing analysis programs to accommodate.

Simulator and Perceptual Models

Detailed descriptions of the motion and visual systems employed in the helicopter simulation are described in Ricard, et al [6]. Because of the complexity of the basic problem, as noted above, these systems were approximated in the simplest fashion possible. In particular, it was assumed that any dynamics associated with motion cueing or sensing could be neglected. It was also assumed that differential delays between motion and visual cues could be neglected. These two assumptions were made to avoid

the necessity of having to introduce additional state variables. The assumption concerning the dynamics associated with motion cueing and sensing has proven to be reasonable in several previous studies. The assumption concerning the differential delays is less tenable, particularly when additional delay is added to the visual channel. Nonetheless, as we will see later, these assumptions appear quite adequate with respect to capturing much of what is in the data.

Thus, we treat the cue generation and perception problem on an informational level; i.e., we determine what control related information is available from the various cues and the corresponding perceptual limitations (mainly delay and thresholds).

The total delays for the various cues reported by Ricard, et al. [6] (including sampling and computational delays) are 77 ms for the VMS motion platform, 82 ms for the g-seat and 62-69 ms for the visual display.* We assumed an "average" delay of 70 ms for all cues.

The visual display provided a view of the ship, the horizon and a dimly lit Head-Up Display (HUD) to provide position reference information, otherwise denied by a limited field of view (48° by 26°). The TV display had a resolution of 9 minutes of arc (.15°). It was assumed that from this external view display, the pilot could obtain both position information (relative to the landing pad) and attitude information. As is normally the case in applying the OCM, it was also assumed that rates of change of these variables could also be perceived.

Visual thresholds for the various variables were computed using the method described in Baron, Lancraft and Zacharias [4]. These calculations are based on an analysis of the potential cues and depend on the geometry of the situation and on the discrimination and/or resolution capabilities of the human controller and of his equipment. In this case, the resolution limit of the display (.15°) is greater than that determined for human trackers in previous studies (.05°), so that it was used for resolution-determined thresholds. The geometric factors of importance are field of view and distance from and above the object being used to obtain the cue. The values of the visual thresholds determined from this analysis are given in Table 1.

* The visual delays were different for translation (62 ms) and rotation (69 ms).

Table 1. PERCEPTUAL THRESHOLDS*

Variable	Visual	Motion Platform	G-Seat
x, ft	1.68	--	--
$\dot{x}$, ft/sec	.72	--	--
y, ft	.07	--	--
$\dot{y}$, ft/sec	.30	--	--
z, ft	.74	--	--
$\dot{z}$, ft/sec	.34	--	--
a_x, ft/sec^2	--	.053	.0636
a_y, ft/sec^2	--	.053	--
a_z, ft/sec^2	--	.053	--
ϕ, deg	.15	--	--
$\dot{\phi}$(p), deg/sec	.6	2.5	3.0
$\dot{p}$, deg/sec^2	--	.41	.49
θ, deg	.15	--	--
$\dot{\theta}$(q), deg/sec	.6	3.6	4.3
$\dot{q}$, deg/sec^2	--	.67	.80
ψ, deg	.15	--	--
$\dot{\psi}$(r), deg/sec	.6	4.2	--
$\dot{r}$, deg/sec^2	--	.41	--

*An entry of -- means that it is assumed that no information on the variable is provided by the modality

The motion platform was assumed to provide the pilot with linear accelerations (a_x, a_y, a_z) and body angular rates and accelerations (p, q, r and $\dot{p}, \dot{q}, \dot{r}$). Thresholds for perception of these variables were based on the values given in Baron, Lancraft and Zacharias [6] and are listed in Table 1.

The g-seat was assumed to provide the pilot with surge acceleration and with pitch and roll rates and accelerations. Perceptual thresholds for this type of cueing have not been measured, but it was felt that the g-seat did not provide cues as faithfully as the motion platform. Therefore, thresholds for g-seat cues were arbitrarily set to be 20 per cent higher than for the corresponding thresholds for platform motion indicated in Table 1.

Performance Criteria

The application of the OCM requires specification of a quadratic cost function. This involves specifying the variables to be considered and the weightings associated with those variables. The weightings can be associated with allowable deviations or with human subjective criteria or limitations. Here, because there was no opportunity to discuss performance objectives with the pilots, we made somewhat arbitrary initial assumptions concerning allowable deviations and then adjusted them via a brief preliminary sensitivity analysis to obtain reasonable agreement with the data for the fixed base, no added delay, no ship motion condition. The weights were then fixed for the rest of the analysis. The output and control variables included in the cost functional along with the corresponding deviations and weightings resulting from the preliminary analysis are listed in Table 2. Though not shown in Table 2, the cost functional also included control rate terms whose weightings were chosen to yield a "neuromotor" response lag of T_N = .1 second in each control channel.

Other Model Parameters

The remaining model parameters to be specified are the operator's time delay and observation and motor noise/signal ratios. We assumed a basic pilot delay of .2 seconds and a motor noise/signal ratio of -25 dB; these values are typical of those obtained in basic tracking experiments and have been used in numerous other applications.

To specify the observation noise, we must choose a base value for noise/signal ratio, P_i, and an attention-sharing algorithm. We assume a value of P_i of .04 (i.e., -14 dB) for all information variables, regardless of source. This value is higher than that associated with simple laboratory tasks in which operators receive extensive feedback and motivation and train to asymptotic performance. We believe the higher noise ratio is probably more appropriate for the experimental conditions and problem complexity being analyzed here.

We assume no attention-sharing between modalities or <u>within</u> the motion and proprioceptive modalities. Within the visual modality, we assume that the pilot shares attention equally among the six basic informational variables (position errors and vehicle attitudes).*

* As is normally the case in applying the OCM, it is assumed that no diversion of attention is required to obtain the rate of change of a quantity from the same indication or cue as the quantity itself.

Table 2. PERFORMANCE COST FUNCTION PARAMETERS

Variable	"Maximum" Allowable Deviation (M_i)	Cost Weighting $\frac{1}{M_i^2}$
Rail-Rail Error, x, ft	10	.01
Bow-Stern Error, y, ft.	10	.01
Heave Error, z, ft.	10	.01
Roll, ϕ, rad.(deg.)	.1(5.7)	100
Pitch, θ, rad.(deg.)	.1(5.7)	100
Yaw, ψ, rad.(deg.)	.0175(1)	3280
Collective XAOS, in.	1	1
Roll Cyclic, YCS, in.	1	1
Pitch Cyclic, XCS, in.	1	1
Rudder, XTR, in.	.1	100

MODEL-DATA COMPARISONS

Twelve experimental conditions were investigated for the main experimental variables of motion cueing, visual delay and ship movement. These conditions are enumerated in Table 3. Each condition was repeated five times by each of 12 pilots, so that

Table 3. EXPERIMENTAL CONDITIONS

CONDITION NO.	SHIP MOTION	VISUAL DELAY	MOTION BASE	G-SEAT
1	OFF	OFF	OFF	OFF
2	"	"	ON	"
3	"	"	OFF	ON
4	"	ON	OFF	OFF
5	"	"	ON	"
6	"	"	OFF	ON
7	ON	OFF	OFF	OFF
8	"	"	ON	"
9	"	"	OFF	ON
10	"	ON	OFF	OFF
11	"	"	ON	"
12	"	"	OFF	ON

there were 60 runs per condition. The means and standard deviations for position errors, pitch and roll angles and control inputs obtained from these 60 runs, by experimental condition, are

given in Table 4. Also shown in the last column of Table 4 is the variability in the various measures that is attributable to pilot difference alone (expressed as a standard error and computed from the ANOVA data in appendix B of Ricard, et al [6]. Unfortunately, this measure cannot be separated by experimental conditions on the basis of the data presently available.

The reader is referred to Ricard, et al [6] for a detailed analysis of the data using univariate analysis of variance. The discussion here will be very brief. Significant effects were found on at least some measures for all the experimental variables, including the replicate and pilot variables. Indeed, the pilot factor was the major source of variance among the experimental variables. Motion cueing and visual delay had significant effects on the error measures but ship movement did not. Visual delay tended to increase the values of all system variables (error, states and controls).* Platform motion had a significant effect on vehicle roll attitude but not on pitch. Cyclic pitch increased with platform motion while cyclic roll decreased. Most pilots used more control when the ship motion was active. In addition to the main effects, there were a number of significant interactions, several of them involving the pilot factor. For example, the effect of visual delay was more pronounced when there was no ship motion.

The OCM as specified above was used to predict the rms values for all the variables of interest in each of the twelve conditions. The results are given in Table 5. These results were obtained with a fixed set of pilot parameters--the only changes made in the model from condition-to-condition were those corresponding to changes in the simulation (e.g., presence or absence of platform motion, g-seat or ship motion). It should also be noted that when used in this fashion, the OCM is intended to be representative of the performance expected of a highly trained, highly motivated pilot who has been allowed to reach asymptotic performance on each condition before data are taken.

Comparison of Tables 4 and 5 shows that all measures predicted by the model are within one standard error of the means, where the standard error is the measure of pilot variability mentioned above. This implies that the model is as good a predictor of a pilot's performance across the conditions as another pilot selected randomly from the group. Given the complexity of the problem and the assumptions made, this is a substantial achievement.

---

* Except for two pilots whose performance, surprisingly, improved with delay.

Table 4. EXPERIMENTAL RESULTS (Mean ± S.D.)

CONDITION NO. 1	1	2	3	4	5	6	7	8	9	10	11	12	Pilot Variability (S.E)[2]
Rail-Rail Error,x,ft	9.5±5.8	7.3±3.0	9.3±4.9	12.2±6.6	8.9±3.8	10.2±4.3	10.9±4.8	7.9±3.3	10.3±5.0	11.0±5.3	8.7±3.4	10.1±3.8	5.41
Bow-Stern Error,y,ft	9.8±5.3	8.3±3.7	9.5±4.6	12.4±5.6	10.1±4.2	10.9±5.1	12.0±5.7	8.4±3.5	10.7±5.1	10.9±4.4	9.8±3.8	10.8±4.1	5.52
Altitude Error,h,ft	4.3±1.9	3.9±1.7	4.1±1.5	5.1±1.9	4.2±1.6	4.6±1.3	4.3±1.4	3.8±1.2	5.0±2.2	5.0±1.9	4.4±1.9	4.9±1.7	2.05
Roll Angle ϕ, rad.	.073±.023	.066±.012	.071±.016	.080±.019	.072±.019	.080±.018	.070±.013	.071±.016	.073±.014	.081±.022	.072±.017	.081±.020	.021
Pitch Angle,θ, rad	.083±.013	.083±.012	.082±.010	.085±.014	.081±.009	.083±.009	.082±.010	.083±.012	.084±.012	.086±.012	.086±.011	.086±.013	.015
Collective XAOS,in.	.42±.21	.46±.25	.36±.14	.54±.36	.53±.35	.46±.20	.45±.19	.50±.28	.48±.24	.50±.38	.48±.28	.53±.29	.343
Roll Cyclic YCS,in.	.41±.13	.33±.12	.41±.16	.41±.13	.36±.14	.43±.14	.38±.12	.37±.15	.40±.13	.45±.15	.36±.12	.43±.17	.246
Pitch Cyclic XCS,in.	.44±.16	.45±.13	.48±.24	.46±.14	.47±.16	.50±.14	.41±.14	.49±.17	.47±.17	.51±.17	.49±.13	.50±.21	.276
Rudder, ,in.	063±.019	.068±.019	.062±.016	.073±.022	.071±.018	.070±.016	.069±.017	.067±.020	.070±.022	.069±.024	.075±.021	.073±.018	.020

1. Condition numbers refer to Table

2. Standard Error attributable to pilot variability alone, independent of condition.

Table 5. MODEL PREDICTIONS FOR DIFFERENT EXPERIMENTAL CONFIGURATIONS

Condition No.[1]		1	2	3	4	5	6	7	8	9	10	11	12
Rail-Rail	, ft	9.5	8.2	9.2	9.9	8.6	9.7	10.0	8.5	9.3	10.6	8.9	10.0
Bow-Stern	, ft	7.0	5.8	6.9	7.4	6.1	7.2	7.5	6.0	7.0	7.9	6.3	7.5
Altitude	, ft	3.1	2.1	2.8	3.2	2.2	2.9	6.5	5.2	5.8	6.5	5.2	5.9
Roll	, rad	.069	.062	.067	.070	.063	.069	.072	.063	.069	.074	.064	.071
Pitch	, rad	.073	.070	.073	.074	.071	.074	.074	.071	.073	.075	.072	.074
Collective	in	.27	.23	.26	.27	.23	.26	.30	.24	.27	.30	.24	.28
Roll Cyclic	, in	.29	.24	.28	.30	.25	.29	.31	.25	.29	.32	.26	.30
Pitch Cyclic,	in	.29	.24	.29	.31	.26	.30	.31	.26	.29	.33	.27	.31
Rudder	, in	.057	.056	.057	.056	.056	.057	.058	.056	.057	.058	.056	.057

[1] See Table 3.

The model predictions are also within one standard deviation of the mean (where the variation is across subjects and replications) for virtually all measures and all conditions. The exceptions to this are, in a few instances, altitude error, cyclic pitch and pitch angle. These exceptions are well within two standard deviations of the mean and most are only marginally more deviant than one standard deviation.

Almost all model predictions are somewhat lower than the corresponding mean values determined experimentally. This suggests that the predictions could be brought in even closer agreement with the measured means by slight increases in the delay or noise parameters of the model.

If we examine the model results as a function of the main experimental variables, additional interesting features emerge.

Effects of Motion Cues. Figure 1 shows the effect of motion cues on rail-to-rail error, averaged over ship motion conditions, for the two delay conditions. It can be seen that the g-seat cues provide some improvement in performance and platform motion cues result in still less error. Further, the effect of motion cues does not appear to be significantly different for the two delay values; i.e., there does not appear to be a delay-motion interaction. These effects of motion cues exactly parallel the trends contained in the data (Figure 3 of Ricard et al [6]). Reference to Table 5 indicates that the remaining error terms show the same trends.

The effect of motion cues on the attitude variables (ϕ and Θ) are shown in Figure 2. For this figure, model results for each motion condition were averaged across the delay and ship motion variables. The g-cues have only a very slight effect on the attitude variables. Platform motion results in demonstrably lower rms attitude values, with the effect for roll being greater than that for pitch. Again, these effects closely parallel those in the data (See Ricard, et al [6] Figure 15) where there was a significant effect of motion on roll, but not on the pitch rms score.

The model does less well in predicting the effects of motion cues on control inputs, particularly pitch cyclic. As can be seen from Table 5, rms control inputs show very little differences between fixed base and g-seat condition, with, if anything, somewhat less activity with the g-seat on. However, control inputs are significantly reduced in the moving base condition. These trends reproduce those for the roll cyclic data. The collective

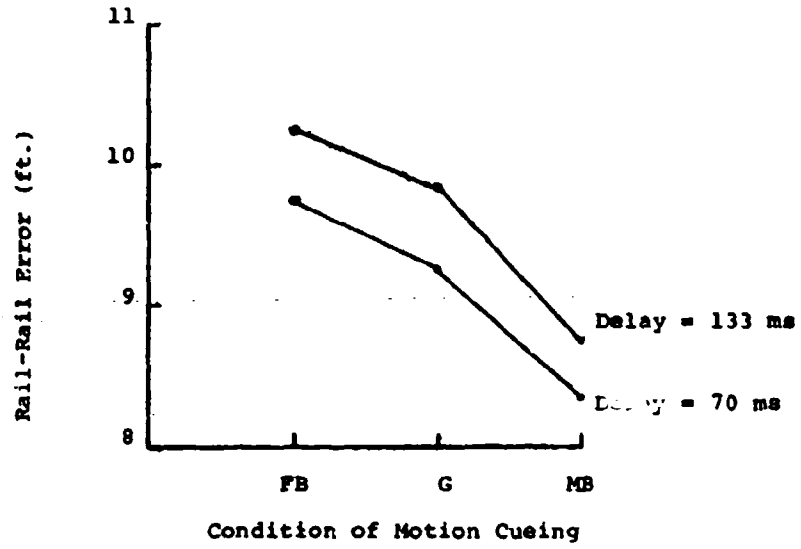


Figure 1. Effects of Motion Cues on Rail-Rail Hover Error

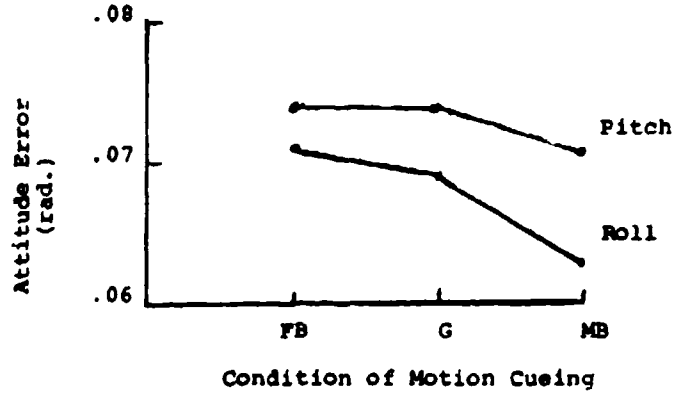


Figure 2. Effects of Motion Cues on Attitude Errors

and rudder data show virtually no difference between the fixed base and g-seat conditions, as is true for the model; however, while there appears to be no consistent effect of platform motion, Ricard reports that most pilots use significantly more collective and rudder for this condition, contrary to the predictions of the model. Finally, the model's trend for pitch cyclic does not agree at all with the observed data. In the experiment, pitch cyclic activity increases when motion cues are added and is greatest for the g-seat condition.

Visual Delay. The OCM predicts that all measurements, errors, states and controls will increase with the addition of delay. This pattern is also evident in the data. However, the increases predicted by the model tend to underestimate those observed in the data; i.e., the OCM is doing a better job of compensating for the delay than are the pilots. Also, while the model accurately predicts no motion-delay interaction, Figure 3 shows that the model does not predict the delay-ship motion interaction that was observed (Ricard et al [6], Figure 12).

Ship Movement. The model predicts an increase in errors, attitude variables and control inputs when ship movement is present. The model results are not surprising from an analytic standpoint, but they are not entirely borne out by the data. For example, heave error increases with ship movement predicted by the model are much greater than those observed empirically. In addition, the model predicts increases in error regardless of the delay condition whereas the data show decreased errors with ship movement for the longer visual delay. Model trends are confirmed for the attitude variables, with ship movement increasing rms pitch and roll (for most pilots). Although the mean rms values of control variables do not exhibit consistent trends (Table 4), Ricard et al [6] does report a higher rudder activity, that is significant, with ship motion and that other control activity measures were also higher for most pilots; these results are in keeping with the model's predictions.

CONCLUSION

This paper has described the results of a brief effort to apply the OCM to analyze data obtained in an experiment aimed at exploring simulator fidelity factors. The experimental task was to maintain a high hover position relative to a moving ship. Twelve skilled pilots performed the task for each of twelve experimental conditions. The main experimental variables were motion cue presentation, visual delay and ship movement.

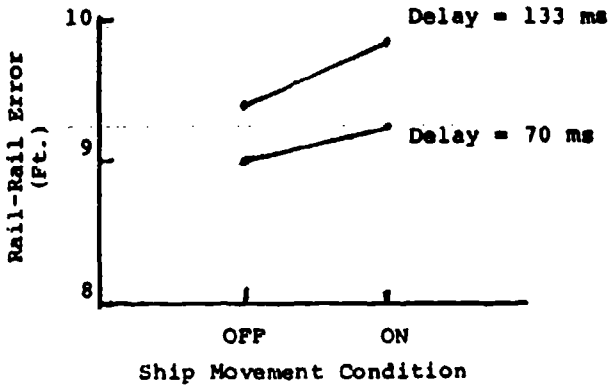


Figure 3. Effect of Ship Movement on Predicted Rail-Rail Hover Error

The simulated hover task was specified in terms needed to apply the OCM. The vehicle dynamics were of substantial complexity, so simplifying assumptions with respect to simulator and perceptual characteristics were made for practical reasons. The parameters of the model relating to human limitations were selected mostly on the basis of previous work, though cost functional weightings were adjusted to give a reasonable "match" to the base experimental condition.

In general, the model predictions agree with the data to a substantial degree, with most of the differences being within the range of experimental variability. There appears to be a general trend for the model to be less sensitive to configuration changes than the mean measures. Unfortunately, while the data treatment

allows us to determine the significance of a difference in performance resulting from a configuration change, it does not tell us the magnitude of the inter-subject variability in this difference (a paired-difference treatment of the data would provide this information). Thus, we cannot determine whether or not the differences predicted by the model are within the pilot-to-pilot variation. Nonetheless, there are two assumptions that could readily explain a lesser degree of model sensitivity to configuration change: 1) that performance has asymptoted in each condition so that the "pilot" has fully adapted to the configuration; and 2) that the basic pilot limitations (i.e., the OCM model parameters) do not change significantly throughout the experiment. It is highly suspect that these assumptions can ever be strictly valid in a factorial experiment of this size, given practical constraints. On the other hand, for this very reason, the model results may be a better measure of the differences due to configuration changes alone than are the data.

The excellent agreement between model and data for the fixed base, no additional delay case suggests that the information the pilot was obtaining from the external scene was adequately accounted for. Although we did not show the results, we also obtained model results under the assumption of no visual thresholds; these results failed to match the data. Thus, it would appear that the method used here for estimating visual thresholds from analysis of the external scene, proposed by Baron, Lancraft and Zacharias [4] is reasonable.

The model does particularly well in predicting the effects of motion cues on position errors and attitude angles. This suggests that the simple informational treatment of motion cues employed here was adequate for this task. Furthermore, both the sensory thresholds assumed for these cues and the assumption of no attention-sharing within the motion-sensing modalities appear reasonable in light of the model-data agreement. A major caveat to these conclusions is that the model did fail to predict the effect of motion on pitch cyclic control, a fact that we are unable to explain based on our limited analysis. It should also be noted that one is unlikely to be able to separate definitively the various sources of pilot observation noise (thresholds, base noise/signal ratio and attention sharing logic) without designing experiments expressly for this purpose (see, e.g., Levison [10]).

The main effects of visual delay were also predicted by the model as was the lack of motion-delay interaction. The model predicts less degradation with delay than is observed in the "mean" rms performance measures. It is not clear that the lower

sensitivity of the OCM to increases of delay is attributable entirely to the assumptions mentioned above. There is the possibility that the reduced stability margins in the added delay case cause the pilots to adapt their strategies in response to an increased tendency for Pilot Induced Oscillation (PIO) and the model does not account for this change. Further analysis would be needed to resolve this question.

The model was less successful in predicting the effects of ship movement, in particular in over-estimating the effects on heave error and in missing the ship movement/delay interaction. On the basis of the present analysis, we cannot tell whether this is simply the result of our approximation to the ship movement input or it is due to some other inadequacy of the model.

In sum, these results demonstrate a significant capability for predicting closed-loop system performance, and the effects on it of simulator deficiencies, for even highly complex tasks. When viewed in light of other recent studies of a similar nature, the results clearly point to the conclusion that the OCM can be very useful for predicting and analyzing engineering requirements for simulators. This includes the potential for design and evaluation of algorithms (such as washout filters or delay compensators) to ameliorate hardware deficiencies.

Several areas for further research are suggested by this effort. A major source of variance in the simulator study was the pilot factor; there were significant differences between pilots, both quantitatively and qualitatively. The development and application of the OCM has stressed modelling the performance of an "average" well-trained, well-motivated pilot. This is highly appropriate strategy for many problems of system design and evaluation. However, there are obviously individual differences and these, too, are important, particularly if one attempts to address training and/or "screening" issues. It is of interest, therefore, to determine the extent to which variations of parameters of the OCM can account for individual differences among skilled pilots or among pilots in various stages of training.

Although the model did quite well in predicting the effects of motion cues, there are still several issues to be addressed. One specific issue is the disparity between modelling results and data in the control activity seen in the mving base condition. More generally, we need a basic and systematic empirical study of g-seats to get at the lmits of their utility. Interpreting such studies using the OCM, especially in terms of perceptual limitations, would be very useful.

Finally, with respect to any of the simulation devices or limits examined here, it is important to investigate their effects on skill acquisition (i.e., rate of learning) and on transfer of training, as well as on performance. It may well be that a simulator enhancement does not improve asymptotic performance but does reduce training time or, conversely, that it improves simulator performance but not training. We believe that the OCM could provide very useful insights and measurements of the progress of training, but it requires research to extend it so that it will be applicable to the problem of skill acquisition.

ACKNOWLEDGEMENTS

The study reported herein was performed under contract N61339-80-C-0055 for the Naval Training Equipment Center. Dr. Gil Ricard of NTEC was the technical monitor. The experimental study was conducted jointly by NTEC and NASA-Langley Research Center at the NASA facility. Messrs. Russell Parrish, Roland Bowles and Burnell McKissick, of NASA, gave essential help by providing an appropriate vehicle model and the empirical data needed for the study.

REFERENCES

[1] Baron, S., Muralidharan, R., and Kleinman, D.L., "Closed-Loop Models for Analyzing Engineering Requirements for Simulators," NASA CR-2965, Feb. 1980.

[2] Ashworth, B.R., McKissick, B.T., and Parrish, R.V., "The Effects of Simulation Fidelity on Air-to-Air Tracking," Proc. of Fifteenth Annual Conference on Manual Control, Wright State University, Dayton, OH, March 1979.

[3] Baron, S. and Muralidharan, R., "The Analysis of Flight Control Simulator Characteristics Using a Closed-Loop Pilot Vehicle Model," BBN Rept. No. 4329, Feb. 1980.

[4] Baron, S., Lancraft, R., and Zacharias, G., "Pilot/Vehicle Model Analysis of Visual and Motion Cue Requirements in Flight Simulation," NASA CR-3212, Oct. 1980.

[5] Baron, S. A Multi-Cue OCM for Analysis of Simulator Configurations," BBN Report No. 4373, April 1980.

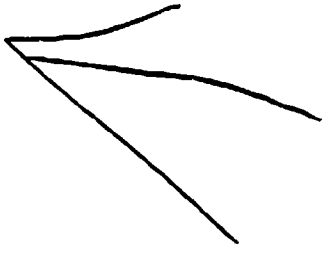
[6] Ricard, G.L., R.V. Parrish, B.R. Ashworth and M.D. Wells, "The Effects of Various Fidelity Factors on Simulated Helicopter Hover," NAVTRAEQUIPCEN 1H-321, December 1980.

[7] Carzoo, S. and Leath, B., "A Linearized Helicopter Model," Sperry Support Services, Report No. SP-710,018, Model 1980.

[8] Dieudonne, J.E., "Description of a Computer Program and Numerical Techniques for Developing Linear Perturbation Models from Nonlinear Simulations," NASA TM 78710, July 1978.

[9] Baron, S., "An Optimal Control Model Analysis of Data from a Simulated Hover Task", Report No. 4522, Bolt Beranek and Newman, Inc., Cambridge, MA, November 1980.

[10] Levison, W.H., "The Effects of Display Gain and Signal Bandwidth on Human Controller Remnant," AMRL-TR-70-93, March 1971.



AD P001627

ANALYSIS OF A AAA SUPERVISORY CONTROL TASK USING SAINT

by
Kuang C. Wei, Blaza Toman, Joseph Whitmeyer
Systems Research Laboratories, Inc.
Dayton, Ohio 45440

Sharon L. Ward
Air Force Aerospace Medical Research Laboratory
Wright Patterson AFB, Ohio 45433

(Systems Analysis of Integrated Networks of Tasks).

ABSTRACT

In this study, the crew's behavior and performance of a supervisory control task in an anti-aircraft artillery (AAA) system was investigated from an integrated-network-of-tasks point of view. The supervisory control task was characterized by a sequence of subtasks interconnected and performed by a commander, a gunner, and a range operator. A subtask flow chart was developed using the structure of SAINT. Distribution of various times to complete subtasks were constructed. Procedural data were categorized into different groups according to mode switching patterns, and analyzed with performance data to find strategies in the commander's decision-making process. Differences in procedure and completion times due to target maneuver, visibility, countermeasures, and C^3 information were examined and quantified.

I. INTRODUCTION

The tracking and firing performance of an individual tracker in an anti-aircraft artillery (AAA) system has been extensively studied in the past [1]-[4]. Human tracking data were collected to develop mathematical models for describing human tracking behavior in different system operational modes. These data and models were also used to evaluate the survivability of aircraft in different engagement scenarios [5]. Interest in the study of the performance of crews as a team in the AAA system has grown recently. Preliminary work has been done using the optimal control approach to model the interaction and decision-making process among crew members [6]-[7]. Due to lack of empirical data to validate the model, a considerable amount of additional work is needed before the model can be put into practical use. A different approach is proposed here to analyze the performance of crews in terms of key tasks completed and primary operating procedures adopted. The method of Systems Analysis of Integrated Networks of Tasks (SAINT) was applied to model crews performance as a sequence of interconnected subtasks.

The supervisory control task was broken down into a series of subtasks undertaken by individual crew members. More specifically, these subtasks included detection, mode decision, acquisition, mode actualization, tracking/breaklock, reacquisition, lead angle computer (LAC) solution, and firing. Task completion time was defined as the time required to complete each subtask. The commander was allowed to select one of three operating modes interchangeably in order to maximize the system performance in terms of hit score. Experiments were designed to collect empirical data which would characterize the distribution of task completion times and estimate the probability of selecting a system mode. Experiments were conducted on a AAA simulator at the

Air Force Aerospace Medical Research Laboratory in WPAFB, Ohio. Thirty-two conditions from a factorial combination of trajectory, visibility, electro-countermeasure (ECM), and C^3 input were studied to determine their effects on task completion time, operating procedure, and performance score.

Distributions of task completion time were constructed from the empirical data. Analysis of variance was performed on the task completion time to determine the effect of the independent variables. The crew's operating procedures were categorized into different groups according to mode switching pattern and peak switching time. A separate analysis of variance was done for each combination of trajectory and visibility to test the effect of procedure on overall performance. Significant effects of procedures were found in certain trajectories coupled with good visibility. Probabilities of mode switching were estimated from the procedural data based on visibility, target maneuver, and ECM presence. The task flowchart, the task completion times, and the probability estimates obtained here will be used to develop a complete SAINT model for the simulation of a AAA supervisory control system on a digital computer.

II. EXPERIMENT

Previous studies indicate that electronic countermeasures, complex trajectories, and target visibility have a degrading effect on tracking and firing accuracy. On the other hand, the effect of early warning information from the command, communication and control (C^3) network on crew performance is of interest from an integrated engagement point of view. Thus thirty-two experimental conditions were designed from a factorial combination of two levels of ECM, visibility, C^3 input, and four levels of trajectory (see Table 1).

Electronic countermeasures were either 1) present in the form of both a noise strobe and a false target or, 2) not present at all. Visibility of 10.8 KM and 2.0 Km were considered. C^3 input was presented in the form of target position on the commander's console display, either available or absent. Input trajectories included two weapon delivery trajectories and two flyby trajectories. Ten replications of each of the 32 conditions were obtained for each team. Five teams of subjects participated in the experiment and each team consisted of a commander, a gunner, and a range operator. This constituted a total sample size of 50 for each condition. The presentation order of the 32 engagement scenarios were randomized within each replication.

The functional roles of each crew member were defined and structured so that subjects could be trained. The subtasks of the commander included 1) assessing the situation, 2) deciding and requesting the desired system mode, and 3) monitoring the situation (including gunner performance, capability limit, and change of target status). The gunner's subtasks were divided into 1) searching for the target using either an optical device or a radar screen, 2) acquiring the target, 3) actualizing the system mode selected by the commander, 4) tracking the target, 5) reacquiring the [illegible], and 6) firing the guns. The subtasks of the range operator were 1) he[illegible]ng detect the target, 2) acquiring/reacquiring the target, and 3) tracking the target's range.

There were three system modes of operation available for the commander's choice. These modes were designated as Mode A (automatic tracking mode),

Mode S (semiautomatic tracking mode) and Mode M (manual mode). The system was considered to be in Mode O if none of the above modes had been actualized by the gunner. The general configuration of the manned AAA system is shown in Figure 1.

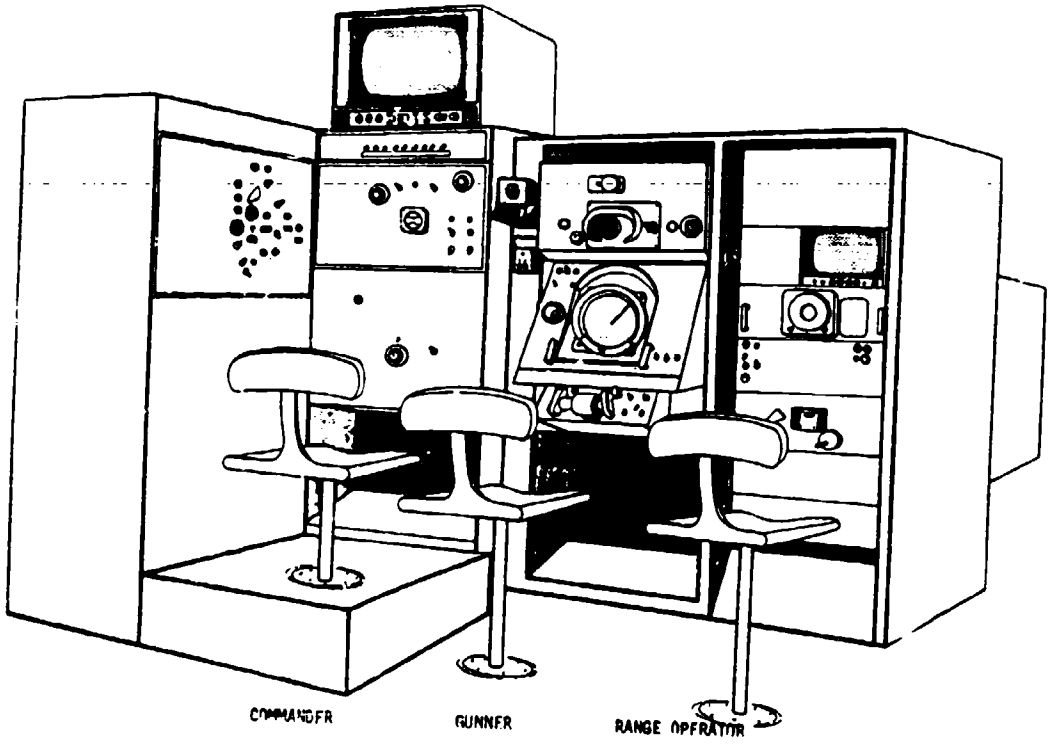


Figure 1. General Configuration of an AAA System

A command originated with the commander and was carried out by the gunner and the range operator. On a typical trial, once the target was detected, the commander would choose a system mode to engage the target. A verbal command was then issued to the gunner and the range operator who would carry out that command. In the event that system capabilities were degraded due to ECM, visibility, or target maneuver, this information was fed back to the commander who would make a new decision between keeping the system in the same operating mode or changing to another mode. Onset and completion time of individual subtask were measured in the experiment. An overall performance measure, "percentage of hits", was developed and it was the ratio of the number of hits placed on the target to the number of rounds fired. This single measure of performance, in contrast to measures used in previous experiments, such as tracking error, tracer error, and miss distance, reflected the performance of the entire crew instead of individual crew members and measured performance for all modes of operation. In other words, the commander's decision as well as the gunner's and the range operator's tracking and firing all contributed to this single performance measure. This score was presented to the crew at the end of each trial as a feedback metric.

III. TASK ANALYSIS USING SAINT

In order to apply the SAINT technique [8] to modeling and simulation of a AAA supervisory control system, crew activities were divided into a sequence of interconnected tasks. Each task was represented and specified in terms of the symbolism and terminology of SAINT. The standard symbol of a task is depicted in Figure 2.

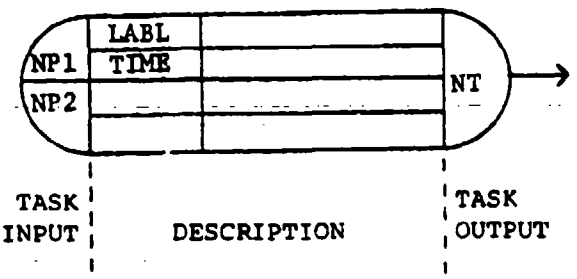


Figure 2. Task Symbol

A task is characterized by its input (precedence) relation with other tasks, description of the task, and its output relation with other tasks. Tasks are represented by nodes and connected by branches. The combination of branches and nodes forms a network model which serves as an abstract representation of a system. Once the network model is constructed from the system or from the task to be analyzed (in this case the AAA supervisory control system) the SAINT computer program can be used to simulate the task flow within the system.

The basic tasks involved in a supervisory control system were obtained by dividing the engagement scenario into separate phases. Each phase was represented by a task ur en by specified crew member(s) or other resources such as the lead-angle-computer. These tasks were detection of target, commanding the system mode, mode actualization, tracking, obtaining the lead angle solution, commanding the onset of firing, firing, and maintaining accuracy of firing. Based on this configuration, a SAINT model for a AAA supervisory control system was developed and is presented in Figure 3. The model presented here is far from complete, yet it provides a structured representation of the underlying system and tasks. It also provides a basis for defining major system variables to be measured from the current experiment and for performing preliminary task analyses. We will briefly describe each task and other relevant components of the model, such as state variables, user functions, and so forth.

Detection of Target was designated as Task 1 with a label DET. The completion time of this task could be determined empirically, as was done here, or could be drawn from one of the eleven distribution functions available in SAINT: constant, normal, uniform, Erlang, lognormal, Poisson, beta, gamma, beta fitted to three estimates, triangular, and Weibull [8]. A moderator function, MODF 1, which is defined by the user, could be used to indicate values of the parameters of the desired distribution function based either on experimental data or some prior knowledge. Resource, RESR, designated the operator(s) or the machine performing the task. In this case, only one crew member was

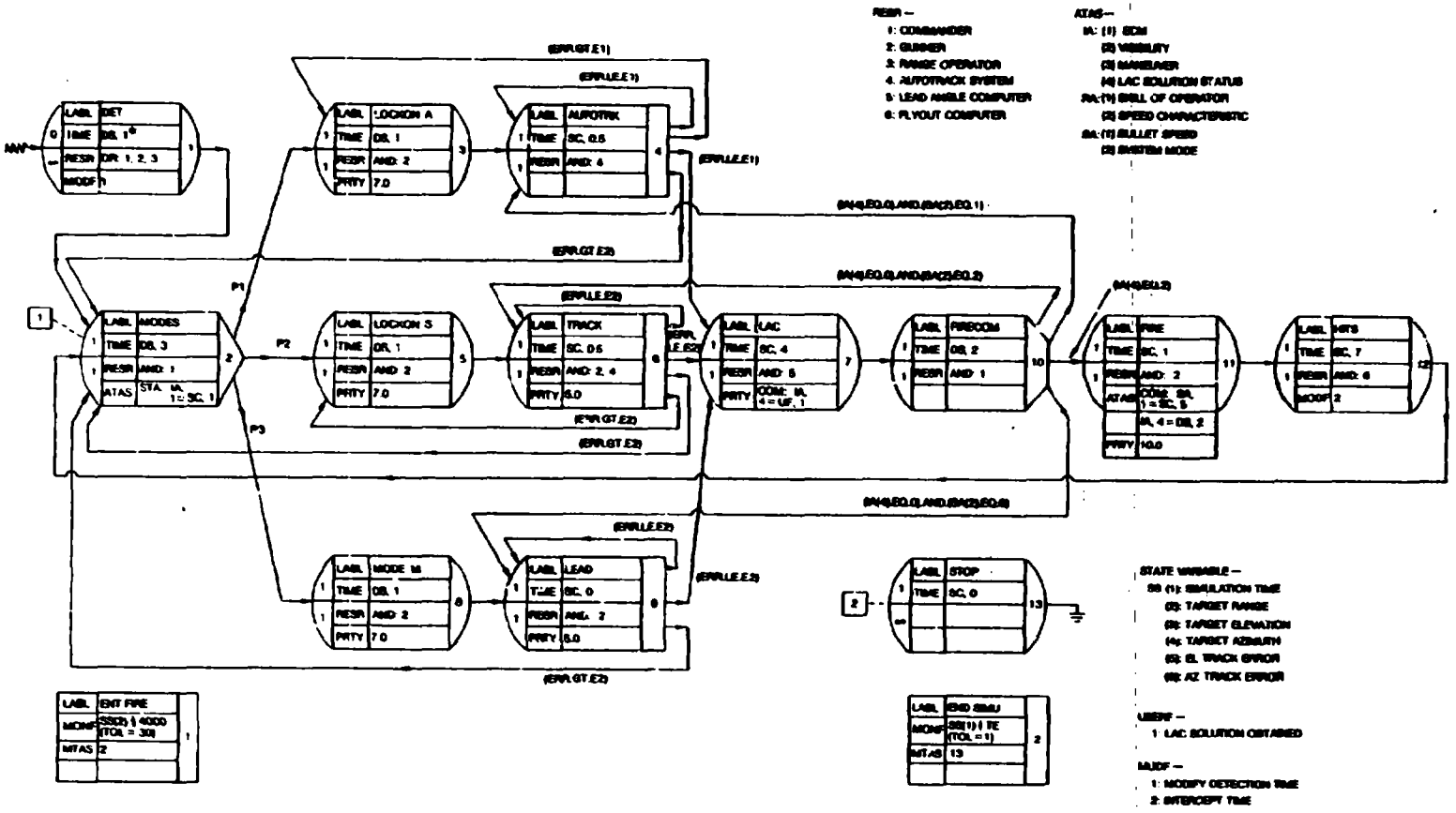


Figure 3. SAINT Model of an AAA Supervisory Control Task

required to detect the target. The commander, the gunner, and the range operator were designated as Resource 1, 2, and 3, respectively. OR indicated that any one of the crew members could have performed the task. On the input side of the task node, the wavy arrow indicated that the first release of the task did not require the completion of a preceeding task. The OO sign meant the detection task would only be performed once and no subsequent release of the task was expected.

Commanding the System Mode, performed by the commander, was designated as Task 2 with a label MODES. Since the commander could select one of three available system modes, the output of this task had three branches. Each branch leads to a different mode actualization task with a branching probability P_i. In the next section, we will generate preliminary estimate of these probabilities based on empirical data collected in this experiment. A monitor function ENT FIRE, denoted by ◫ --→, was reserved to be used as an indicator that the target was within the effective firing range and a manual mode should be adopted to engage the target immediately. Attribute assignment, ATAS, was used to assign value of various attributes: information attributes (IA), system attributes (SA), or resource attributes (RA). For instance, the ECM condition was represented by IA (1), with a value 1 meaning the presence of ECM and a value 0 meaning no ECM. In RESR row, AND indicated that all the resources listed, in this case only the commander, were required to perform the task.

Mode Actualization, performed by the gunner, was designated as Task 3 (LOCKON A), 5 (LOCKON S) or 8 (Mode M) corresponding to the three different modes which could have been commanded. In Modes A and S, the task involved making the proper mode switch throw, acquiring the target, and lockon to the target. In Mode M, the task only required making the proper mode switch throw. A task priority, PRTY, was developed to provide the scheduling of a task in case of a resource conflict. Since both Tasks 5 (LOCKON S) and 6 (TRACK) required Resource 2 (the Gunner) to perform them and Task 5 should have been completed first if both tasks were waiting to be performed, a higher priority (7.0) was given to Task 5.

Tracking, performed by the gunner and/or the autotrack system, was designated as Task 4 or 5 corresponding to Mode A or S. In Mode A, the autotrack system would track the target's range and angular position automatically. If the tracking error, ERR, was within certain system tolerance, E1, the system would continue to track and release the next task (lead angle solution). If the tracking error exceeded E1, the system would breaklock and Task 3, the mode actualization task, would be released again. If the tracking error were large enough to cause a complete loss of target (exceeded system criterion E2) the commander would be informed so that he could consider a possible system mode change. A conditional-take-all branching, which is a deterministic branching mechanism, was used to perform the required error testing and branching. Should a probability branching be desired, probabilities of releasing each subsequent task need to be assigned a priori or estimated from the empirical data. In Mode S, the only difference from Mode A was that the gunner would track the angular position of the target.

Obtaining the Lead Angle Solution, performed by the lead angle computer or the gunner, was designated as Task 7 for Modes A and S or Task 9 for Mode M. In the former case, the lead angle computer would take a fixed time to calculate a lead angle and direct the barrel pointing accordingly. An indicator

light would come on when a solution was reached. In the latter case, the gunner would calculate a lead angle and direct the barrel pointing based on his estimation of target velocity and the miss distance of previous rounds fired.

Commanding the Onset of Firing, performed by the commander, was represented by Task 10. A conditional-take-first branching was used here to select the first branch which satisfied the accompanying condition. For instance, IA(4) denoted the solution status of lead angle computer. When IA(4) equaled 1, it meant a solution had been achieved and subsequently, a firing command would be issued which would release Task 11 and result in actual firing. When IA (4) equaled 0, it meant a solution had not been achieved and no firing command would be issued. Instead, either Task 4, 6, or 9 would be released.

Firing, performed by the gunner, was denoted by Task 11. A high priority (10.0) was given to this task compared with tasks of mode actualization and tracking. The attributes of the initial speed and firing angle were assigned at the completion of this task. These attributes were used as parameters in the following task which involved the computation of intercept and miss distance of rounds fired.

Maintaining Accuracy of Firing, performed by the flyout computer, was designated as Task 12. The time required to complete this task was a variable depending upon the target range and firing angle. A moderator function was therefore reserved for modification of the completion time on the basis of the assigned attribute values and state variable values. System variables which changed value continuously over time were referred as state variables. Dynamic equations (algebraic or differential equation) which governed time-dependent behavior of state variables were used in the subroutine STATE to update the value of these variables at each time step. Variables of interest included the simulation time SS(1), the range, elevation and azimuth angle of the target, SS(2), SS(3) and SS(4); and the elevation and azimuth tracking error, SS(5) and SS(6). One example of the STATE subroutine would be to apply existing human operator models [3] [4] to generate tracking error under various operational conditions. For those cases in which dynamic equations were not available for a specific variable, a user-generated subroutine UINPUT could be used to read in the values of this variable directly from input and store them in an array for later use.

Additional user-generated subroutines USERF were defined to provide computation of other system functions and data communication between tasks. Examples are the computation of lead angle in Mode A, S, and M, the intercept time and the miss distance or hit score whenever a round was fired. The simulation stopped when the simulation time first exceeded a designed time TE, specified in input data. The [2] monitor function END SIMU and Task 13 were designed to end the simulation.

IV. RESULTS AND DISCUSSION

Analysis of Task Completion Time

Task completion time for detection of target (T_D), command system mode (T_{MD}), and acquisition (T_A)* were collected from the experiment.

* Acquisition task begins at the completion of detection and ends at the completion of mode actualization.

Examination of histograms of these times indicated that the data were skewed to the right (See Figure 4). These sample distributions were compared to both the expoential and lognormal distributions. Detection time, T_D, was found to be approximately lognormally distributed; time to command the system, T_{MD}, was approximately exponential; neither distribution yielded a satisfactory fit to time of acquisition, T_A. Further work will be done to find an appropriate description of these sample distributions, so that these times can be used in future SAINT simulations. These data are summarized in Table 1, which shows the means and standard deviations of the log-transformed values.

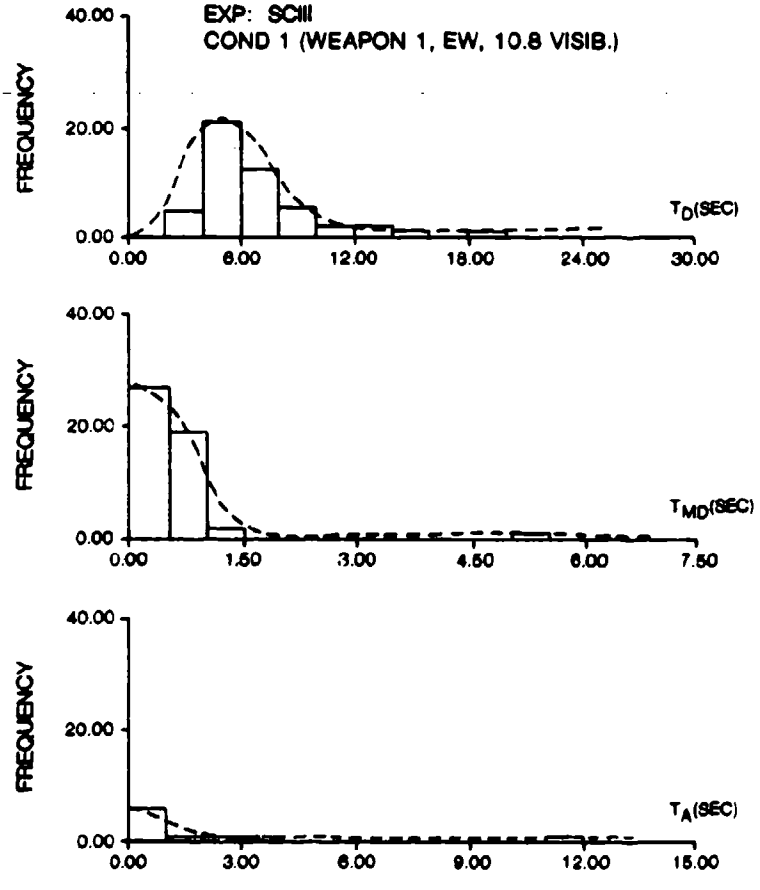


Figure 4. Frequency Distribution of T_D, T_{MD}, and T_A

Separate analyses of variances (ANOVA) were performed on the task completion time to determine the effect of independent input variables. Prior to these analyses, the data were transformed using the logarithmic transformation so that the distributions were approximately symmetric. Due to a large number of empty cells in the conditions using the WEAPON 2 trajectory under the 2.0 km visibility the data set was divided into two parts.

One consisted of all conditions using the trajectories WEAPON 1, FLYBY 1 and FLYBY 2, the other consisted of the 10.8 km visibility conditions of trajectory WEAPON 2. The first set of data was analyzed by 3 x 2 x 2 x 2 x 5 mixed model ANOVA with main factors trajectory, visibility, ECM, C^3 presence or absence and team (random). The second set of data was analyzed by 2 x 2 x 5 mixed model with main factors ECM, C^3 and team (random).

For the three trajectory model of target detection time, T_D, ECM had a significant main effect, and the interactions of ECM with visibility, C^3 with visibility, ECM with C^3, and visibility with ECM and C^3 were all significant. The interpretation of these results is as follows: for relatively easy conditions (high visibility and all of these three easy trajectories), the presence of ECM actually shortened detection time somewhat; C^3 had no effect. For low visibility, T_D increased if both (1) ECM was present and (2) C^3 information was not available. For the model with only the WEAPON 2 trajectory and high visibility, the absence of C^3 information significantly increased target detection time. From what data were available from this trajectory in the low visibility condition, it appears that T_D was much higher for this group of conditions than any other conditions, and that neither ECM nor C^3 altered time of detection.

Visibility had a significant effect on mode decision time, T_{MD}, for the three-trajectory model. High visibility increased T_{MD}, presumably because more time was available to make the decision. Mode decision times were higher for the WEAPON 2 trajectory than for any of the other three, and high visibility decreased T_{MD}, rather than increasing it. This reversal may be due to increased task difficulty for this trajectory in the low visibility condition.

The interaction of visibility and C^3 had a significant effect on acquisition time, T_A, in the three-trajectory model. This was due to a decrease in T_A when C^3 information was available in low visibility conditions. The one trajectory model showed no significant differences due to ECM or C^3 for the high visibility conditions. Data from the low visibility conditions were too scanty to draw any conclusions from them.

Analysis of the Effects of Procedure

In this experiment, the percentage of hits was used to measure the effectiveness of the crew's procedure. The procedures under analysis were not predetermined before the experiment; instead, they were identified by examination of the actual behavior of the crews during the experiment. This allowed the crew to exercise different procedures on their own and arrive a set of preferred procedures. The following is a list of procedures which were compared:

1) Mode A maintained during entire trial
2) Mode M maintained during entire trial
3) Switch from Mode A to Mode S early in the trial
4) Switch from Mode A to Mode S late in the trial
5) Switch from Mode A to Mode M early in the trial
6) Switch from Mode A to Mode M late in the trial
7) Switch from Mode A to Mode S, then to Mode M

Analysis of variance was applied to compare percentage of hits among procedures such that the trajectory and visibility factors were constant and trials from different teams were used interchangeably. This resulted in eight ANOVA designs, one for each combination of trajectory and visibility, which were unbalanced but had no empty cells. Only two of the eight ANOVA's showed a significant difference in performance due to procedure ($\alpha = 0.05$). Specifically, in the FLYBY 1 trajectory with visibility 10.8 km the best procedure for maximizing percentage of hits was Procedure 3, followed by Procedure 1, then Procedure 4. This is illustrated in Figure 5 in which three procedures (time history of commander's mode selection) were ranked from top to bottom according to their performance score.

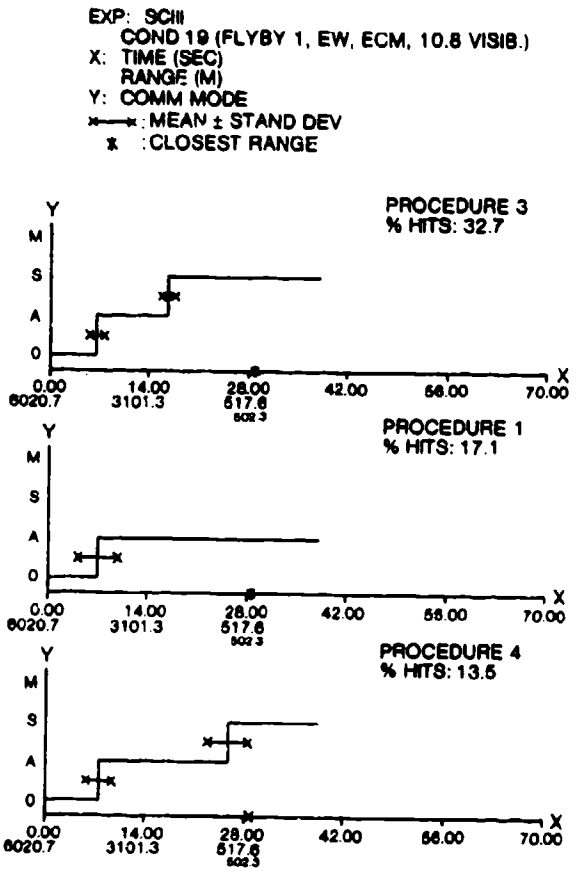


Figure 5. Procedure vs Performance Score - FLYBY 1

There are two interesting aspects to this result. First, Procedure 1 was considered in the past to be a preferable procedure to engage a flyby or non-maneuvering trajectory. Nevertheless, the result in this study suggested that an alternate procedure, Procedure 3, was preferable in the event of good visibility with possible presence of ECM. The second aspect is that the timing of mode change has a critical impact on the performance score. In other words, a mode change at peak time about 18.0 sec in Procedure 3 resulted in a considerably higher percentage of hits than if the mode change occurred later (at 25 sec), as in Procedure 4.

In the Flyby 2 trajectory with visibility 10.8 km, Procedure 1 was significantly better than Procedures 3 and 4. This is not surprising because Procedure 1 was considered in general to be a preferable procedure for a flyby trajectory. One reason that Procedure 3 did not emerge as a better procedure, as in Flyby 1 case, may have been due to the fact that in this trajectory the target was much further away from the AAA system, which made it difficult to hit in manual tracking model.

The effect of procedure on percentage of hits revealed in the empirical data can be compared with the results obtained in the SAINT simulation once the SAINT model is completed. Based on the comparison, additional experimentation can be designed to assess the impact of procedure on performance score and to validate the model. In addition, extensive SAINT simulations can be used to determine preferrable procedure(s) for a given engagement condition.

Estimate of Mode Selection Probability P_i

The probability that the commander would select a certain system mode could be defined across time or in more detail as a function of time. For this application, it was impractical to estimate the probability at each time step. The estimates of mode selection probability obtained here were event-based instead. Empirical data indicated that C^3 input had no significant effect in the determination of procedure. Therefore, target visibility, target maneuvering and ECM presence were chosen as the events to predict the values of mode selection probability. The fact that any one of the three events were time-dependent made the mode selection probability an implicit function of time.

An index function $I_v(t)$ was defined for the target visibility event. $I_v(t)$ took a value "0" if the target was not visible at time t; otherwise, it took a value "1". Similarly, index functions $I_M(t)$ and $I_c(t)$ were defined for the target maneuver event and the electro-countermeasure event. $I_M(t)$ equaled 1 if the target was taking a maneuvering course; otherwise, it equaled 0. $I_c(t)$ equaled 1 if the electro-countermeasure was present, otherwise, it equaled 0. The mode selection probability P_i (I_v, I_M, I_c) was expressed in terms of the index functions. To estimate the mode selection probability P_i from empirical data, the percent of sample runs using each procedure were tabulated in Table 2. Notice that we combined Procedures 3 and 4 (defined in the analysis of procedure section) into one procedure characterized by a Mode A to Mode S switch because the switching time per se was not critical here. Similarly, Procedures 5 and 6 were combined into one. The column "other" consisted mostly of unsuccessful attempts to detect or engage the target.

From Table 2, data were grouped according to the events: visibility, maneuver and ECM, as listed in Table 3.

Event Index (I_v, I_M, I_c)	Condition
(0,*,*)	5-8, 21-24, 29-32
(1,1,*)	1-8
(1,0,1)	19, 20
(1,0,0)	17-18, 21-22

Table 3. Condition Grouping for Events

*: Both 0 and 1 included

Since the "other" procedure column was not of interest in figuring the probability, it was disregarded. Percent use of procedure was recomputed for each condition. Estimated mode selection probabilities were obtained by computing the average percent selecting Mode A, S and M in each group of

conditions corresponding to particular events. It should be pointed out that time-dependency of events were considered in the grouping of conditions and in estimating the mode selection probability. Also, conditions were grouped so that the probabilities of mode selection within each group showed similar patterns and were reasonably consistent with each other. For example, the event of the target not being visible was determined by the visibility (which was constant in time) in this experiment and the target range (which was a function of time). By examining the percent use of procedure across all the procedures in Cond. 5-8, 21-24, 29-32, a consistent estimate of probability of selection of Mode A for the event (0, *, *) was obtained. It is interesting to note that mode switching from Mode A to S or A to M occurred after the target became visible. Hence the probability of selecting Mode A with the target not visible equaled 1.0, disregarding the "other" column. Estimated probabilities are presented in Table 4.

Event Index / Est. Proba.	Not Visible	Visible		
		Maneuvering	Not Maneuvering	
			ECM	No ECM
	(0,*,*)	(1,1,*)	(1,0,1)	(1,0,0)
P_1	1.0	0.14	0.53	0.65
P_2	0.	0.38	0.47	0.35
P_3	0.	0.48	0.	0.

Table 4. Estimated Probability of Mode Selection

*: Both 0 and 1 included

Since the commander does not change mode constantly, it is expected that additional triggering mechanisms are needed in the mode selection task (MODES) to trigger the release of the task. The probability estimates obtained here are useful once the task is released and completed.

V. SUMMARY REMARKS

Crew performance in a AAA supervisory control task was studied using the SAINT method. The task was divided into a sequence of interconnected subtasks. A flowchart of subtasks was developed in which each subtask was characterized by a task completion time and a set of branching probabilities. An experiment was designed to collect crew procedural data from which distributions of task completion time were constructed and summary statistics were obtained. Estimated probability of mode switching in one subtask was calculated from the procedural data as a function of target visibility, target maneuver, and ECM presence.

Separate analyses of variance were performed on the task completion time data and the procedural data. Significant effects of the independent variables, C^3 input, and visibility, on task completion times were found. Significant effects of procedures, in terms of mode switching pattern, on performance score were also found in certain trajectories coupled with good visibility. It was also noticed that in some cases the timing of mode switch

had critical impact on the final performance score. It is anticipated that the data obtained in this study will be used to further refine a AAA supervisory control simulation model using SAINT. Existing tracker models and attrition models could also be incorporated into the simulation model to provide an overall crew-in-the-loop simulation.

ACKNOWLEDGEMENTS

This work was supported by the Air Force Aerospace Medical Research Laboratory, Wright-Patterson AFB, Dayton, Ohio, under Contract F33615-79-C-0500. The authors wish to thank Mr. W. Summers and Lt. M. Felkey for their valuable suggestions and support.

TABLE 1. SUMMARY STATISTICS OF T_D, T_{MD}, AND T_A (LOG_{10} SEC)

COND	VISIB.	ECM	C^3	$\log_{10} T_D$ MEAN	$\log_{10} T_D$ SD	$\log_{10} T_{MD}$ MEAN	$\log_{10} T_{MD}$ SD	$\log_{10} T_A$ MEAN	$\log_{10} T_A$ SD
	WEAPON 1 Trajectory								
1	10.8 KM	N	Y	.80	.21	-.31	.36	.32	.47
2			N	.72	.23	-.29	.23	.19	.12
3		Y	Y	.72	.23	-.32	.18	.69	.37
4			N	.73	.21	-.30	.23	.29	.84
5	2.0 KM	N	Y	.61	.18	-.34	.22	.07	.28
6			N	.57	.21	-.36	.22	1.66	-
7		Y	Y	.69	.24	-.39	.22	.78	.46
8			N	.92	.24	-.28	.35	.02	.81
	WEAPON 2 Trajectory								
9	10.8 KM	N	Y	.69	.19	-.14	.42	.27	.75
10			N	.73	.26	-.10	.40	.17	.71
11		Y	Y	.75	.22	-.13	.41	.42	.35
12			N	.89	.19	-.08	.36	.36	.38
13	2.0 KM	N	Y	1.29	.18	-.08	.43	.55	.27
14			N	1.32	.08	.12	.49	.29	.39
15		Y	Y	1.33	.07	-.17	.36	.50	-
16			N	1.34	.08	.07	.56	.66	.33
	FLYBY 1 Trajectory								
17	10.8 KM	N	Y	.74	.19	-.33	.30	.14	.49
18			N	.73	.21	-.30	.27	.02	.66
19		Y	Y	.64	.16	-.31	.24	.84	.39
20			N	.65	.21	-.31	.30	.26	.35
21	2.0 KM	N	Y	.69	.23	-.34	.28	-.10	-
22			N	.59	.24	-.34	.26	1.32	-
23		Y	Y	.64	.20	-.35	.30	.21	.43
24			N	.93	.26	-.29	.25	.10	-

TABLE 1. (CONTINUED)

COND	VISIB.	ECM	C^3	$\log_{10} T_D$ MEAN	$\log_{10} T_D$ SD	$\log_{10} T_{MD}$ MEAN	$\log_{10} T_{MD}$ SD	$\log_{10} T_A$ MEAN	$\log_{10} T_A$ SD
	FLYBY 2 Trajectory								
25	10.8 KM	N	Y	.78	.23	-.18	.38	.72	.43
26			N	.73	.17	-.25	.36	.20	.61
27		Y	Y	.68	.19	-.20	.39	.40	.52
28			N	.64	.21	-.20	.40	.24	.44
29	2.0 KM	N	Y	.64	.26	-.33	.26	.22	.46
30			N	.68	.22	-.28	.26	.17	.28
31		Y	Y	.71	.28	-.29	.26	-.03	.37
32			N	.96	.14	-.28	.26	.33	.43

TABLE 2. PERCENT USE OF PROCEDURE

COND	VISIB.	ECM	C^3	Mode A	Mode M	A→S switch	A→M switch	A→S→M switch	Other
	WEAPON 1 Trajectory								
1	10.8 KM	N	Y	12.0	0.	22.0	56.0	0.	10.0
2			N	12.0	0.	22.0	66.0	0.	0.0
3		Y	Y	0.	0.	24.0	64.0	0.	12.0
4			N	0.	0.	24.0	60.0	0.	16.0
5	2.0 KM	N	Y	18.0	0.	48.0	28.0	0.	6.0
6			N	18.0	0.	52.0	24.0	0.	6.0
7		Y	Y	14.0	0.	60.0	26.0	0.	0.0
8			N	32.0	0.	36.0	28.0	0.	4.0
	WEAPON 2 Trajectory								
9	10.8 KM	N	Y	0.	10.0	0.	74.0	0.	16.0
10			N	0.	20.0	0.	60.0	10.0	10.0
11		Y	Y	0.	16.0	0.	74.0	0.	10.0
12			N	0.	22.0	0.	76.0	0.	2.0
13	2.0 KM	N	Y	36.0	22.0	0.	0.	0.	42.0
14			N	20.0	18.0	0.	0.	0.	62.0
15		Y	Y	30.0	16.0	0.	0.	0.	54.0
16			N	26.0	16.0	0.	0.	0.	58.0
	FLYBY 1 Trajectory								
17	10.8 KM	N	Y	54.0	0.	34.0	0.	0.	12.0
18			N	56.0	0.	38.0	0.	0.	6.0
19		Y	Y	48.0	0.	46.0	0.	0.	6.0
20			N	50.0	0.	40.0	0.	0.	10.0
21	2.0 KM	N	Y	64.0	0.	26.0	0.	0.	10.0
22			N	60.0	0.	26.0	0.	0.	14.0
23		Y	Y	54.0	0.	32.0	0.	0.	14.0
24			N	64.0	0.	26.0	0.	0	10.0

TABLE 2. (CONTINUED)

CONL	VISIB.	ECM	C^3	Mode A	Mode M	A→S switch	A→M switch	A→S→M switch	Other
	FLYBY 2 Trajectory								
25	10.8 KM	N	Y	54.0	0.	32.0	0.	0.	14.0
26			N	64.0	0.	28.0	0.	0.	8.0
27		Y	Y	32.0	0.	40.0	20.0	0.	8.0
28			N	48.0	0.	28.0	14.0	0.	10.0
29	2.0 KM	N	Y	86.0	0.	10.0	0.	0.	4.0
30			N	94.0	0.	0.	0.	0.	6.0
31		Y	Y	92.0	0.	0.	0.	0.	8.0
32			N	88.0	0.	0.	0.	0.	12.0

REFERENCES

1. Kleinman, D.L. and T.R. Perkins, August 1974, "Modeling Human Performance in a Time-Varying Anti-Aircraft Tracking Loop", IEEE Trans. Automat. Contr., vol. AC-19, No. 4, pp. 297-306.

2. McRuer, D.T., and E.S. Krendel, 1974, "Mathematical Models of Human Pilot Behavior," AGARD-AG-188.

3. Kou, R.S., et al., 1978, "A Linear Time-Varying Anti-aircraft Gunner Model Based on Observer Theory," Proceedings of the International Conference on Cybernetics and Society, Vol. II, pp. 1421-1426.

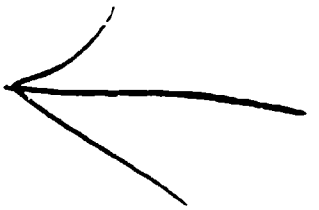
4. Wei, K.C., Bianco, W., and Vikmanis, M.M., "An Anti-aircraft Gunner Model with Delayed Tracer Measurements," Proceedings of 1980 Joint Automatic Control Conference, San Francisco, CA.

5. Summers, W.C. and Whitmeyer, J., "P001/OBS3/6: A Human Operator Augmented AAA Simulation Model (U)", Air Force Aerospace Medical Research Lab. Tech. Report AFAMRL-TR-81-108, CONFIDENTIAL, WPAFB, OH. May 1981.

6. Zacharias, G., Baron, S. and Muralidharan, R., "A Supervisory Control Model of the AAA Crew," Technical Report No. 4802, Bolt, Beranek and Newman, Inc., Cambridge, MA., Oct. 1981.

7. Hale, C., and Valentino, G., "Application of Optimal Control Principles to Describe the Supervisory Control Behavior of AAA Crew Members," Proceedings of the Sixteen Annual Conference on Manual Control, Cambridge, MA., May 1980.

8. Wortman, D., et al., "Simulation Using SAINT: A User-oriented Instruction Manual," Air Force Aerospace Medical Research Lab Tech. Report AMRL-TR-77-61, WPAFB, OH., July 1978.



AD P001628

HUMAN RESPONSE MODELING AND FIELD TEST VALIDATION OF AN AIR DEFENSE GUN SYSTEM SIMULATION

Ying-Kuh Yin, K. Sahara, V. Dea and A. Netoh
GENERAL DYNAMICS, POMONA DIVISION

SUMMARY

A mathematical model of a gunner was developed to represent the response of a well-trained and motivated operator of a dual axis compensatory tracking task of a high performance air defense gun system. The model is based on the quasi-linear transfer function approach. The parameters used in the model were derived from the tracking experiments of 16 gunners who operated a real time man-in-loop simulation against a variety of targets.

This human math model was then integrated into a 3-D Fortran optic track and fire simulation. It was validated by comparng the critical parameters of the simulation results with field test results. The comparison were made by both Monte Carlo methods as well as by single runs. This simulation is the mathematical counterpart of the hardware, the software and the operator of an advanced air defense gun system.

INTRODUCTION

During the course of developing an advanced air defense gun system, a Fortran computer simulation program was written in order to facilitate the system design, hardware and software integration, and performance prediction. The optic manual control mode required a mathematical model of a human operator's response . Techniques of modeling such an operator as a control element has long been established by pioneering work such as McRuer's cross-over model[1], Hess's dual loop model[2] and Kleinman's optimal control model[3]. However, we were most concerned with the application of an human model in a digital computer simulation in order to study the system performance rather than the model itself. The transfer function approach was chosen because of its simplicity and more importantly because the modelling could proceed in parallel with system design and thus also served as a design tool.

REAL TIME MAN-IN-LOOP SIMULATION

In order to provide the means of parameter quantification of the human math model, we used a real time man-in-loop simulation. The simulation hardware consists of a PDP 11/34 computer system, a Megatek graphics display, and a control handle set which enables the operator to interface with the computer. A silhouette of the target appears on the

screen according to a pre-defined trajectory. The operator closes the optic track loop when he attempts to keep the reticle of the optic sight on the target by using the thumb force transducer on the handle. The whole simulation runs in real time. Time histories of the critical variables were recorded for off line analysis. Fig. 1 shows a pictorial description of the simulation.

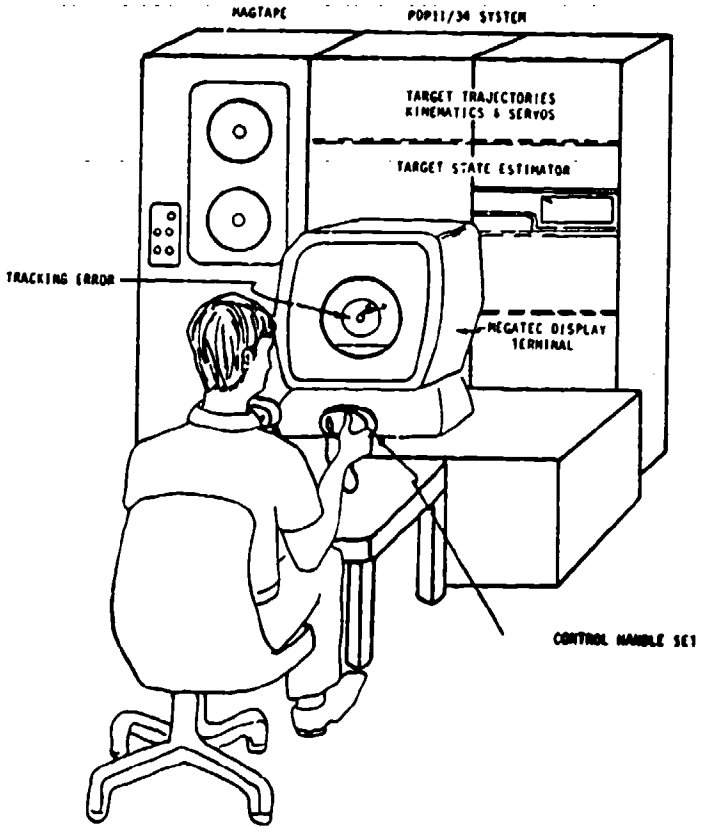


Fig.1 REAL TIME MAN-IN-LOOP SIMULATION

HUMAN MATH MODEL

The transfer function which represents the deterministic part of an operator's response in a compensatory tracking task is taken from McRuer[1]. the transfer function of the deterministic part is given below and a block diagram of the model is shown in Fig. 2.

$$K_p e^{-\tau s} \frac{(\tau_L S+1)}{(\tau_I S+1)} \frac{1}{(\tau_N S+1)}$$

Where

KP	:	gain
τ	:	reaction time delay
$(\tau_N S+1)$	:	first order neuromuscular lag.
$\frac{\tau_L S+1}{\tau_I S+1}$	:	human compensation

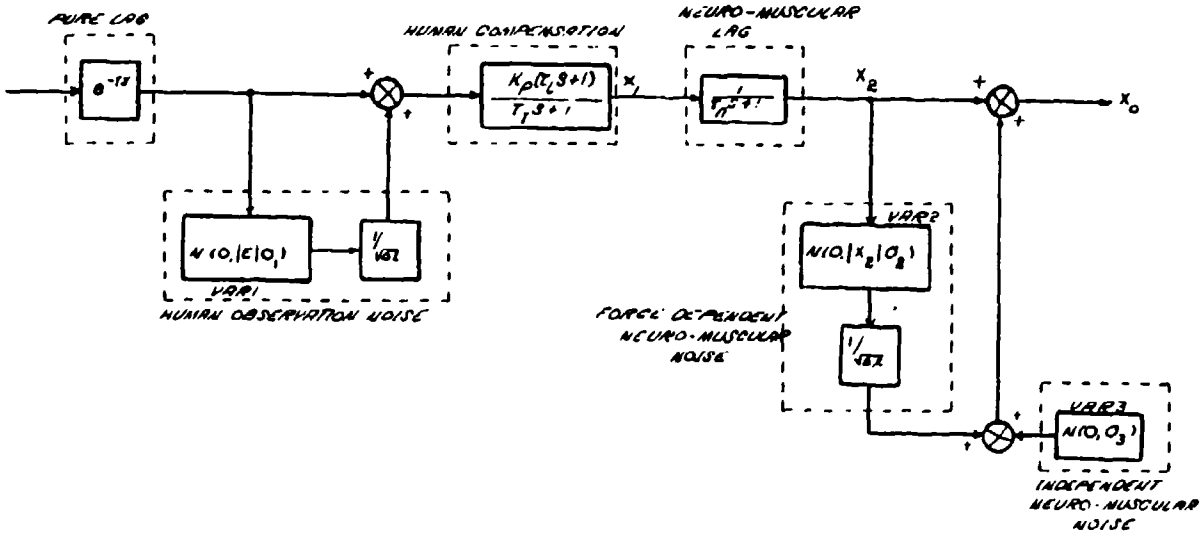


Fig.2 HUMAN MATH MODEL

The remnant is modelled by three noises which represent the uncertainty of the operator's estimates of visual inputs as well as control outputs plus a third independent noise source. The observation error is modelled by a Gaussian noise with variance propotional to the

error signal. The error signal represents the difference between the target line of sight and optic sight boresight center line. The neuromuscular noise is proportional to the thumb force output. The third noise is independent of outside stimulus. The need of this term was due to the real time man-in-loop tracking experiments of very low speed or near stationary targets. The numerical values of the model is listed in Table 1.

Table 1

τ : (from .072 to .162 sec)	τ_N : .1 sec
K_p : (from 292 to 382 volt/sec)	Var1: (from .00067 to .002)
τ_I : (from 18.0 to 22.0 sec)	Var2: (from .004 to .012)
τ_L : (from 2.0 to 5.0 sec)	var3: (from .004 to .012)

OPTIC TRACK LOOP

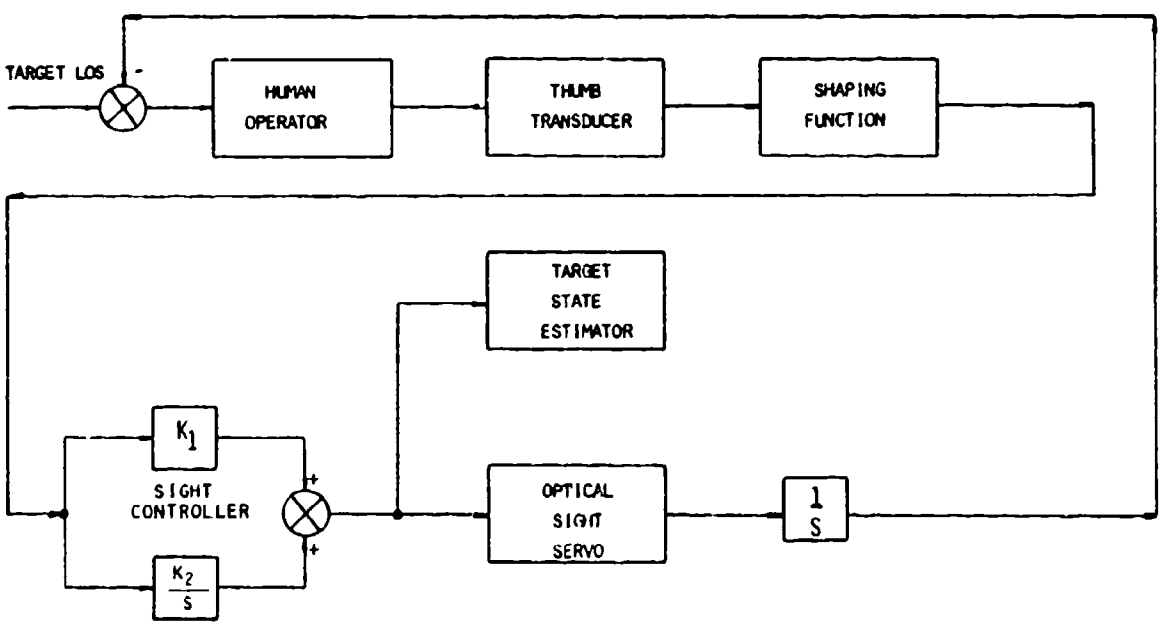


Fig.3 OPTIC TRACK LOOP

A single axis of the optic track system is depicted in Fig. 3. The

operator,upon seeing an angular tracking error through the optic sight, outputs a rate command through the thumb transducer to the sight controller. The controller consists of a propotional plus integral elements. The sight servo then follow the command and position the sight. Since the sight servo has a very wide bandwidth compared with this control loop, it is modelled as unity. The outputs from the sight controller are used together with the range information from the range finder to generate the target state estimates to direct the gun.

The implementation is the same for both the azimuth and elevation channels. However, the two axes are coupled through the observation and force output noises to simulate the inability of the operator to decouple exactly the two axes.

MODEL VALIDATION BY FIELD TESTS

During the field test, the range tracking system was used to record the trajectories of live targets. The time history of all critical system variables such as optical sight angular postions,sight rate commands,target range,as well as all parameters related to the gun were recorded on digital tapes. Furthermore, video films taken with a camera mounted on the commander's eyepiece (while the gunner was doing the tracking) were manually reduced to give tracking error.

Simulation validation was based on an incoming fixed wing aircraft and a crossing helicopter run. The target trajectories are shown in Figures 4 and 5. The target trajectory data was smoothed and then fed into the Fortran simulation. The simulation results were then compared with the field test results against the same target. The comparison were made by both the Monte Carlo methods as well as by single run match-ups. Before each Monte Carlo simulation run, the parameters listed in table 1 were randomly selected within the limits shown. For each time increment, the fifth lowest and the fifth highest values of the 100 run Monte Carlo set were recorded and plotted to form the 90 percent bounds. The field test results for the same variable were examined to see whether it was within the 90 percent bounds. We have also provided single runs which are the first run of the one hundred Monte Carlo run.

Fig 6 is the tracking error which was manually reduced from the video tape taken during the field test when the high speed incoming fixed wing target was tracked optically. Fig. 7 shows the tracking error of the target by the Fortran simulation.

Results against the fixed wing aircraft are presented in Figures 8,9,and 10. Time history of operator's command rates and optical sight angular postion in both azimuth and elevation channels recorded during field test are plotted in Figure 8. Fig 9 shows the corresponding results from the Fortran simulation. The field test data falls within the 90 percent bound most of the time. Fig 10. presents a single run . This run shows that the simultion has less high frequency noise than the field test.

Similarly Fig. 11 ,12, and 13 presents the field test , 100 Monte Carlo runs and single run respectively against the helicopter target. The comparison is again quite good.

It should be noted that the time mark of the field test results is different from that of the Fortran simulation. This is because that the field test time is based on mission start time while the simulation uses the trajectory tape begining as time zero. However, the two results were carefully lined up in real time frame. The comparison can be most easily made by overlaying the curves.

DISCUSSION

Field test and simulation comparison have demonstrated that the quasi-linear transfer function representation of an human operator works reasonably well for both high speed fixed wing aircraft and slow helicopter targets. However, the noise model still can be improved. For example, the independent noise should be modelled based on the typical human operator'natural frequency instead of being a purely white noise.

REFERENCE

[1] McRuer,D.T., Graham,D.,Krendel,E.,and Riesener,W.Jr., "Human Pilot Dynamics in Compensatory Systems," Air Force Flight Dynamics Laboratory,AFFDL-TR-65-15,1965

[2] Hess,R.A.,"Dual-Loop Model of the Human Controller," Journal of Guidance and Control, Vol. 1,No. 4, 1978

[3]Kleinman,D.L.,and Perkins,T.R., "Modeling Human Performance in a Time-Varying Anti-Aircraft Tracking Loop," IEEE Transaction on Automatic Control, Vol. AC-19,No. 4,August 1974.

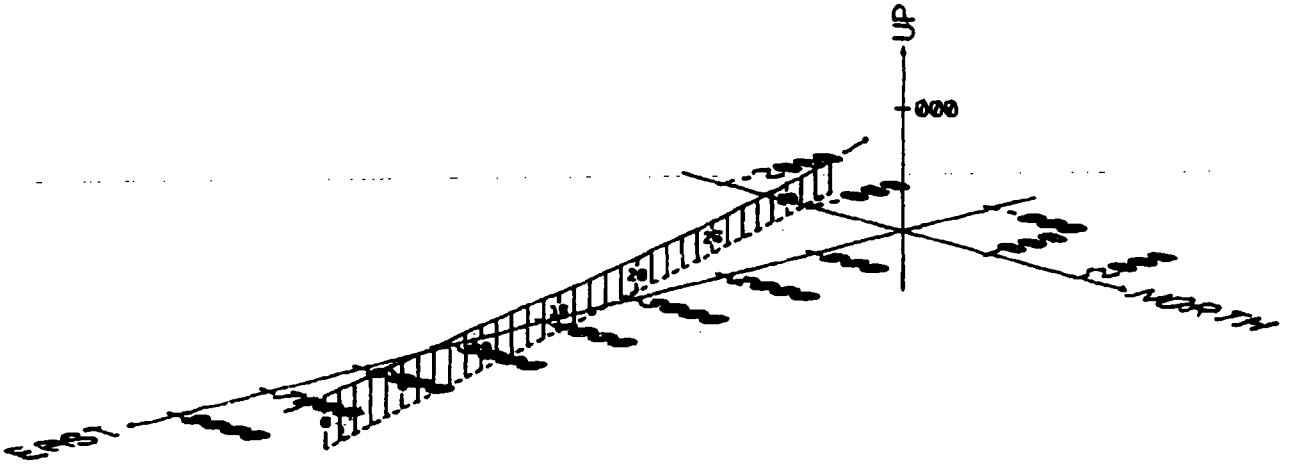


Figure 4. Fixed Wing Aircraft Trajectory.

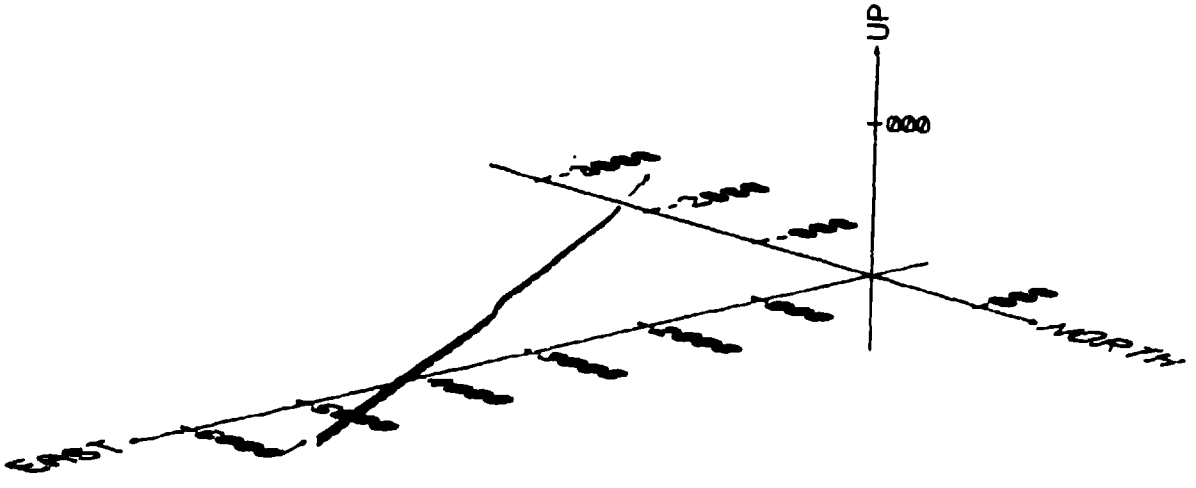


Figure 5. Helicopter Trajectory.

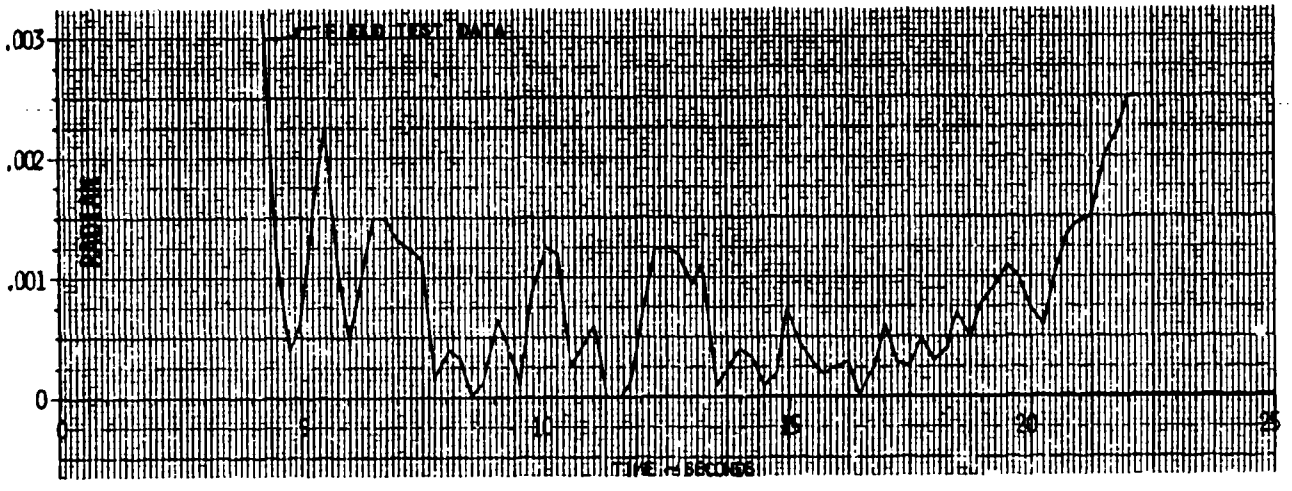


Figure 6. Tracking Error for Fixed Wing Target: Field Test

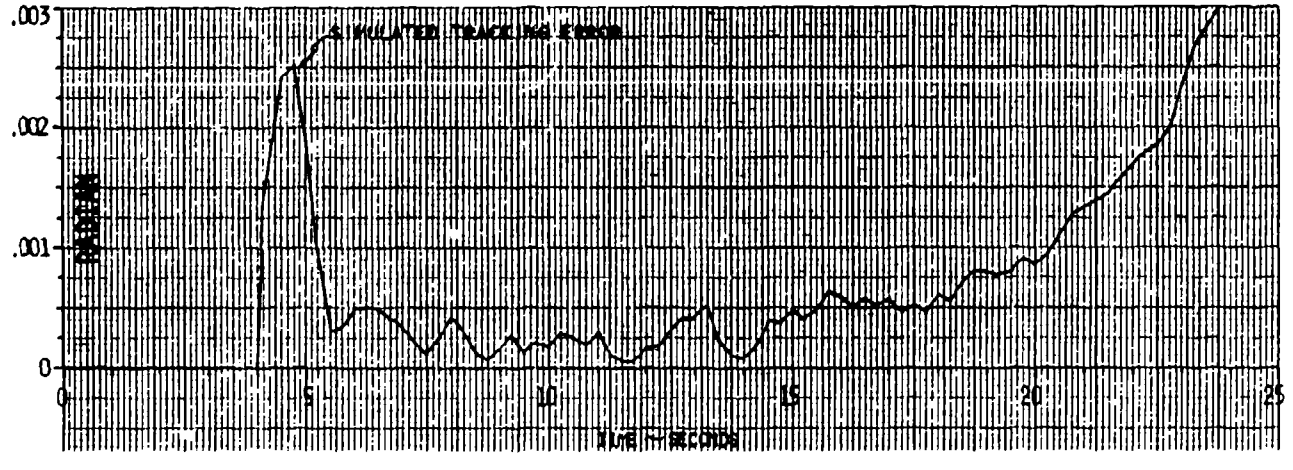


Figure 7. Tracking Error for Fixed Wing Target: Simulation

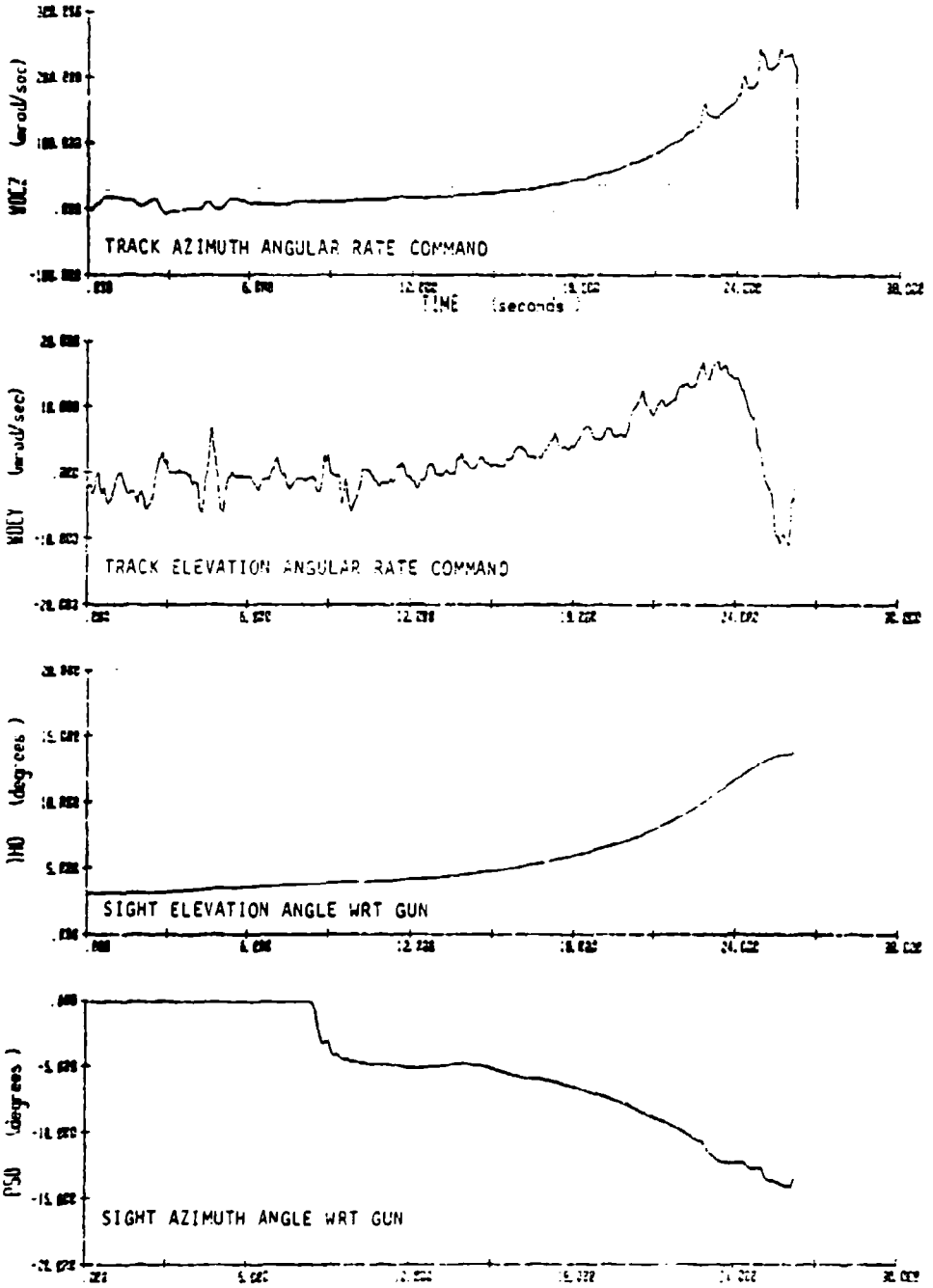


Figure 8. Field Test Data For Fixed Wing Aircraft.

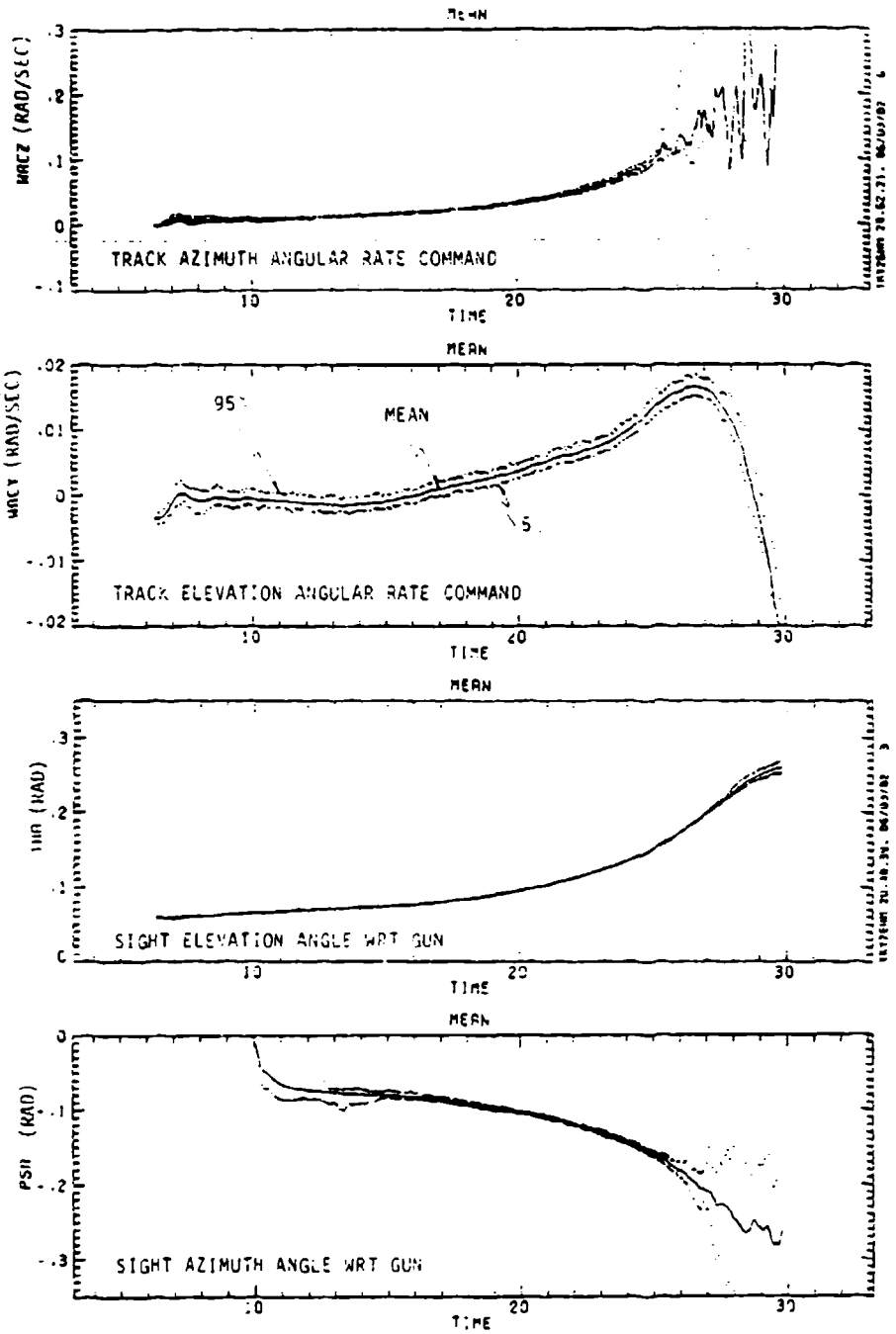


Figure 9. Simulation Results for Fixed Wing Aircraft (100 RUNS)

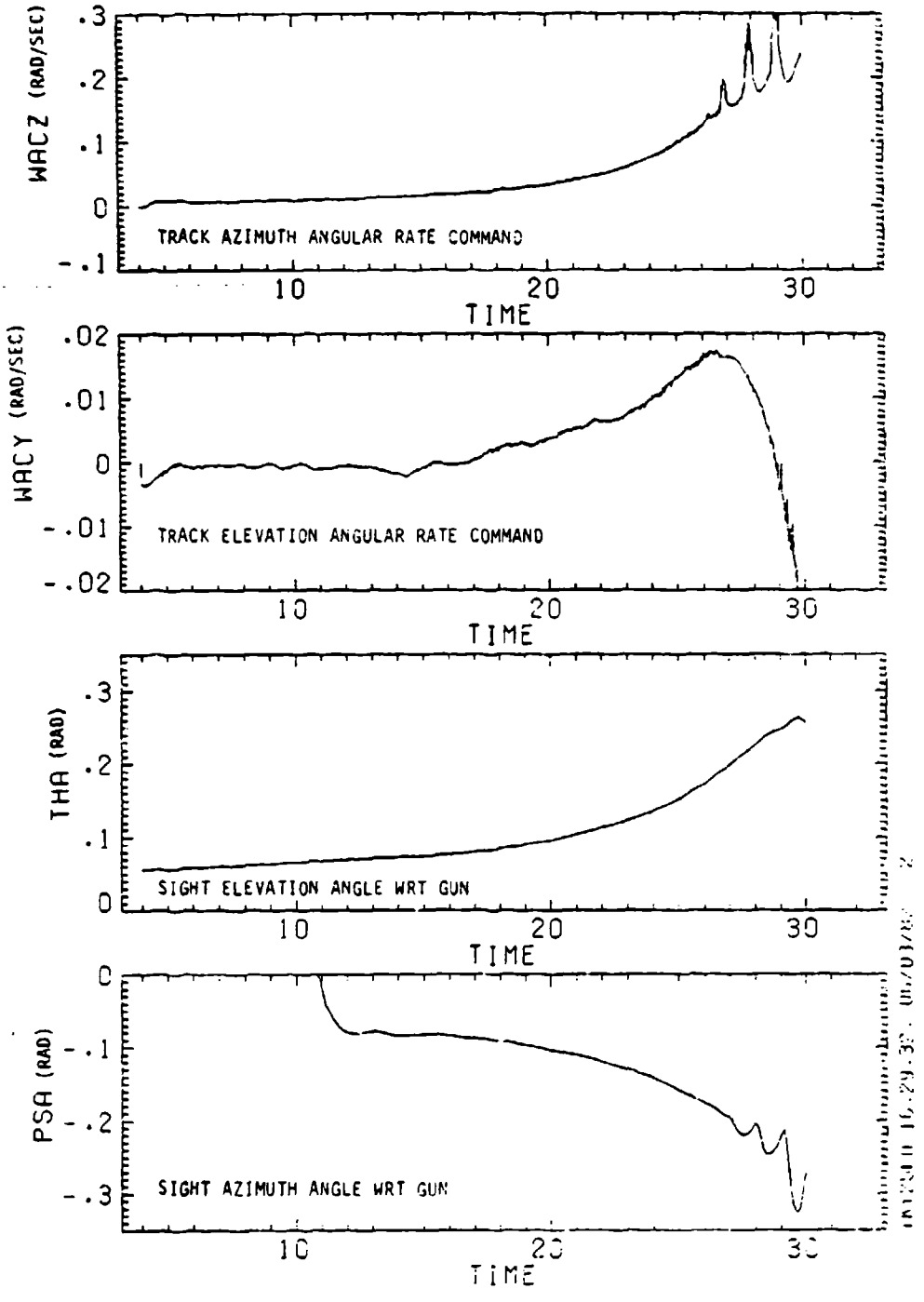


Figure 10. Simulation Results for Fixed Wing Aircraft (1 RUN)

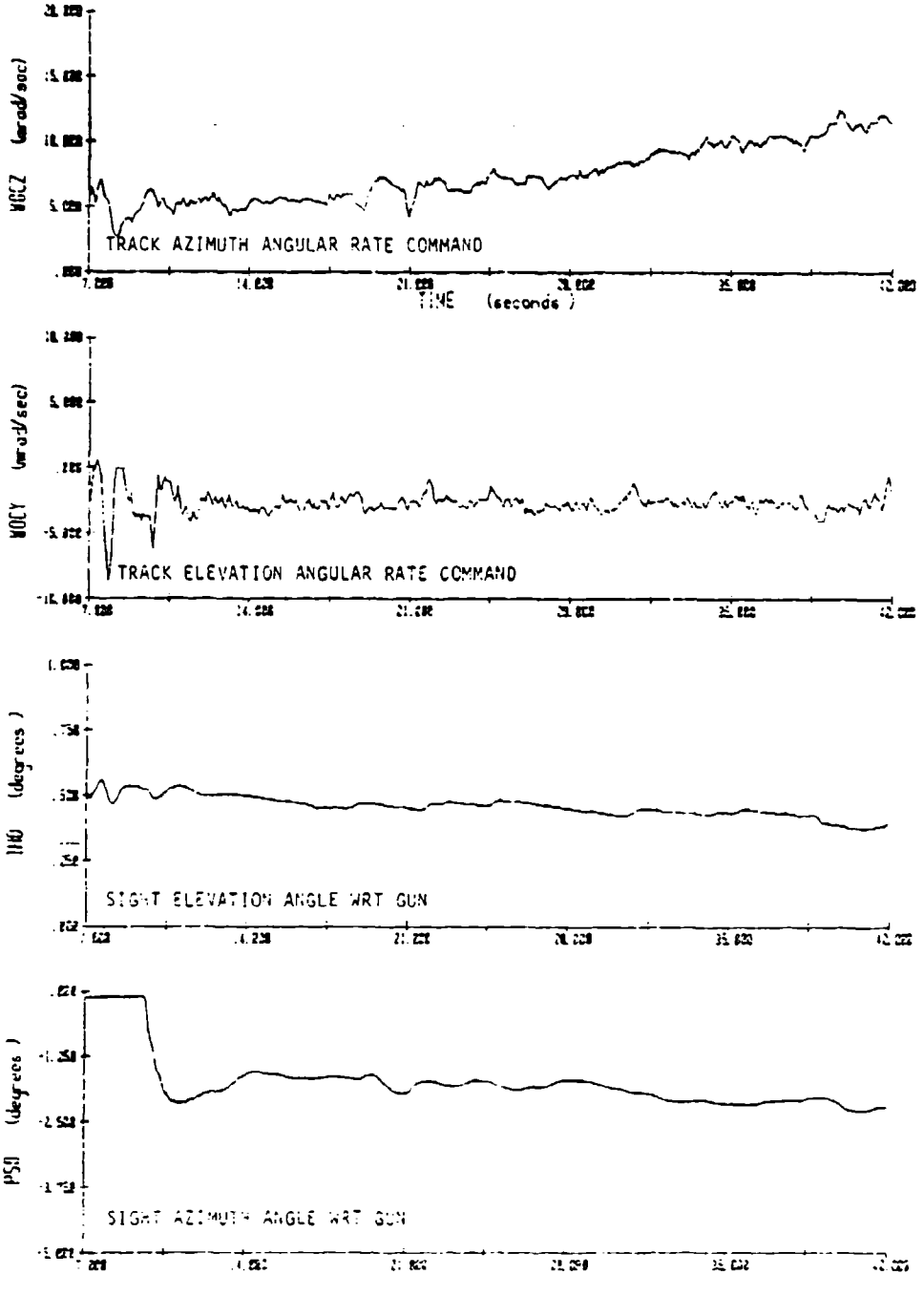


Figure 11. Field Test Data For Helicopter.

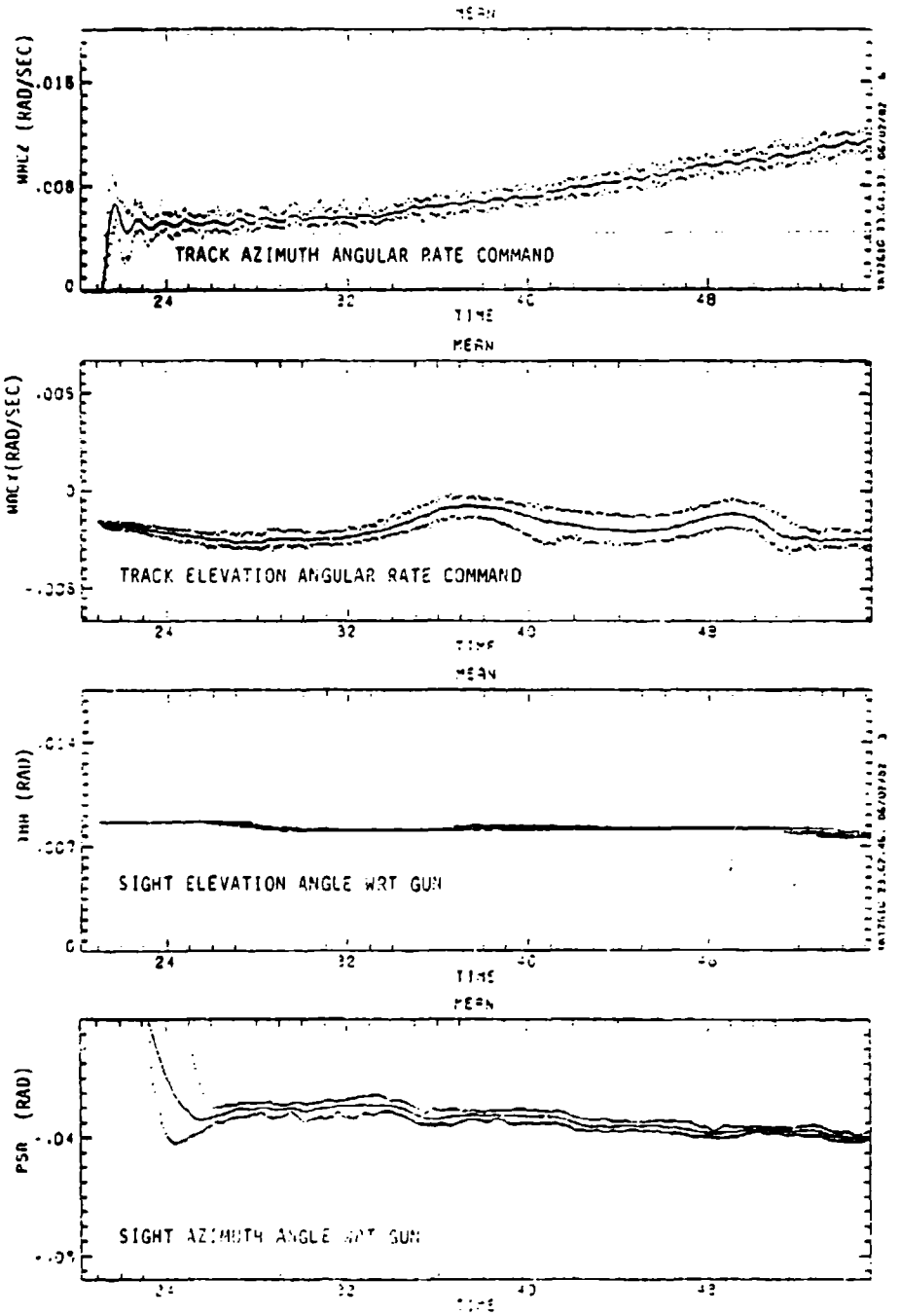


Figure 12. Simulation Results for Helicopter.
(100 RUNS)

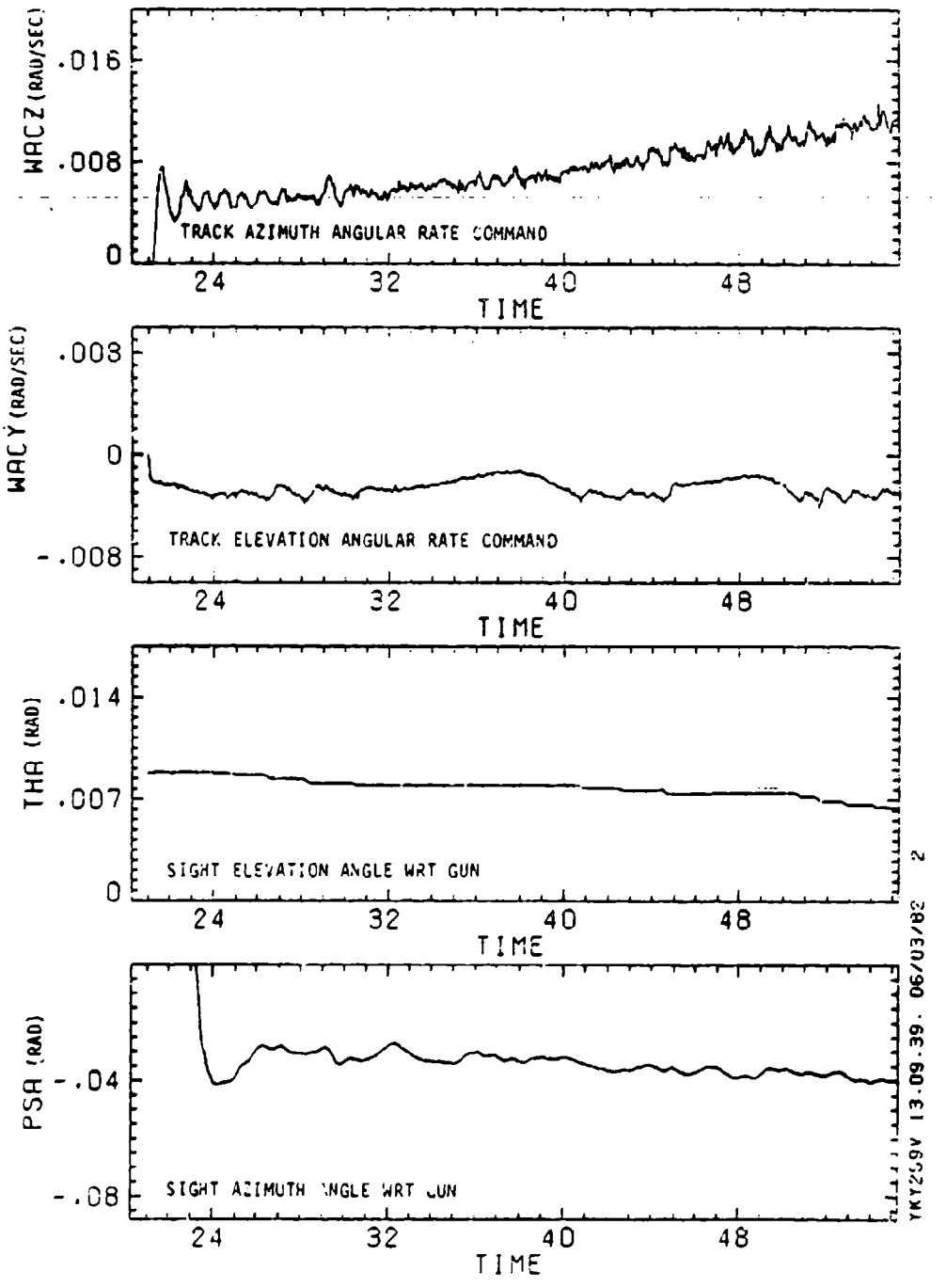


Figure 13. Simulation Results for Helicopter (1 RUN)

AD P001629

Effects of Practice on Pilot Response Behavior

by

William H. Levison
Bolt Beranek and Newman, Inc.
50 Moulton Street
Cambridge, MA 02238

Proceedings of the
Eighteenth Annual Conference on Manual Control
Dayton, OH
June 8-10, 1982

ABSTRACT

This study is directed toward development of an analytic tool for the design of training procedures and for the evaluation of trainee performance in tasks relevant to aircraft flight management. Laboratory tracking data have been analyzed to determine changes in pilot-related parameters of the optimal control model that best correlate with practice effects. Data from two independent studies show that, with continued practice, subjects lower their remnant and increase their response gain. These effects can be accounted for primarily by a reduction of the observation noise/signal ratio and, to a lesser extent, by a reduction of the motor time constant. Preliminary analysis suggests that, in part, the apparent practice-related change in motor time constant may reflect a relative insensitivity of performance to piloting strategy early in training when pilot noise levels are high, rather than a change in response bandwidth capabilities.

INTRODUCTION

This study has been directed toward the development of an analytic tool for the design of training procedures and the assessment of trainee performance in the kinds of monitoring, decision, and control tasks required for aircraft flight management. A more specific goal is to extend the optimal control model (OCM) of pilot/vehicle systems [1,2] into a predictive tool that relates the acquisition of continuous estimation and control strategies to the perceptual cueing environment. The first phase of this effort, the results of which are summarized in this paper, has been to analyze existing

manual control data to determine the effects of practice on pilot response behavior and to quantify changes in independent model parameters that best correlate with practice effects.

Data from two experimental studies are analyzed. The study referred to as the "Motion Cue Study" was conducted to explore the effects on pilot performance of a delay between platform motion and visual cueing [3]. Five groups of subjects were first trained to asymptote under various motion-cueing conditions: (a) fixed-base, (b) combined visual and motion cueing with no additional delays, and (c) different amounts of additional delay added to the motion cueing. The subject populations then transitioned to the zero-delay condition for further training. Data considered in this paper were obtained from the fixed-base and zero-delay motion populations during the pre-transition phase.

The study referred to as the "Performance Analyzer Study" was conducted as part of a research project to develop a tracking task for which performance would be particularly sensitive to task- and environment-related stress [4]. The tracking tasks performed in this study were fixed base and used a simulated vehicle for which the response characteristics contained an unstable root (divergence time constant of 0.5 seconds). The reader is referred to the literature for additional details on these two studies.

Four steps were involved in analyzing the effects of practice on pilot response behavior: (1) computation of standard engineering measures of control behavior (specifically, describing function and remnant) at different stages of practice; (2) use of these measures to identify independent model parameters, (3) statistical analysis to determine which of the apparent practice-related parameter changes are significant, and (4) interpretation of parameter changes in terms of practice-related differences in human operator information-processing capabilities.

The work reported in this paper has dealt mainly with the first three steps; adequate interpretation of results will necessitate further research. As discussed below, this effort was to some extent a methodological study as well as a study of practice effects.

METHODOLOGY

Parameter Identification

The quasi-Newton gradient search procedure [5-7] was used to identify independent -- or "pilot-related" -- parameters of the OCM. This procedure adjusts the independent model parameters to minimize a scalar matching error that is composed of time domain measures (typically, error, error-rate, control, and control-rate variances) and frequency-domain measures (estimates of describing function gain and phase shift, and pilot remnant).

Four classes of model parameters were identified from the tracking data: observation noise/signal ratios, motor noise/signal ratios, time delay, and motor time constant. Nomenclature and definitions for these parameter are as follows:

PY_e = Observation noise/signal ratio relevant to tracking error, in dB. This quantity is defined as the ratio of the observation noise covariance associated with perception of tracking error, normalized with respect to tracking error variance. (Tracking error was a zero-mean process.)

$PY_{\dot{e}}$ = Observation noise/signal ratio relevant to perception of tracking error rate.

$PY_{\ddot{e}}$ = Observation noise/signal ratio relevant to perception of tracking error acceleration. This parameter was identified for motion-base tracking only; the subject was assumed not to obtain useful error acceleration information from visual cues.

PUP = Pseudo-motor noise/signal ratio, in dB, defined as the ratio of pseudo motor noise covariance normalized with respect to the variance of the commanded control rate. The prefix "pseudo" signifies a noise process that only influences the pilot's response strategy; it does not reflect an actual noise process injected into the system. See Reference 8 for a detailed discussion of the motor noise aspect of the OCM.

TD = Time delay, in seconds.

TN = Motor time constant, in seconds, which relates to the

relative cost penalty on control-rate variance.* An increasing motor time constant reflects decreasing response bandwidth.

Independent model parameters were identified from data obtained early in training (referred to as "early" practice) and after training to a near-asymptotic level of mean-squared error performance ("late" practice). Data from between 2 to 4 experimental trials were averaged for a given subject performing a given task at a given level of practice. To minimize the effects of learning within a set of replications, trials yielding similar mean-squared error scores (for a given subject) were selected for the early practice phases. Thus, "early" data reflect trials in close succession, but not necessarily consecutive. There was no evidence of learning in the final stages of practice; typically, the final four practice trials were averaged for the "late" data.

Pilot parameters were identified for each subject in one of two ways. First, the full set of parameters was identified; this procedure is termed the "unconstrained" search procedure. Next, parameters were re-identified with the following constraints placed on the search scheme: (a) time delay fixed at an appropriate value for all subjects, both practice conditions; (b) motor noise ratio fixed at the average value for early and late practice for a given subject, and (c) observation noise/signal ratios constrained to be the same for all observation noise parameters identified (e.g., $PY_e = PY_{\dot{e}}$ for a given subject at a given level of practice). This procedure, which is termed the "constrained" search procedure, considers only the overall observation noise/signal ratio and the motor time constant to be influenced by training.

The constrained search procedure is justified because, for the data base considered in this study, time delay, motor noise, and differences among observational noise variables were not significantly influenced by practice. (One should not conclude that these model parameters were redundant or non-identifiable; rather, they were not indicators of the learning process.) As the reader will see, results are more readily digested if the search procedure is narrowed down to the important quantities.

---

* The search procedure actually identifies the relative cost weighting associated with control-rate variance. This weighting is then converted to an equivalent motor time constant for presentation.

Significance Testing

Two techniques were employed to determine the significance of practice-related parameter differences: paired-difference t-tests performed on the model parameters, and the qualitative cross-comparison scheme proposed by Levison [6,7]. Application of the t-test was the same as would be applied to the basic tracking data (say, mean-squared error scores), except that the identified pilot parameters served as the "data". As a practical matter, this procedure was limited to analysis of population means, with subject-paired early/late differences used in the computation of the "t" statistic.*

The cross-comparison technique was applied to both individual subjects and to subject populations. To perform this test, three sets of model parameters were identified for a given tracking task: (1) the set that best matched the early data, (2) the set that best matched the late data, and (3) the set that best matched the average of early and late performance. Four scalar matching errors were computed:

J(E,E) = matching error obtained from the early data, using parameters identified from the early data (i.e., best match to the early data).

J(E,A) = matching error obtained from the early data, using the average parameter set.

J(L,L) = best match to late data.

J(L,A) = matching error obtained from late data, using the average parameter set.

Finally, an average "matching error ratio" was computed as $[J(E,A)/J(E,E) + J(L,A)/J(L,L)]/2$.

The matching error ratio relates to the degradation in model-matching capability when an average set of parameters is used instead of the best-fitting parameters. If the average parameter set matches the data as well as the best-fitting set, then the matching error ratio is unity, and we accept the null hypothesis that there is no significant effect of practice on

---

* Significance testing for individual subjects would require parameter identification for a number of individual trials per subject, which would generally entail significant computational requirements.

pilot parameters. Conversely, a matching error ratio substantially greater than unity indicates that parameters must be changed across conditions to maintain minimum matching error, and we reject the null hypothesis.

This method can be used to test the parameter set as a whole, to test groupings of parameters, and to test individual parameters. Parameters were tested singly in this study. For each such test, remaining independent parameters were re-optimized to minimize matching error. Thus, this cross-comparison scheme provided a conservative test of significance in that it tested the matching errors that remained after the potential "tradeoffs" among parameters were accounted for.

Because of the complex nonlinear relationship between model parameters and matching errors, a quantitative transformation has not been found to relate matching error to the probability of the null hypothesis being either valid or invalid. Hence, the reference to this procedure as the "qualitative" test. As a matter of engineering judgement, a matching error ratio of 2.0 or more was considered in this study to reflect a qualitatively significant practice effect.

RESULTS

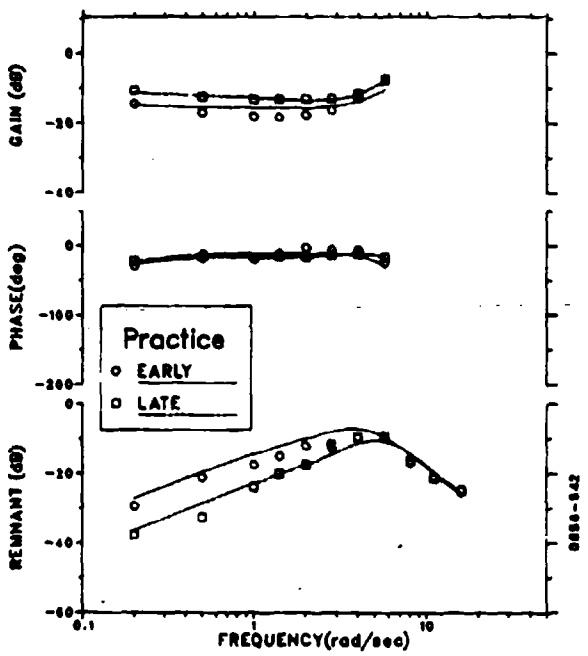
Effects of Practice on Pilot Response Behavior

Effects of practice on pilot describing functions are shown in Figure 1. Figure 1a shows results for a single subject from the static (fixed-base) group of motion-cue study, and Figure 1b shows results from a subject who trained initially with platform motion. Average results of six subjects are presented in Figure 1c. Discrete symbols indicate experimental data, whereas the smooth curves are model matches to the data.

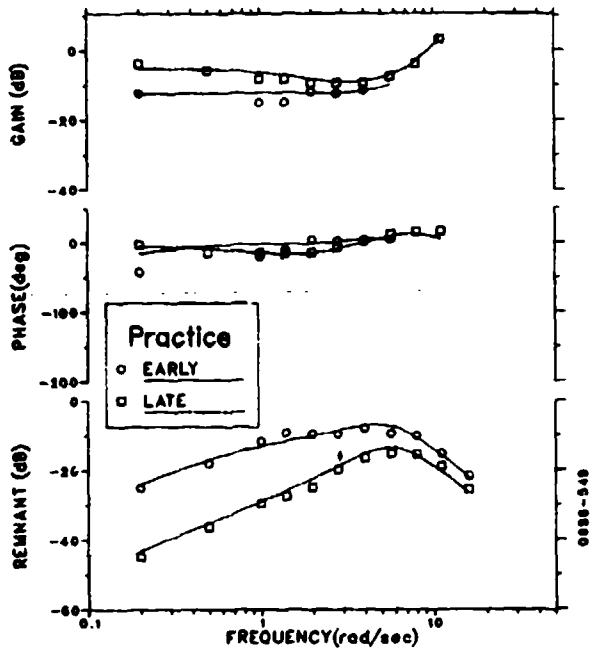
These curves, which are representative of the entire subject populations, show similar trends. Performance obtained early in practice, relative to asymptotic performance, is characterized by lower pilot gain, minimal differences in phase shift, and higher remnant.

Practice-related effects shown in these figures should not be considered simply characteristics of the pilot, but of the pilot/vehicle interaction. That is, one should not be surprised to find quantitatively (and perhaps qualitatively) different practice effects for tasks using substantially different vehicle dynamics. The goal of the subsequent model analysis is to relate

a) Subject VS, Static Group, Motion-Cue Study



b) Subject RK, Motion Group, Motion-Cue Study



c) Average of 6 Subjects, Performance Analyzer Study

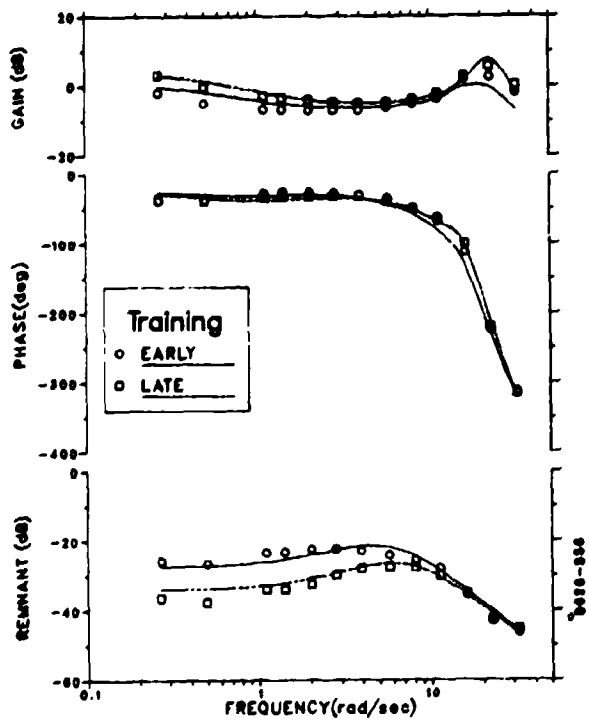


Figure 1.

Effect of Practice on Pilot Response Behavior

these changes to parameters that are more reflective of changes in pilot capabilities and relatively independent of the details of the tracking task.

Effects of Practice on Independent Model Parameters

Table 1 shows the effects of practice on average pilot parameters for the static and motion subject populations participating in the motion study. Results of the unconstrained and constrained search procedures are similar. Both observation noise and motor time constant decreased with practice for the two populations, with time constant differences being greater for the static group. Table 1a shows negligible changes in motor noise, an increase with practice in the time delay, and apparently greater practice effects on position-related noise as compared to rate-related noise. However, time delay changes and differences between position and rate noise were found to be largely insignificant, thereby justifying the constrained search reflected in Table 1b.

Results of the significance test of practice-related parameter differences are given in Table 2. The t-tests indicated that only observation noise differences are significantly influenced by practice; this was true for both the static and motion groups. The qualitative cross-comparison test confirmed this conclusion for the motion group. For the static group, however, changes in motor time constant, but not observation noise, were found to be significant.

The discrepancy between the conclusions yielded by the two significance testing methods may be due, in part, to appreciable subject-to-subject differences. Figure 2, which shows parameters identified for individual subjects, reveals an apparent negative correlation between practice-related effects on observation noise and motor time constant for the static group. That is, subjects exhibiting relatively large changes in motor time constant tended to exhibit relatively small changes in observation noise, and vice versa. This finding suggests some sort of tradeoff between decrements in motor time constant and decrements in observation noise in the early stages of practice. Because the parameter differences shown in Figure 2 were found to be largely significant (i.e., matching error ratio greater than 2), this presumed tradeoff does not simply reflect an insensitivity of the model-matching scheme. We explore shortly the possibility that this tradeoff reflects an insensitivity of performance to piloting strategy.

TABLE 1

Effects of Practice on Average Pilot Parameters, Motion Cue Study

a) Unconstrained Parameter Search

Motion Condition	Practice State	Statistic	PY_e	$PY_{\dot{e}}$	$PY_{\ddot{e}}$	PUP	TD	TN
STATIC	EARLY	MEAN	-10.5	-15.6	-	-41.9	.187	.205
		SD	3.1	1.9	-	11.1	.032	.084
	LATE	MEAN	-22.5	-18.7	-	-40.6	.221	.135
		SD	2.1	2.3	-	10.1	.044	.021
MOTION	EARLY	MEAN	-15.5	-14.9	-16.8	-55.4	.152	.115
		SD	8.2	1.3	4.6	7.9	.083	.006
	LATE	MEAN	-31.8	-14.6	-24.4	-56.8	.192	.099
		SD	0.4	6.6	1.6	5.6	.024	.024

b) Constrained Parameter Search

Motion Condition	Practice State	Statistic	PARAMETER PY	PARAMETER TN
STATIC	EARLY	MEAN	-15.2	.2021
		SD	1.4	.0826
	LATE	MEAN	-[illegible].9	.135
		SD	1.0	.022
MOTION	EARLY	MEAN	-15.4	.124
		SD	3.4	.015
	LATE	MEAN	-24.1	.099
		SD	0.5	.022

Statistics for 5 subjects, static; 4 subjects, motion; 2-4 trials/subject.

TABLE 2

Significance Test of Practice-Related Parameter Differences, Motion Cue Study

Motion Condition	Search Procedure	Test Procedure	Parameter PY_e	$PY_{\dot{e}}$	$PY_{\ddot{e}}$	PUP	TD	TN
STATIC	Unconstrained	t-test	*	-	N/A	-	-	-
	Constrained	t-test	*		N/A	N/A	N/A	-
	Constrained	Qualitative	-		N/A	N/A	N/A	*
MOTION	Unconstrained	t-test	-	-	*	-	-	-
	Constrained	t-test	*			N/A	N/A	-
	Constrained	Qualitative	*			N/A	N/A	-

N/A = not applicable.

- = not significant.

* = alpha significance level 0.05 or less for the t-test, matching error ratio greater than 2.0 for the qualitative test.

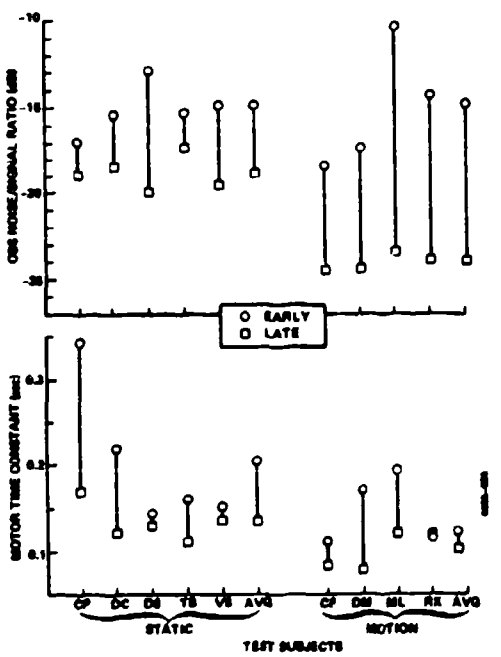


Figure 2. Effects of Practice on Pilot-Related Model Parameters, Motion Cue Study

Results of the model analysis were more consistent for the performance analyzer study. Table 3 shows that practice influenced primarily the observation noise/signal ratio for the

subjects participating in this study. Both methods for determining significance indicated that only the noise parameters were significantly effected. On the average, motor time constants and observation noise/signal ratios were lower than those obtained in the motion-cue study, perhaps because of the more severe information-processing requirements imposed by the unstable plant dynamics used in the performance analyzer study.

Sensitivity Analysis

The process of developing a model for control-strategy learning would be simplified (or, at least, easier to discuss) if we could relate the process of skill acquisition to changes in a single independent model parameter. As suggested above, the apparent negative correlation between changes in time constant and changes in observation noise may reflect a certain insensitivity of performance to variations in pilot strategy. That is, what we measure as a difference in motor time constant may not reflect a change in the pilot's response bandwidth limitation, but rather that, early in practice, the subject deviates from optimality in such a way as to give the appearance of a degraded response bandwidth.

In the interest of developing the most parsimonious model possible, let us postulate the following:

1. The only "real" effect of practice on "pilot-related" model parameters is to reduce the observation noise/signal ratio from an initially high level to a substantially lower-level.

2. The presence of a high level of observation noise affords the pilot a certain amount of flexibility that is absent when noise levels are relatively low. Specifically, he has the option of reducing his gain below would what be predicted as "optimal" in order to trade control activity for tracking error, while maintaining a near-optimal performance level (where "performance" is a weighted sum of mean-squared error and control-rate).

In order to test this hypothesis, various sets of "data" were generated by the model and then reanalyzed to see if experimental trends could be replicated. First, a motor time constant of 0.1 seconds was specified, and the observation noise ratio was adjusted (equal for position and rate noise) to yield the mean-squared error score found experimentally on the average for static tracking early in practice. Nominal values were

TABLE 3

Effects of Practice on Pilot-Related Model Parameters, Performance Analyzer Study

a) Model Parameters

Practicing State	Statistic	Parameter PY_e	$PY_{\dot{e}}$	PUP	TD	TN
EARLY	MEAN	-18.0	-20.0	-44.2	.141	.105
	SD	2.3	2.4	9.3	.032	.029
LATE	MEAN	-24.2	-22.2	-42.2	.135	.0809
	SD	1.4	0.7	7.3	.013	.0145
EARLY	MEAN *	-19.5		-45.9	.135	.0863
LATE	MEAN *	-22.4				.0786

* Serarch performed under following constraints: $PY_e = PY_{\dot{e}}$, PUP = -45, TD = .135.

Statistics for 6 subjects, 2-4 trials/subject.

b) Significance Tests

Analysis Procedures	Parameter PY_e	$PY_{\dot{e}}$	PUP	TD	TN
Unconstrained search, t-test	*	-	-	-	-
Constrained search, qualitative test	*		N/A	N/A	-

N/A = not applicable.

\- = not significant.

* = alpha significance level .05 or less for the t-test, matching error ratio greater than 2.0 for the qualitative test.

Statistics for 6 subjects, 2-4 trials/subject

specified for remaining model parameters. Next, certain manipulations were performed to force the pilot model to modify its tracking strategy, subject to the constraint that total performance (a weighted sum of error and control-rate variances) not be materially increased. This procedure was repeated, with observation noise adjusted to reproduce tracking error variance appropriate to static tracking late in practice. Finally, these model-generated "data" were then analyzed in the same manner as the tracking data for the purposes of identifying independent model parameters.

Analysis of these pseudo-data tended to reflect the trends of the experimental data. Specifically, motor time constants of approximately 0.2 and 0.15 seconds were identified for "data" corresponding to early and late practice stages, respectively, and observation noise ratios associated with perception of error displacement were somewhat greater than noise ratios associated with rate perception.

Thus, there appears to be some support for the notion that apparent practice-related changes in motor time constant are due in part, if not entirely, to certain tradeoffs in control strategy that are available to the pilots when performance is limited by high observation noise levels. Additional work is needed to validate this hypothesis; for example, to show that it also accounts for results obtained in the motion case and in the performance analyzer study.

DISCUSSION

This study was, in part, an exploration of analysis methodologies as well as a study of practice (or learning) effects. We discuss methodological considerations first, then proceed to the discussion of the results of primary interest.

Methodological Considerations

Within the framework of the quasi-Newton search procedure, two techniques for identifying independent model parameters were explored: an unconstrained search in which all relevant independent model parameters are identified, and a constrained search in which certain parameters are fixed and others forced to conform to certain interrelationships. Both procedures are useful and, in fact, needed in most instances. In general, one must first perform an unconstrained search to determine which parameters are (and are not) influenced by the experimental

variable(s) of interest. After parameters have been identified in this manner and tested for significance, one can justifiably impose constraints on the search procedure. A constrained search then allows one to focus on the important effects and, as is clear from a comparison of Tables 1a and 1b, it greatly facilitates presentation of the results of the analysis.

Two procedures were also explored for testing the statistical significance of parameter differences: the standard t-test, and a sensitivity analysis that we have termed the "qualitative cross-comparison" procedure. Both methods yielded the same conclusions in situations where practice effects were consistent across subjects. Conclusions were different when subject behavior was inconsistent; but, since we have no way of determining which method is "right", it is prudent to retain both techniques in our repertoire of analysis tools.

To some extent, the nature of the data base will determine which testing procedure to use. Since the t-test requires a number of replications per condition, the qualitative testing procedure must be used to compare parameters identified from two experimental trials (as might be desired if significant trial-to-trial learning occurs). The qualitative procedure can also be used to minimize computational requirements when multiple trials per condition are available, as this method allows one to directly test the replicate-averaged data.

Either method may be used for testing average results for a subject population. If parameters are identified for individual subjects first (a recommended procedure if substantial subject differences are expected), applying the t-test to subject-paired parameter differences avoids the necessity for further model analysis. If, on the other hand, parameters have been identified only for subject-averaged data, the cross-comparison scheme must be used, as there is only one set of model parameters per experimental condition.

Practice Effects

There are a number of reasons for attempting to characterize the effects of practice in terms of a single independent model parameter. First, it is this author's opinion that, in general, the model having the greatest potential for predictive capability is the most parsimonious model -- i.e., the model that explains the greatest amount of experimental data with the fewest independent parameters. Reduction of practice effects to changes in a single independent parameter does not necessarily impose limitations on modelling the complexity of the learning process.

We have shown that, at the very least, pilot gain and remnant are influenced by practice; one may reasonably suspect that more subtle aspects of the operator's performance capabilities -- such as the accuracy and precision of his internal model of the task structure and requirements -- are also modified during the learning process. Presumably, these factors can be treated as dependent model parameters -- i.e., aspects of behavior that can be predicted and need not be pre-specified.

Another reason for focussing on a particular model parameter is that only one parameter -- observation noise -- was consistently influenced in a significant manner by practice for the data base explored in this study. Changes in motor time constant were much less consistent, varying from subject-to-subject. Hence the hypothesis, at least partially supported by sensitivity analysis, that changes in motor time constant reflect a relative insensitivity of overall closed-loop performance to changes in piloting strategy in the high (pilot) noise environment characteristic of early stages of practice.

Finally, practice-related changes in observation noise/signal ratio make more intuitive sense than practice-related changes in motor time constant.* Recall that observation noise is the mathematical device by which we account for most of the stochastic portion of the pilot's response ("pilot remnant"). Thus, observation noise reflects time variations and nonlinearities in the pilot's response strategy as well as injected noise. It is reasonable to assume that, with continued practice, the operator varies his response strategy less during the course of an experimental trial, develops a more stable internal model of the task structure, and learns to respond in a more linear fashion.

Analysis of various data bases suggests that the motor time constant is not a truly independent model parameter in the sense of representing an inherent human operator information-processing limitation. It appears to reflect, in part, the operator's adaptation to the task environment. For example, this parameter appears to vary in a systematic manner with the bandwidth of the plant dynamics [7]. In addition, analysis of the data obtained from the motion-cue study showed that the motor time constant decreased almost immediately by about one-third when the

---

* One might reasonably expect time delay to decrease with practice. However, practice-related changes in this model parameter were inconsistent and not statistically significant.

subjects, trained initially fixed-base, were provided with concurrent visual and motion cues. It is highly unlikely that an inherent inability or reluctance to generate large rates of change of control would suddenly be modified.

These results lead to an alternative hypothesis concerning variations in motor time constant; specifically, that the motor time constant reflects, in part, limitations on the operator's ability to construct an internal model of the task environment. As information to the operator is improved, either by extending system bandwidth or by enhancing perceptual cueing, the more accurate and precise the internal model, and the lower the motor time constant.

As a final comment, the reader is reminded that the attempt to characterize the learning process in terms of variations in one or two independent model parameters is strictly a short-term research objective that is intended to provide some insights into the learning process and to point the way to further research. The long-term goal is to develop a model that will allow one to predict the effects of task environment -- especially perceptual cueing -- on the rate at which the operator learns appropriate estimation and control strategies. It is anticipated that the next phase of this study program will be to explore hypotheses regarding the operator's ability to construct his internal model of the task environment.

REFERENCES

1. Baron, S. and Levison, W.H., "The Optimal Control Model: Status and Future Direction", Proc. of the International Conf. on Cybernetics and Society, Cambridge, MA, October 1980, pp. 90-101.

2. Levison, W.H., "The Optimal Control Model for the Human Operator: Theory, Validation, and Application", Proc. of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics, Edwards Air Force Base, CA, January 1982.

3. Levison, W.H., Lancraft, R.E. and Junker, A.M., "Effects of Simulator Delays on Performance and Learning in a Roll-Axis Tracking Task", Proc. of the Fifteenth Annual Conference on Manual Control, Wright State University, Dayton, OH, March 1979.

4. Zacharias, G.L. and Levison, W.H., "A Performance Analyzer for Identifying Changes in Human Operator

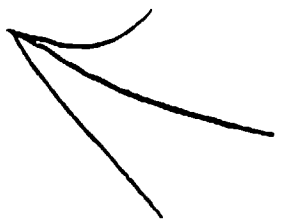
Tracking Strategies", AMRL-TR-79-17, Aerospace Medical Research Labortory, Wright-Patterson Air Force Base, March 1979.

5. Lancraft, R.E. and Kleinman, D.L., "On the Identification of Parameters in the Optimal Control Model", Proc. of the Fifteenth Annual Conference on Manual Control, Wright State University, Dayton, OH, March 1979.

6. Levison, W.H., "A Quasi-Newton Procedure for Identifying Pilot-Related Parameters of the Optimal Control Model", Proc. of the Seventeenth Annual Conference on Manual Control, Los Angeles, CA, June 1981.

7. Levison, W.H., "Effects of Whole-Body Motion Simulation on Flight Skill Development", Report No. 4645, Bolt Beranek and Newman, Inc., Cambridge, MA, November 1981 (ADA 111-115).

8. Levison, W.H., Baron, S. and Junker, A.M., "Modelling the Effects of Environmental Factors on Human Control and Information Processing", AMRL-TR-76-74, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, OH, August 1976.



AD P001630

A MODERN APPROACH TO PILOT VEHICLE ANALYSIS AND THE NEAL-SMITH CRITERIA

Barton J. Bacon & David K. Schmidt
School of Aeronautics and Astronautics
Purdue University
West Lafayette, IN 47907

Extended Abstract

In the past, the longitudinal handling qualities of an aircraft were determined almost entirely by the modal characteristics of the classical rigid body modes (short period and phugoid). These modes, which dominate conventional aircraft's dynamics and their modal parameters (i.e., damping and natural frequency), exhibit a definite correlation with pilot opinion ratings. Unfortunately, beyond the realm of conventional aircraft, criteria based on these parameters alone are inadequate. The addition of other modes, whether due to structural dynamics or to augmentation, has been shown to seriously affect pilot opinion rating.

In the early 70's, Neal and Smith [1] sought a solution to this high-order system problem by "pilot-in-the-loop-analysis". Reviewing pilot comments, the Neal-Smith investigation drew on some important hypotheses: "Pilot rating is a strong function of the pilot compensation required to achieve good low frequency performance and the oscillatory tendencies of the pilot-vehicle system." They developed a closed-loop methodology to extract measures of the desired quantities (i.e., pilot compensation, oscilatory tendencies), and these measures were shown to correlate with pilot ratings of the 57 configurations flight tested. The measures are the resonant peak of the closed-loop pilot-vehicle frequency response, and the pilot's phase compensation at the closed-loop bandwith (see Figure 1).

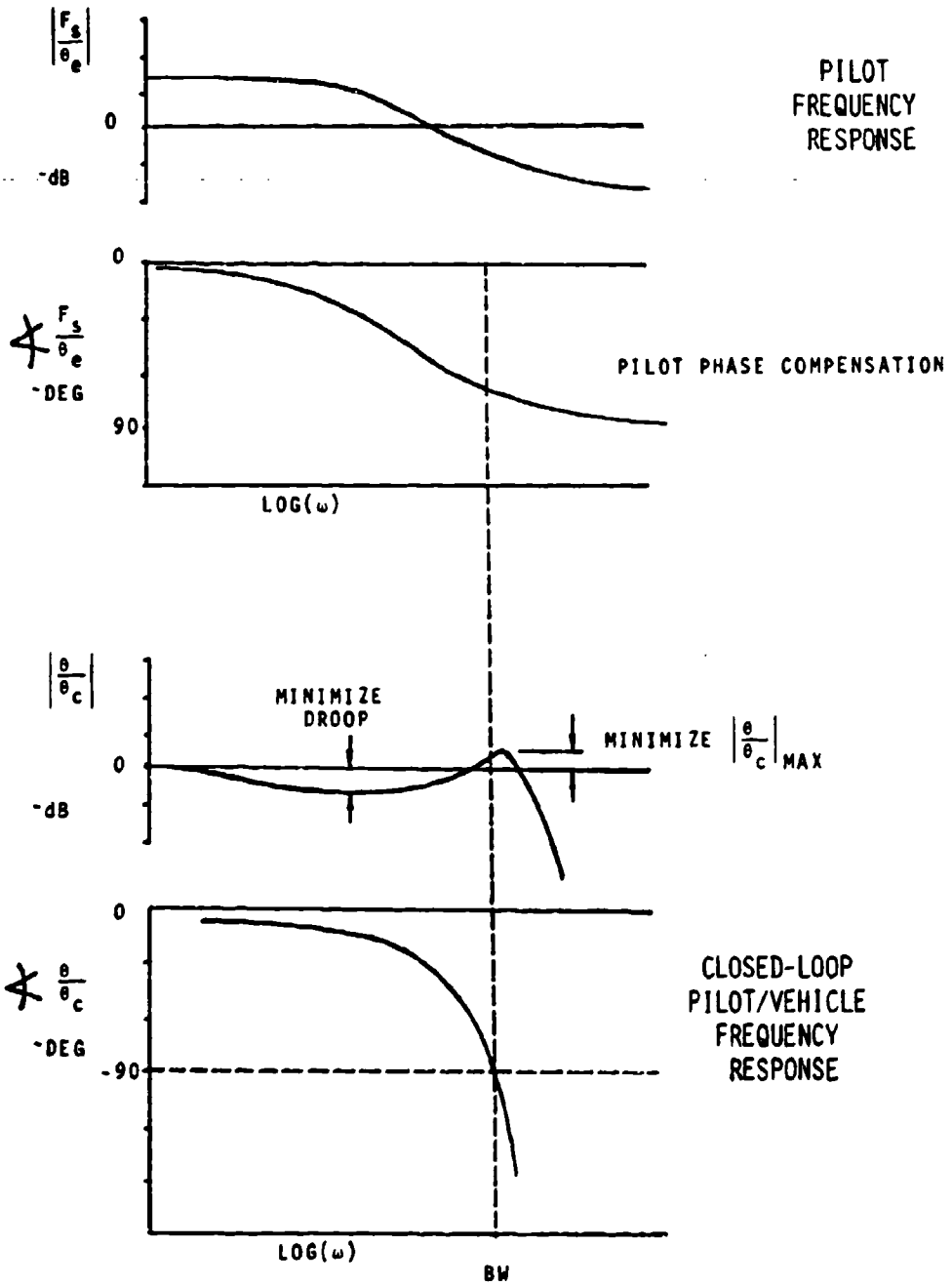

$\left|\frac{F_s}{\theta_e}\right|$
0
-dB
PILOT
FREQUENCY
RESPONSE
0
$\angle \frac{F_s}{\theta_e}$
-DEG
90
PILOT PHASE COMPENSATION
LOG(ω)
$\left|\frac{\theta}{\theta_c}\right|$
MINIMIZE
DROOP
MINIMIZE $\left|\frac{\theta}{\theta_c}\right|_{MAX}$
0
-dB
0
$\angle \frac{\theta}{\theta_c}$
-DEG
-90
CLOSED-LOOP
PILOT/VEHICLE
FREQUENCY
RESPONSE
LOG(ω)
BW

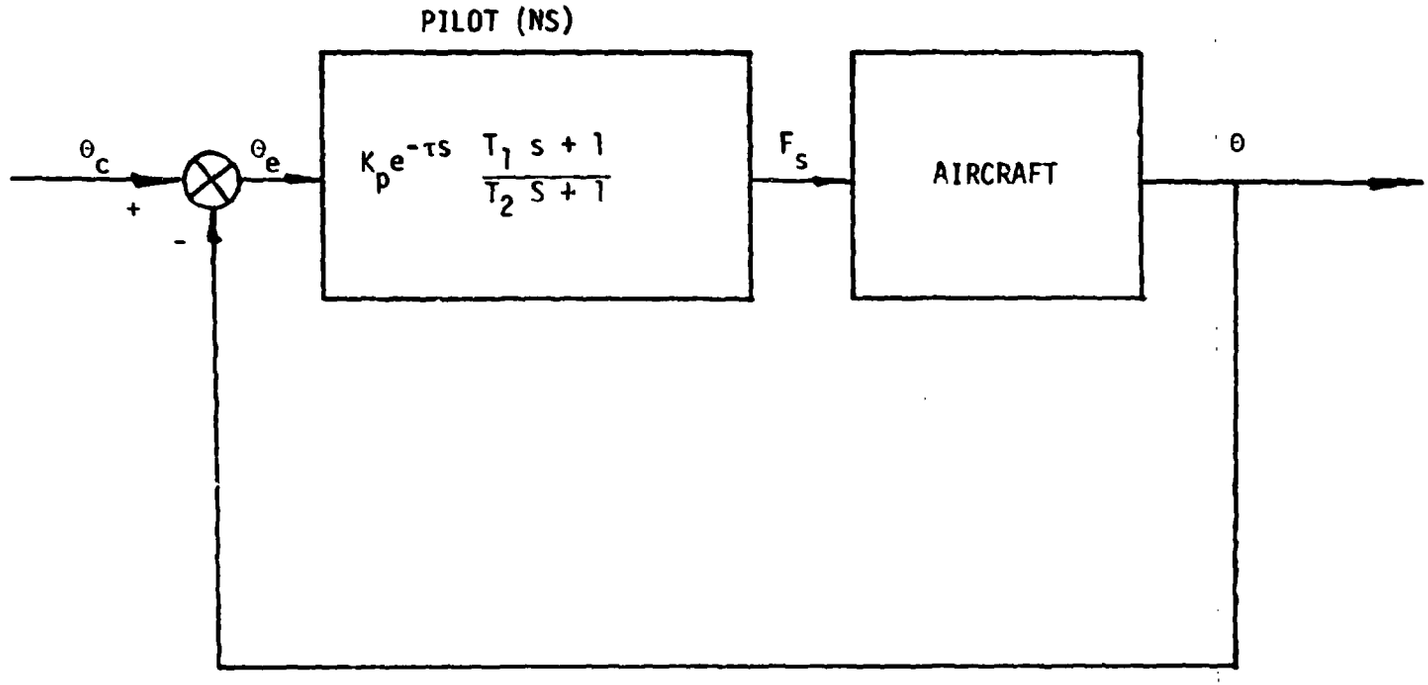
Figure 1

Unfortunately, the methodology suffers from difficulty in obtaining the parameters of a pilot describing function, taken to be a simple lead-lag with transport delay and feed forward gain. (see Fig. 2). These parameters are chosen such that certain closed-looped objectives, set forth by Neal and Smith, are met by this model of compensatory pitch-attitude tracking task. Choosing these parameters poses the problem. The algorithm can often become a graphical nightmare and lead to non unique results. However, the overall hypothesis has merit, and is considered by many handling qualities researchers as a most promising technique. So extending the method to yield unique results in a straight forward fashion would appear most useful. This is the subject of this paper.

Therefore, the goal is to present an alternate method for finding the quantitative handling quality parameters recognized as important by the Neal-Smith study. While further confirming the Neal-Smith hypothesis, this paper develops a methodology that expresses directly the pilot's objective in attitude tracking. More importantly, the method relieves the researcher of the tedious task of finding pilot model parameters, and in fact yields a better model that may be applied directly to closed-loop criterion.

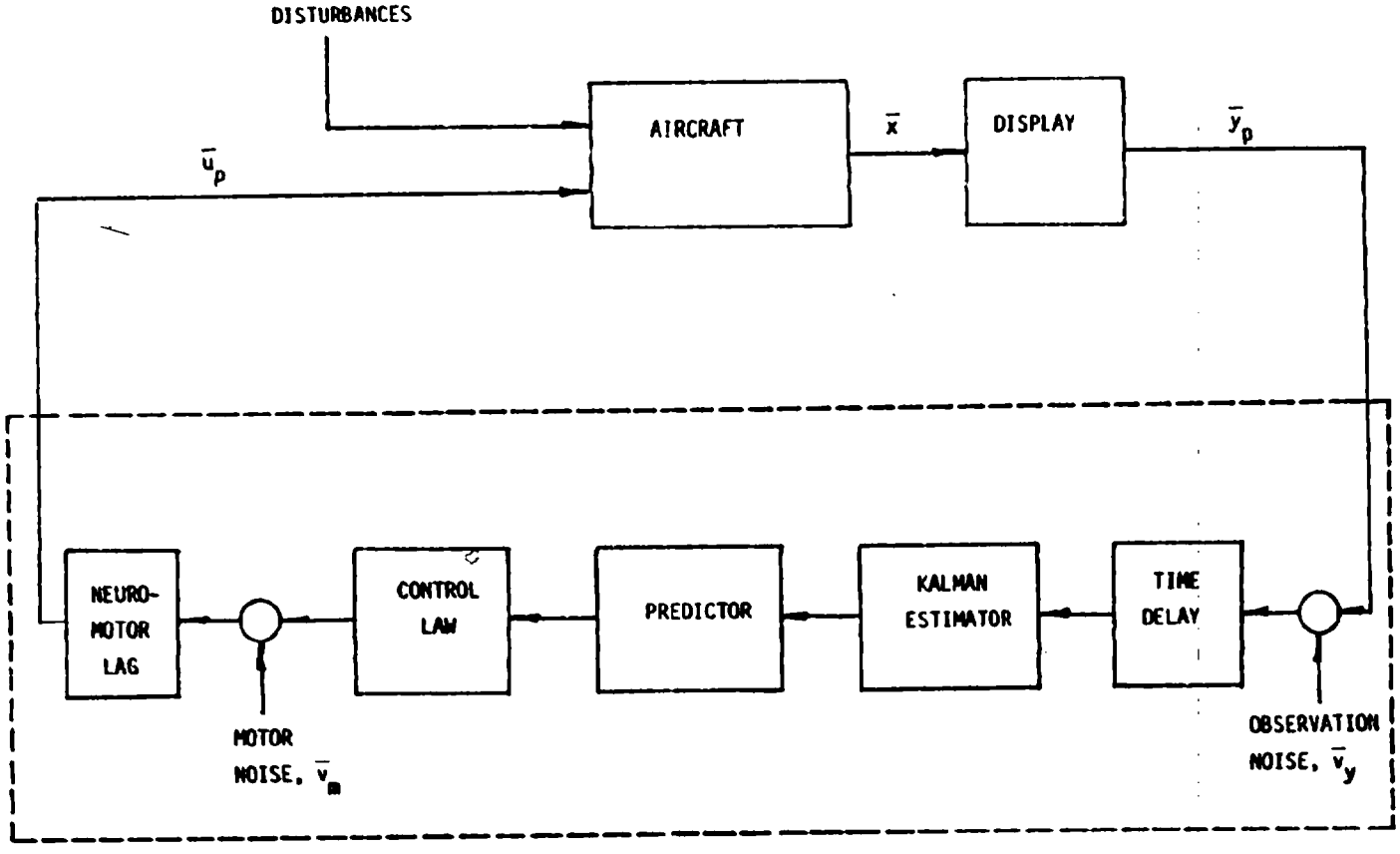
In 1970, Kleinman, Barron, and Levison [2] published a mathematical model of human response using optimal control and estimation theory. We will show that by this modeling approach we obtain a better pilot model that is easier to obtain, and in so doing incorporate multivariable, time-domain techniques directly with the classical approaches, and specifically the Neal-Smith method.

Use of the optimal control model (OCM) (see Fig. 3) reflects more correctly the actual experimental situation. The OCM incorporates the



CLASSICAL MODEL STRUCTURE

Figure 2



OCM SCHEMATIC

Figure 3

visual threshold effects, as well as time delay and neuromuscular lag associated with the human pilot. More importantly, the OCM can be used to predict complex pilot compensation likely to be present in the control of high-order dynamics, in contrast to being restricted to only lead-lag compensation. Finally, the objective of the pilot in the tracking task had to be inferred indirectly in the analysis of Neal and Smith, but is expressed directly in the objective function of the OCM. Specifically, the objective function found (from previous research) to be representative of a pilot's in the task may be expressed as

$$J(F_s) = \frac{1}{T}\int_0^T (16\,(\theta_c - \theta)^2 + (\dot{\theta}_c - \dot{\theta}) + g(\dot{F}_s)^2)dt$$

where $(\theta_c - \theta)$ is the attitude error, and F_s is the stick force input of the pilot.

Admittedly, the OCM block diagram varies drastically from the classical, single loop structure. However, the OCM may still be applied to estimate the pilot's equalization in a compensatory tracking task. To substantiate this fact, we have used the OCM to estimate spectral densities of the system's signals (e.g. plant and pilot input and outputs) and then compared tracking errors from both model structures. Integrating the spectral density over the frequency domain gives the variance of the signals in question by applying the following relation.

$$R_x(o) = \frac{1}{\pi}\int_o^\infty S_x(\omega)\, d\omega = \sigma_x^2$$

If the compensatory describing function is accurately estimated via the OCM we should obtain comparable error and state variances. And indeed the estimated compensatory tracking describing function produced tracking errors within 4% of the OCM.

By validating that the OCM can be used to estimate the compensatory tracking task describing function, one has the flexibility to study the pilot frequency response in the same framework used in the Neal-Smith report. Hence, both open-looped and closed-looped frequency responses are available along with the pilot frequency response. Also, the desired measures of pilot phase compensation and closed-loop resonance peak may be directly extracted from the OCM. And finally, the link has been made between the frequency domain and the time domain and multivariable systems techniques.

Complete results are presented in Reference [3].

References

[1] T.P. Neal and R.E. Smith, An In-flight Investigation to Develop Control System Design Criteria for Fighter Airplanes. AFFDL-TR-70-74, Vol. I, December (1970).

[2] D.L. Kleinman, S. Baron and W.H. Levison, "An Optimal Control Model of Human Response Part I: Theory and Validation". Automatica, Vol. 6, pp. 357-369, Pergamon Press, (1970).

[3] Bacon, B.J., and Schmidt, D.K., "A Modern Approach to Pilot/Vehicle Analysis and the Neal-Smith Criteria", AIAA paper no. 82-1357, 1982 AIAA Atmospheric Flight Mechanics Conf., San Diego, CA.

A STRUCTURED METHODOLOGY FOR ANALYZING HUMAN INFORMATION PROCESSING IN COMMAND SYSTEMS

By

Joseph G. Wohl, Krishna R. Pattipati, David L. Kleinman, Nils R. Sandell, Jr., and Elliot E. Entin*

ALPHATECH, Inc.
3 New England Executive Park
Burlington, MA 01803

ABSTRACT

Human operators play critical information processing and decisionmaking roles in command, control and communication (C^3) systems. These systems are characterized by extreme complexity, and functional and geographic diversity. In this paper, we present a methodology, based on the concept of hierarchical decomposition, for the quantitative analysis and design of such systems. Our approach begins with an analytic decomposition of the C^2 process into functions, along with the concomitant association of objectives to functions. The functions, along with the associated objectives, in turn, are subdivided into well-defined entities or tasks that are amenable to analytic modeling and/or empirical description. The task level analysis is integrated into an evaluation of specific C^2 functions. The existing modeling technology for task analysis and integration is discussed, along with the research needed to augment this technology base to arrive at a unified methodology for the description, analysis and, ultimately, the design of C^3 systems.

---

*Joseph Wohl is the Vice-President of Research and Development; Krishna Pattipati and Elliot Entin are members of the Technical Staff; David Kleinman is a consultant; and Nils Sandell, Jr., is the President of ALPHATECH, Inc.

SESSION 4: MODEL BASED DESIGN AND CONTROL INTEGRATION

Moderator: Sheldon Baron, Bolt Beranek and Newman, Inc.

4

Visual Cue Requirements for Flight Control Tasks

Ronald A. Hess
Ames Research Center
Moffett Field, California

and

Arthur Beckman
University of Southern Colorado
Pueblo, Colorado

ABSTRACT

A method for determining basic visual cue requirements for piloted flight control tasks is presented. The method emphasizes a classical, loop-by-loop synthesis procedure for determining required pilot equalization and open-loop crossover frequencies in well-defined flight control tasks. The results of recent pertinent laboratory tracking experiments are discussed involving measurement of pilot dynamics in single and multiple-loop tasks. The role of pursuit and preview tracking in the inner attitude loops of multiple-loop tasks is outlined and their importance in determining visual cue and display requirements is discussed. It is finally shown that flight-director laws evolve quite naturally via this procedure.

AD P001631

VALIDATION OF AN ADVANCED COCKPIT DISPLAY DESIGN METHODOLOGY VIA WORKLOAD/MONITORING TRADEOFF ANALYSIS

By

Dr. Jonathan Korn
Dr. Sol W. Gully
ALPHATECH, Inc.
3 New England Executive Park
Burlington, MA 01803
And
Dr. David L. Kleinman
University of Connecticut
Storrs, CT

ABSTRACT

Validation of a model-based methodology for the design and evaluation of advanced cockpit information-display systems is undertaken. The technique is applied to determine information and display requirements in an A10 terrain-following flying task. Four candidate display systems, including a flight director system, are proposed and rank-ordered across dimensions of workload and performance. Validation of the analytical predictions is accomplished through man-in the-loop simulation experiments. It is concluded that the methodology, which can determine the limit information to the best quantities needed by the pilot to perform effectively, can be a valuable tool in the development of advanced cockpit information systems.

INTRODUCTION

The increased complexity of advanced tactical air-to-air/close air support aircraft has elevated the pilot to a role of systems manager. The pilot must allocate sensors, evaluate threats, select weapons, employ ECM, etc., in addition to the omnipresent task of flying the aircraft. The increase in the number of systems and subsystems has required the display of increasingly more information to the pilot about his aircraft and the environment. Unfortunately, this has increased the pilot's workload. A methodology, therefore, is required to determine and limit information to the best informational set needed by the pilot to perform his various tasks effectively.

If pilot and aircraft are to function as an effective combat system, it is essential that the workload associated with the basic flying task be reduced through either display or control automation. Thus, the pilot would be in a

* This work performed under the following contract to AMRL: F33615-80-C-0528

position where he can allocate a greater fraction of his capacity for decision making relative to weapon assignment, etc. Nowhere is this problem more critical than in high-speed, low-level, adverse-weather interdiction where the workload demands of the flying tasks are severe.

The design of automated control/display systems for reducing workload in piloting tasks is generally accomplished through extensive recourse to man-in-the loop simulations. This can be an expensive and time-consuming approach, especially when large numbers of competing designs and/or parameter sets must be evaluated and iterated. While simulation is ultimately necessary in the development of control/display systems, it would be desirable to have a computer-based tool that could perform a preliminary evaluation on a wide variety of systems on a relative performance basis. In this manner, one would only need to retain those systems exhibiting promise for further evaluation by manned simulation.

A model/computer-based methodology has recently been proposed [1-2] that offers considerable potential for application to tactical aircraft display systems design and evaluation. Although this methodology has been applied to rank-order control/display systems for a CH-47[1], the design technique has remained largely unvalidated due to the lack of subsequent manned simulation to assess the preliminary analytic results. In this study the design procedure, as proposed by Kleinman and Curry, is extended to a practical cockpit scenario. Subsequently, we validate the model predictions via man-in-the-loop simulation experiments. The piloting control task considered is a representation of a zero-visibility, high-speed, low-level terrain following scenario with an A10 aircraft. The aircraft dynamics are approximated by their short-period longitudinal equations in the analysis and simulation.

The focus of this study is on display, design, and evaluation, including optimal synthesis or aggregation of information states, for the pilot control task. The design methodology, when applied to display systems evaluation, follows a three-level procedure. At the first, or information level, the methodology determines the relative importance of each system variable to closed-loop task performance and determines the optimal synthesis of information in the contex of a flight director signal. At the second, or display element level, the information requirements are integrated to propose several different realistic display systems. Human generated information processing limitations are included at this level, and for each candidate design, performance versus workload curves are generated. The different display systems are then rank-ordered across dimensions of performance and workload. The proposed displays that resulted via the application of the methodology to the terrain following scenario included: (1) a terrain predictor/flight path vector display, (2) a tunnel display, (3) an integrated tunnel-predictor display, and (4) an optimally designed flight director display.

The display format level is the third and final phase of the design process. Here, specific display formats are suggested for the presentation of the candidate display systems, guided by the sensitivities, attentional demands, etc., that are predicted at the display element level. The format level is the precursor to manned simulation experiments. In this effort, a man-in-the-loop

experimental program was conducted at the University of Connecticut to evaluate the four candidate display systems. The objective of these experiments was to validate the overall display design procedure, including the analytic rank-ordering, the workload assessment, and the flight director design synthesis processes.

The analytical nucleus to the display design methodology is the Optimal Control pilot Model (OCM) [3]-[5]. A detailed description of the OCM, in conjunction with the three-level display design process, is given in [6]. In this paper we first introduce the control task of interest, including the A10 aircraft dynamics and the terrian profile characteristics. We apply then the display design procedure, which yields the major analytical results of this study, including the predictions of attentional allocation, workload demands, and performance rank-ordering. Finally, the experimental program is described, and the experimental results discussed.

APPLICATION OF THE CONTROL TASK TO THE OCM

A-10 LONGITUDINAL AIRCRAFT DYNAMICS

As indicated, the control task of interest is high-speed terrain following for a representative low-level attack flight condition. The basic set of longitudinal equations being used is the short-period dynamics. These are the perturbation equations written about straight and level flight, and they describe the aircraft longitudinal rotations about its center of mass.

The (two-degree of freedom) transfer functions of interest are [7]

$$\frac{\alpha}{\delta}(s) = \frac{1}{U_o}\,\frac{Z_\delta s + (U_o M_\delta - Z_\delta M_q)}{\Delta(s)} \tag{1}$$

$$\frac{q}{\delta}(s) = \frac{(M_\delta + Z_\delta M_{\dot w})s + (Z_\delta M_w - M_\delta Z_w)}{\Delta(s)} \tag{2}$$

$$\Delta(s) = s^2 - (U_o M_{\dot w} + Z_w + M_q)s + (M_q Z_w - U_o M_w) \tag{3}$$

where q = aircraft pitch rate, α = angle of attack, U_o = nominal air speed, δ = elevator angle, and Z_δ, M_δ, M_q, $M_{\dot w}$, Z_w are the pertinent stability derivatives.

In the present work the stability derivatives were derived from data currently used on the Aerospace Systems Division A-10 nonlinear hybrid simulation, and, the numerical values are given in [6]. The nominal air speed U_o = 468 ft/sec.

In addition, the aircraft pitch angle, θ, and its altitude, h, are necessary in the modeling process. Therefore

$$\dot\theta = q \quad ; \quad \dot h = U_o(\theta - q) \tag{4}$$

TERRAIN PATH MODEL

The next step is to select an appropriate terrain path model that serves as the excitation "signal" to the system. An easily implemented one dimensional model for the terrain height Π as a function of time consists of a Markov process passed through a 2nd order filter*[6]. The Markov process is generated by

$$\dot{z}(t) + \frac{2U_o}{\lambda} z(t) = \frac{2U_o}{\lambda} \xi(t) \tag{5}$$

where $\xi(t)$ is a white Gaussian noise, and λ represents the spatial terrain variations. Next, $z(t)$ is filtered through

$$H(s) = \frac{\omega^2 \eta_o^2}{s^2+2\zeta\omega U_o s+\omega^2 U_o^2} , \tag{6}$$

where $\omega = 2\pi/D$ is the natural frequency of the terrain (D = "terrain period") and ζ is the damping ratio. In order to represent a realistic terrain profile, the terrain parameters are selected appropriately [6]. The parameter values chosen to give the most appealing terrain characteristics where

$$D = 10000' \quad ; \quad \lambda = D/\pi \quad ; \quad \zeta = 1/\sqrt{2}$$

The white noise ($\xi(t)$) intensity was chosen to yield a reasonable terrain RMS value of ~136 ft.

OCM APPLICATION

The aircraft dynamics and the terrain states are combined into a single state equation as required by the OCM. The augmented state is defined as

$$x' = [\Pi \ \dot{\Pi} \ \ddot{\Pi} \ \alpha \ q \ \theta \ e] \tag{7}$$

where $e(t) \triangleq$ flight path error $= \Pi(t) - h(t)$. The total state space equation is now

$$\dot{x}(t) = Ax(t) + B\delta(t) + E\xi(t) \tag{8}$$

---

*The term Π should represent a spatial rather than temporal terrain model. However, assuming a constant velocity, U_o, and neglecting the flight path flactuations, we can write $\Pi(r) \simeq \Pi(U_o t) = U_o \Pi(t)$, where $r \simeq U_o t$ represents the traveled distance.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{-2U_o^3\omega^2}{\lambda} & -\left(\frac{4\zeta\omega}{\lambda}+\omega^2\right)U_o^2 & -\left(2\zeta\omega+\frac{2}{\lambda}\right)U_o & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Z_w & 1 & 0 & 0 \\ 0 & 0 & 0 & U_o(Z_wM_w'+M_w) & U_o(M_w'+M_q) & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{\pi}{180^\circ}U_o & 0 & \frac{-\pi}{180^\circ}U_o & 0 \end{bmatrix}$$

$$B' = \begin{bmatrix} 0 & 0 & 0 & \frac{Z_\delta}{U_o} & M_\delta+Z_\delta M_w' & 0 & 0 \end{bmatrix}$$

$$E' = \begin{bmatrix} 0 & 0 & \frac{2U_o^3\omega^2}{\lambda} & 0 & 0 & 0 & 0 \end{bmatrix}$$

The numerical values of A, B, E are:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -.025 & -.21 & -.71 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1.37 & 1 & 0 & 0 \\ 0 & 0 & 0 & -3.08 & -1.68 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 8.17 & 0 & -8.17 & 0 \end{bmatrix}$$

$$B' = \begin{bmatrix} 0 & 0 & 0 & -0.06 & -6.84 & 0 & 0 \end{bmatrix}$$

$$E' = \begin{bmatrix} 0 & 0 & .025 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The terrain following task requires the pilot to follow the path-over-terrain $\bar{H}(t)$, as closely as he can, subject to his inherent limitations. Thus, the quantity he should minimize is the flight path error (FPE) $e(t)$. In addition, the pilot should avoid large vertical acceleration values, or g's. These requirements on $e(t)$ and $g(t)$ (or, equivalently on $q(t)$) are expressed in the cost functional associated with the OCM, with the pilot's subjective weightings q_e and q_q, viz.,

$$J(\delta,f) = E\{q_e e^2(t) + q_q q^2(t) + q_{\dot\delta}\dot\delta^2(t)\} \quad , \tag{9}$$

where the control rate weighting, $q_{\dot\delta}$, is determined by the pilot's neuromuscular time constant τ_N [3]-[6]. The cost functional $J(\delta,f)$ is minimized with respect to the control δ and the attentional allocation vector f.

The nominal altitude of the trajectory above the terrain ($\Pi(t)$) the aircraft must follow is 200 ft., and we assume that the maximal excursion, or FPE, should be no more than 20%, or $e_{max} = 40$ ft. Also, we assume that the maximal vertical acceleration to be tolerated by the pilot is $a_{z,max} = 1g = 32.2$ ft/sec^2. Thus, we compute the FPE and pitch rate weighting according to [1],[2]-[6], viz.,

$$q_{max} \cong \frac{180°}{\pi} \cdot \frac{a_{z,max}}{U_o} \cong 4°/sec$$

$$q_e = \frac{1}{e_{max}^2} = \frac{1}{40^2} = 6.25 \cdot 10^{-4} ft^{-2}$$

$$q_q = \frac{1}{q_{max}^2} = \frac{1}{4^2} = 6.25 \cdot 10^{-2} \, deg^{-2} \, sec^2$$

The assignment of the neuromuscular time constant $\tau_N = .15$ sec completes the specification of the cost functional Eq. (9). It must be indicated that J is the "nominal" cost functional. With new display systems incorporated in the information set, it is possible that the pilot will attempt to minimize additional indicators. This will be explained in detail in the sequel.

DISPLAY SYSTEMS DESIGN

We present now the major analytical results of the display systems design for the candidate terrain-following task.

INFORMATION LEVEL ANALYSIS

The information level analysis is used to determine information requirements in the preliminary stages of display design and is the cornerstone to the subsequent display level analysis. There are two basic tasks at the information level: the information requirements assessment and the flight-director design.

Information Requirements

At this step we assume that all state variables, and possibly their rates are accessible to the pilot. Only the plausible rate variables should be considered. After some preliminary analysis, the following set of variables was explicity considered:

$$y' = [\Pi \; \dot{\Pi} \; \ddot{\Pi} \; e \; \dot{e} \; h \; \dot{h} \; \theta \; q] \qquad (10)$$

In the OCM application we use the cost functional defined by Eq. (9). The resulting terrain following performance, in terms of absolute values of the flight path error, is not of immediate interest at this level. We are interested in the relative importance of each state variable to the control task. Thus, the performance index (9) is minimized subject to $\Sigma f_{ci} = f_c$, $f_{ci} \geq 0$, where f_{ci} is the attentional allocation to the ith indicator. This procedure

is repeated for several values of control attention, f_c, and the summation is taken over the nine elements of Eq. (10). The optimal f_{ci} that ensue provide the relative importance of each of the states, i.e., the information requirements for the control task.

The OCM analysis yields the relative attentional allocation (f_{ci}/f_c), summarized in Table 1 and Figure 1. The attentional allocations are examined at four levels of total control attention, f_c = 0.9, 0.7, 0.5, 0.3. The only variables that command significant pilot attention are the terrain "vertical acceleration" $\ddot{\Pi}(t)$, the FPE rate, $\dot{e}(t)$, and the FPE e(t). All other system states demand negligible attention in the pilot model.

TABLE 1. RELATIVE ATTENTIONAL ALLOCATION-INFORMATION LEVEL

f_c	$f_{\ddot{\Pi}}/f_c$	$f_{\dot{e}}/f_c$	f_e/f_c
0.9	0.46	0.34	0.1
0.7	0.46	0.36	0.1
0.5	0.42	0.36	0.1
0.3	0.37	0.33	0.1

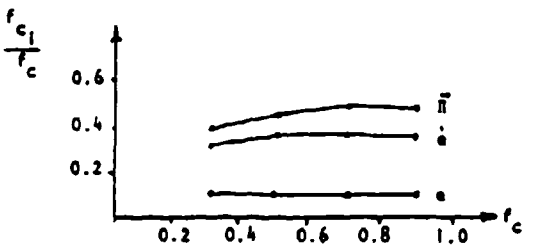


Figure 1. Relative Attentional Allocation-Information Level.

Notice that there is a consistency in these results, i.e.,

$$\frac{f_e}{f_c} = \text{const.}, \quad \frac{f_{\ddot{\Pi}}}{f_c} \cong \text{const.}, \quad \frac{f_{\dot{e}}}{f_c} \cong \text{const.}$$

Thus, the relative importance of the key state variables does not change over different levels of control attention f_c.

The results, which indicate that $\ddot{\Pi}(t)$, $\dot{e}(t)$, and e(t) is the critical information base, must now be interpreted. A major design issue, at this point, is in determining whether or not a separate display of each critical variable is required for use by the pilot. It is clear that a FPE indicator should be included since it is essential that the pilot knows his instantaneous position with respect to the desired path $\Pi(t)$. Such a display is easily constructed using, e.g., a radar altimeter. Also, the pilot can easily extract the FPE rate from a well designed display of e(t), and therefore, a separate $\dot{e}(t)$ indicator is not needed.

The major portion of the pilot's attention is allocated to $\ddot{\Pi}(t)$. In practice, this variable is difficult to derive and display explicity. Also, it is not clear how such a display might be used by the pilot. One must, therefore, synthesize a display in which $\ddot{\Pi}(t)$ will be incorporated in a practical manner.

One interpretation of the requirement for rate information $\ddot{\Pi}$(and $\dot{e}$) is the necessity to display the path-over-terrain at a future time, i.e., at some distance ahead. Such displays will be synthesized, modeled, and compared at the display element level.

Flight Director Design

The basic concept of a flight director is to provide the pilot with (synthesized) information that is useful for control, thus rendering the piloting task easier in some sense. This section describes the flight director design procedure that uses the quadratic synthesis technique of the OCM to determine the optimal (linear) aggregation of system states to be used as a display.

The flight director signal is a linear combination of the system states,

$$y_{FD}(t)=h'x(t) \tag{11}$$

The gains h are chosen so that if $y_{FD}(t)$ is kept "small" by the pilot, the resulting aircraft motion will be desirable. Since the pilot is in control of the vehicle, there are two issues that relate to the harmony between $y_{FD}(t)$ and pilot response. The first concerns the nature of the control task as viewed by the pilot. Thus, the task of keeping $y_{FD}(t)$ small should not conflict with the overall pilot-control task requirements. The second issue relates to the required form of the pilot compensation, as y_{FD} and δ are in one-to-one correspondence. From a reduced workload point of view, one should design a flight director signal $y_{FD}(t)$ such that the transfer function from input $\delta(t)$ to $y_{FD}(t)$ is approximately K/s. The required pilot compensation would then be simple proportional feedback

$$\delta(t)\approx K \cdot y_{FD}(t) \tag{12}$$

From the OCM, the pilot's control strategy is given by

$$\tau_N\dot{\delta} + \delta = -L\hat{x}(t) + v_\delta(t) \tag{13}$$

where $\hat{x}(t)$ is the state estimate, $v_\delta(t)$ is a white motor noise, and the gains L (and τ_N) are obtained by minimizing the cost function $J(\delta)$ that is associated with the terrain following task, Eq. (9). To begin with, we suggest the obvious design choice

$$y_{FD}(t) = Lx(t) \tag{14}$$

i.e., h' = L. Such a design, however, does not consider the possibility that the flight director, once added to the display panel, modifies the pilot's control task and hence changes the cost functional $J(\delta)$. Excluding y_{FD} from the cost functional implies that the pilot's control objectives are basically the same as before introducing this signal. This is not a reasonable

assumption. Indeed, including the y_{FD} with $J(\delta)$, in addition to the other terms, implies that one of the pilot's direct control objectives is to keep the y_{FD} small. Thus, we assume that the director signal y_{FD} is explicitly controlled.

The control cost functional, modified to weight deviations of $y_{FD}(t)$ is now

$$J'(\delta) = J(\delta) + E\left\{q_{FD}y_{FD}^2(t)\right\} \tag{15}$$

The weighing term q_{FD} is selected

$$q_{FD} = \frac{1}{y_{FD,max}^2} \tag{16}$$

to be consistent with the choice of the q_{yi}'s. The maximal flight director excursion is computed according to the rule

$$y_{FD,max} = \sum_i \gamma_i |l_i| x_{i,max} \tag{17}$$

where the l_i are the entries of the gain vector L, and the γ_i is 0 or 1 to indicate which variables are of concern in forming $y_{FD,max}$. We select

$$\gamma_i = \begin{cases} 1 \text{ if } x_i \text{ is a positional variable} \\ 0 \text{ if } x_i \text{ is a rate variable} \end{cases} \tag{18}$$

Thus, the flight director signal is at its maximum value when all error displacements are at their design limits.

With the pilot cost functional modified as in Eq. (15), the pilot model control is now obtained by minimizing a new cost functional, the result being

$$\tau_N \dot{\delta} + \delta = -L\hat{x}(t) + v_\delta(t) \tag{19}$$

But since the cost functionals of Eq. (9) and (15) are not the same, the optimal gains L in Eq. (19) differ from those in Eq. (13). Hence, the flight director signlas of Eq. (14) and the required pilot control gains in Eq. (19) are no longer in harmony. This mismatch can be corrected via the iterative process of computing feedback gains and flight director signals as shown in [1].

This algorithm has given rapid convergence in the terrain following problem. Only four iterations have been needed, and the resulting converged gains were within 10 percent of the initial gain values obtained from Eq. (9).

The numerical results of the flight director design process are:

1. design parameter, $y_{FD,max} \simeq 3$
2. flight director weighting, $q_{FD} = \frac{1}{y^2_{FD,max}} = \frac{1}{3^2} \simeq .1$
3. flight director signal

$$y_{FD} = Lx = \left[3 \cdot 10^{-3} \quad .13 \quad .11 \quad .76 \quad -.48 \quad -1.2 \quad .06 \right] x$$

We do not yet examine terrain following performance with the flight director display, as this is the subject of the display element level in wich all proposed displays are compared.

DISPLAY ELEMENT LEVEL ANALYSIS

The next step in the display design methodology is to evaluate control performance for several candidate display systems. The evaluation process begins with the definition of a performance metric for the terrain following task.

Often, the control performance requirements of the pilot-vehicle combination are specified in terms of allowable excursions or desired RMS deviations in system states. The design specifications are generally a function of mission requirements or flight conditions. Recall that the nominal altitude of the aircraft over the terrain is 200', and that the maximal FPE deviation "allowed" is 40'. Thus, $e_{max} = 40'$ is chosen as the design tolerance in the control performance metric, which is defined as

$$P_c = \int_{-e_{max}}^{e_{max}} \frac{1}{\sqrt{2\pi\sigma_e^2}} \exp\left\{ \frac{e^2(t)}{2\sigma_e^2} \right\} d[e(t)] \qquad (20)$$

This performance metric measures the probability that the aircraft does not deviate more than $e_{max} = \pm 40'$ from the desired path $\Pi(t)$. The maximum level of control performance is selected as $P_{c,max} = .99$.

The OCM parameters to be used in all candidate display systems are:

observation noise (for all eventual indicators), $\rho_{yi} = .20$ dB;
motor noise, $\rho_\delta = -25$ dB; time-delay, $\tau_D = .15$ sec.

The neuromuscular time-constant has already been specified as $\tau_N = .15$ sec.

Using these parameter values and the performance metric P_c, Eq. (20), we now evaluate and compare the candidate display systems.

Display System Synthesis

STATUS DISPLAY

The status display system consists of a FPE $e(t)/\dot{e}(t)$, and pitch and pitch rate, $\theta(t)$, $q(t)$, indicators. This rudimentary display is not a truely synthesized system at which this study is aimed; nevertheless, we use the status display as a benchmark in the display system comparison process. The selection of $[e, \dot{e}, \theta, q]$ for the status display set is natural [6], and all subsequent display systems include the status display set. Naturally, the status display system alone should yield the worst terrain following performance. Also, we assign indifference threshold values, a_i to $e(t)$, $\dot{e}(t)$, $\theta(t)$, and $q(t)$, according to the rule [6] $a_i = |y_{i,max}|/8$. Following the usual assumption that rate variable thresholds are one half of the thresholds on the corresponding position variables, we obtain

$$a_e = \frac{1}{8} e_{max} = \frac{1}{8} \cdot 40' = 5\,; \quad a_{\dot{e}} = \frac{1}{2} a_e = 2.5'/\text{sec}\,;$$

$$a_q = \frac{1}{8} q_{max} = \frac{1}{8} \cdot 4°/\text{sec} = 0.5°/\text{sec}\,; \quad a_\theta = 2 \cdot a_q = 1°$$

PREDICTOR DISPLAY

Grunwald and Merhav have shown [8]-[9] that acceleration cues are vital in visual field control and that they are obtained by estimating the future vehicle path. These findings are in close agreement with our information level analysis, in which it was shown that the terrain "acceleration", $\Pi(t)$, demands the largest amount of attention among all system states. Such information can best be derived (or estimated) by the pilot when the vehicle's future altitude is displayed relative to the future terrain-path. We define T as the prediction time (i.e. $T \simeq D_o/U_o$, where D_o is the distance ahead at which the prediction is displayed), and the predictor signal is then given by

$$e_p(t) = \Pi(t+T) - h(t+T) \tag{21}$$

The underlying assumption in the predictor display is that the predictor vehicle path is along the velocity vector $\bar{U}_o$. We rewrite Eq. (21) to reflect this assumption, viz.,

$$e_p(t) = \Pi(t+T) - \left[h(t) + T\dot{h}(t)\right] = \Pi(t+T) - h(t) - \frac{\pi T U_o}{180°}\left[\theta(t) - \alpha(t)\right] \tag{22}$$

Also, the OCM assumes that the pilot derives rate information from the predictor indicator, $\dot{e}_p(t)$. The appropriate equation for this signal is

$$\dot{e}_p(t) = \dot{\Pi}(t+T) - \dot{h}(t) - T\ddot{h}(t)$$

$$= \dot{\Pi}(t+T) - \frac{\pi U_o}{180°}\theta(t) + \frac{\pi U_o}{180°}\left[1 + TZ_w\right]\alpha(t) + \frac{\pi T Z_\delta}{180°}\delta(t) \tag{23}$$

Such a display is easy to simulate and implement using e.g., a forwardlooking radar. An actual display that was used in the experimental validation is shown in Figure 6. It is important to indicate that the projection of $\vec{U}_o$ on a normal surface located T seconds flight time ahead is required (represented by the cross in Figure 6.

Equations (22-(23) in their present form cannot be modeled directly in the OCM steady-state analysis, where only events at time t are treated. This problem can be "solved" by replacing the signals $\Pi(t+T)$ and $\dot{\Pi}(t+T)$ with their (optimally) predicted future values, viz.,

$$\hat{\Pi}(t+T) = E\left\{\Pi(t+T) \mid \Pi(t), \dot{\Pi}(t), \ddot{\Pi}(t)\right\} \; ; \quad \hat{\dot{\Pi}}(t+T) = \left\{E\,\dot{\Pi}(t+T) \mid \Pi(t), \dot{\Pi}(t), \ddot{\Pi}(t)\right\} \tag{24}$$

Such an approximation can be easily obtained from the 3 x 3 upper-left block of A in Eq. (8). This results in estimates which are a linear combination of the terrain states $\Pi(t)$, $\dot{\Pi}(t)$, $\ddot{\Pi}(t)$ [6]. Specifically we may write

$$\begin{aligned} \hat{\Pi}(t+T) &= p_{\Pi}(T)\Pi(t) + p_{\dot{\Pi}}(T)\dot{\Pi}(t) + p_{\ddot{\Pi}}(T)\ddot{\Pi}(t) \\ \hat{\dot{\Pi}}(t+T) &= r_{\Pi}(T)\Pi(t) + r_{\dot{\Pi}}(T)\dot{\Pi}(t) + r_{\ddot{\Pi}}(T)\ddot{\Pi}(t) \end{aligned} \tag{25}$$

where, as indicated, the p and r coefficients are a function of the prediction time, T. Thus, we rewrite Eqs. (22)-(23) as a linear combination of the system state:

$$\begin{bmatrix} e_p(t) \\ \dot{e}_p(t) \end{bmatrix} = C_p(T)x(t) + D_p(T)\delta(t)$$

$$= \begin{bmatrix} p_{\Pi}(T)-1 & p_{\dot{\Pi}}(T) & p_{\ddot{\Pi}}(T) & \dfrac{\pi TU_o}{180^\circ} & 0 & \dfrac{-\pi TU_o}{180^\circ} & 1 \\ r_{\Pi}(T) & r_{\dot{\Pi}}(T) & r_{\ddot{\Pi}}(T) & \dfrac{\pi U_o}{180^\circ}(1+TZ_w) & 0 & \dfrac{\pi U_o}{180^\circ} & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \dfrac{\pi TZ_\delta}{180^\circ} \end{bmatrix} \delta(t) \tag{26}$$

The indifference thresholds on the predictor are assumed identical to those of the FPE, $e(t)/\dot{e}(t)$.

The last issue which must be addressed is the selection of the prediction time T. The present A10 HUD uses the value T≃4 sec. This value has been selected for use in the present display analysis and in the subsequent experiments. It has been shown [6], however, that the optimal prediction time for the given control task and assumed terrain characteristics is $T^* = 1.5$ sec. The numerical values of $C_p(T)$ and $D_p(T)$ of Eq. (26) for T = 4 sec. and $T=T^*=1.5$ sec. are given in the Appendix.

THREE DIMENSIONAL PERSPECTIVE TUNNEL DISPLAY

The idea of a displaying 3-D perspecitve tunnel which envelopes the trajectory over the terrain is not entirely new. It has recently been studied and simulated by Grunwald in a helicopter approach context [10]. Such a display is a computer-generated, "through-the-windshield" perspective view of a tunnel that follows the contours of the terrain. In practice, it seems as if the tunnel was flying towards the observer. The pilot, on his part, tries to maintain the aircraft as close to the tunnel's center as possible. Figure 7 shows the tunnel geometry and display which was used in the experimental validation.

The present FPE, $e(t)$, is not explicity available to the pilot from the tunnel display. He can, however, derive sufficient rate and acceleration information, as dictated by the information level results, from a continuance of future flight path errors displayed by the perspective tunnel. It is necessary now, to translate the information provided by the tunnel into an analytical model for application of the OCM.

A plausible approach to the modeling problem has been suggested by Tomizuka [11] in the so-called "finite preview" problem. If the tunnel is sufficiently long, we may assume an infinite preview time. The finite preview problem is then reduced to a common optimal tracking problem, and can be treated as such.

Both approaches, albeit plausible, require major modifications in the OCM methodology and, therefore, are not considered here. The modeling approach taken in the present study treats the tunnel display in the OCM framework. As indicated, the tunnel provides a continuance of future flight path errors, assuming a straight flight path. Formally, this information base may be represented as $e(t+\tau)$, $\tau\varepsilon[o,t_p]$ where, in general, $t_p<\infty$. However, such a representation is impractical from a modeling point of view. One approximation is to replace $e(t+\tau)$ with N_p "indicators" which represent the FPE values at distinct points in the future, viz., $e(t+\sigma_i)$, $\sigma_i = t_{pi}/N_p$ where i varies from 1 to N_p. The indicators $e(t+\sigma_i)$ would then be treated as independent N_p predictor display systems. Such a model, although requiring an extensive set of equations, if N_p is chosen to be large, can be implemented in the OCM framework. This modeling approach is carried further by assuming that the pilot concentrates on a single distance ahead, i.e., $N_p=1$. Since the underlaying assumption of the OCM is that the pilot adopts an optimal control policy, it is equally valid to assume that, in the tunnel display, he choses the optimal prediction distance/time, T^*, when looking down the tunnel path. We, therefore, replace the N_p $e(t+\sigma_i)$ observations with a single indicator $e(t+T^*)$. The tunnel display model, $\kappa(t)$, is then simply

$$\begin{bmatrix}\kappa(t)\\ \dot{\kappa}(t)\end{bmatrix} = \begin{bmatrix}e(t+T^*)\\ \dot{e}(t+T^*)\end{bmatrix} = C_p(T^*)x(t) + D_p(T^*)\delta(t) \qquad (27)$$

where the numerical values of $C_p(T^*)$ and $D_p(T^*)$ $T^*=1.5$ sec. are given in the Appendix.

Despite initial similarities, there is a fundamental difference between a simple optimal predictor display where $T=T^*=1.5$ sec., and a tunnel display. Although both displays are represented by the same equation (26), the tunnel display does not include the velocity vector's tip, $\vec{U}_o$ (Fig. 6) projected on the normal surface located T^* seconds flight time ahead. The end point of $\vec{U}_o$ is <u>estimated</u> rather than <u>displayed</u> in the perspective tunnel. This fact is reflected in the OCM by large observation thresholds on $\kappa(t)$, $\dot{\kappa}(t)$. Given the $\pm 40'$ tunnel dimensions, as implemented in the subsequent experiments, we assume

$$a_\kappa = 20', \; a_{\dot{\kappa}} = \frac{1}{2} a_\kappa = 10'/\text{sec}$$

These thresholds are significantly larger than the indifference thresholds used for the simple predictor display (5 ft and 2.5 ft/sec respectively). Naturally, such large thresholds tend to degrade terrain following performance. To overcome this problem the following display system is suggested.

INTEGRATED TUNNEL/VELOCITY VECTOR DISPLAY

In this system we simply superimpose the velocity vector's ($\vec{U}_o$) trace on the existing tunnel display as shown in Fig. 8. Again, the value used is $T=4$ sec and not T^*, in accordance with the current A10 display system. It is obvious that incorporating this new information will reduce the thresholds $a_\kappa/a_{\dot{\kappa}}$, since the pilot now has a reference point about which he will "center" $\Pi(t+T^*)$. Since $\vec{U}_o$ is projected at $T=4$ sec., and the pilot's "focus" in the tunnel is at $T=1.5$ sec., the thresholds $a_\kappa/a_{\dot{\kappa}}$ are not reduced to the $a_e/a_{\dot{e}}$ values, as in the predictor display, but rather to an intermediate value. We select

$$a_\kappa = 10', \; a_{\dot{\kappa}} = 5'/\text{sec}$$

The information base now includes both tunnel $\kappa(t)/\dot{\kappa}(t)$, and a 4 second predictor, $e_p(t)/\dot{e}_p(t)$, (in addition to the status display). The $e_p, \dot{e}_p$ threshold values remain unchanged, 5', 2.5'/sec. respectively.

FLIGHT DIRECTOR SYSTEM

The status information base is now augmetned with the flight director position and rate observation $y_{FD}(t)$, $\dot{y}_{FD}(t)$ as discussed previsouly. The $y_{FD}/\dot{y}_{FD}$ observation equations are given by

$$y_{FD}(t) = Lx(t) = [3.10^{-3} \; .13 \; .11 \; .76 \; -.48 \; -1.2 \; .06] \; x(t) \qquad (28)$$

$$\begin{aligned} \dot{y}_{FD}(t) &= L\dot{x}(t) = LA_o x(t) + LB_o \delta(t) \\ &= [-3.10^{-3} \; .03 \; .05 \; .91 \; .31 \; -.48 \; 0]x(t) + 3.2\delta(t) \end{aligned} \qquad (29)$$

Also, using the fact that $y_{FD,max} \simeq 3$, the indifference thresholds are

$$a_{FD} = \frac{1}{8} \cdot y_{FD,max} = \frac{1}{8} \cdot 3 \simeq .4; \; a_{\dot{FD}} = \frac{1}{2} a_{\dot{FD}} \simeq .2$$

Implementation of the flight director in a practical manner, using future terrain path information in lieu of $\dot{\Pi}$ and Π, is described in the sequel.

Thus, we have proposed and obtained analytical models for five candidate display systems. The next task in the analysis procedure is to evaluate control performance and attention allocation for these systems.

Control Performance: Modeling Results

The performance of each of the display systems is evaluated in terms of

1. control performance, and
2. acceleration stress levels

Following [1]-[6], we introduce now the concept of control and monitoring workload. The control workload metric is based on the fractional attention the pilot allocates among the various display indicators. It is assumed that a pilot distributes a total amount of attention, or workload, $f_T \approx 0.8 < 1.0$ between the tasks of control and monitoring, leaving about 20 percent of his capacity for other duties (e.g., communications). Let f_c and f_m denote, respectively, the control and monitoring attentions, or workloads. Thus. $f_c + f_m = f_T$. The attention allocated for control, f_c, is distributed among all of the display variables $y_1, y_2, \ldots, y_{Ny}$, where y_i and $y_{i+1} = \dot{y}_i$ (i = odd) are obtained from the same display indicator. If $f_{ci} > 0$ is the attention allocated to y_i for control purposes, then the constraints on f_{ci} are

$$\sum_{i=odd} f_{ci} = f_c \quad ; \quad f_{c,i+1} = f_{ci} \quad i = 1, 3, 5, \ldots \tag{30}$$

The pilot allocates his attention among displays, spending the larger f_{ci} on displays that are most useful for control.

With f_{ci} selected, the pilot-vehicle model yields predictions of the performance metric, P_c, Eq.(20). Using this prediction, we can study the tradeoffs between f_c and P_c for any given display system. Figure 2 is a typical performance/workload curve. It shows the performance attained for a given workload, as well as the workload required to obtain a given performance level. In Figure 2 the intersection of the line $P_c = P_{c,max}$ with the P_c versus f_c curve gives the minimum amount of control attention required, $f_{c,req}$, for the given system to meet P_c specifications. The difference between this amount of attention, and the total available for the entire task is the residual workload available for monitoring

$$f_{m,avail} = f_T - f_{c,req} \tag{31}$$

The process of comparing the candidate display systems is now clear. As an example, one may observe Figure 3. Clearly display system 1 is superior to display system 2, since less control workload (or required control attention)

is needed to meet the required performance level. Moreover, more attention is available for monitoring duties when using display system 1.

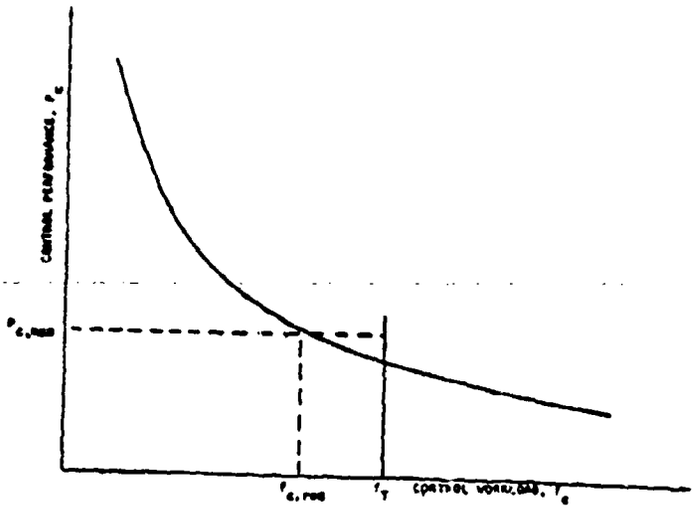


Figure 2. Conceptual Control Performance Versus Workload Curve.

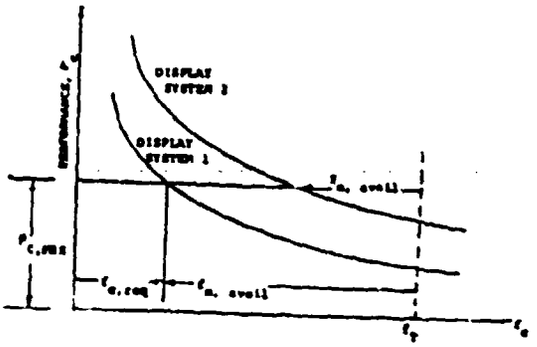


Figure 3. Guidelines for Evaluating Control/Display Designs.

In the terrain following task we compute the control performance metric P_c, Eq. (20), for the candidate systems given the constraints f_T = 0.2, 0.4, 0.6, 0.8 and 1.0. Also, the pertinent RMS flight path errors (e_{rms}) and g-levels (g_{rms}) at the same f_T values are obtained. These results are summarized in Table 2.

TABLE 2. PREDICTED CONTROL PERFORMANCE (P_C), FPE, (ft) AND g-STRESS (g's) RESULTS

Display System	f_T	0.2	0.4	0.6	0.8	1.0
Status	P_c	.84	.92	.95	.96	.97
	e_{rms}	28.6	22.6	20.4	19.0	18.0
	g_{rms}	.54	.50	.48	.47	.46
Tunnel	P_c	.87	.95	.97	.98	.99
	e_{rms}	26.2	20.4	18.0	16.6	15.8
	g_{rms}	.50	.46	.44	.42	.42
Predictor	P_c	.93	.98	.99	>.99	>.99
	e_{rms}	22.4	17.6	15.6	14.4	13.8
	g_{rms}	.44	.41	.40	.38	.39
Tunnel + Predictor	P_c	.94	.99	>.99	>.99	>.99
	e_{rms}	21.0	16.2	14.6	13.6	13.0
	g_{rms}	.45	.42	.41	.40	.39
Flight-Director	P_c	>.99	>.99	>.99	>.99	>.99
	e_{rms}	15.0	12.2	11.4	11.0	10.6
	g_{rms}	.41	.39	.38	.37	.27

Using these numerical results, the display systems attain the rank ordering as shown in Figure 4. Using Eq. (31) we are now able to compute the required control-attention, $f_{c,req}$ (workload), and the available monitoring attention, $f_{m,avail}$, for each of the candidate display systems. We assume f_T=.8. These results are summarized in Table 3.

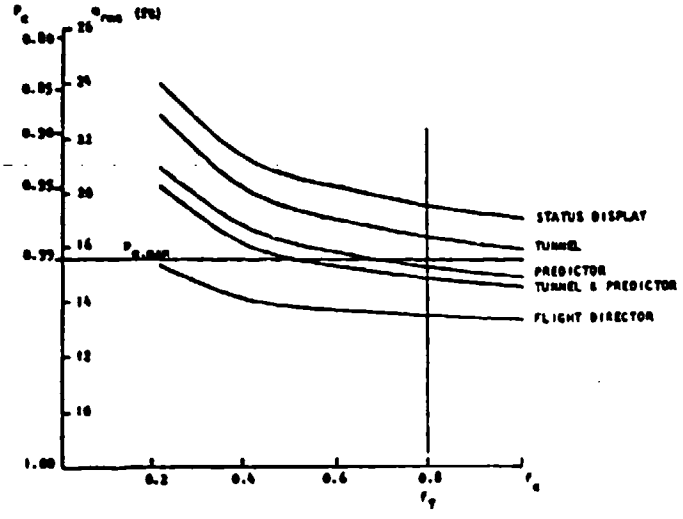


Figure 4. Display System Rank Ordering - Analytical Results.

TABLE 3. WORKLOAD AND MONITORING ATTENTION RESULTS

Display System	$f_{c,req}$	$f_{m,avail}$
Status	>>.8	0
Tunnel	>.8	0
Predictor	.64	.16
Tunnel + Predictor	.48	.32
Flight-Director	<.2	>.6

It is evident from these results that some form of synthesized display must be considered in the A10 aircraft, as the rudimentary status display yields an unsatisfactory control performance. It is also clear that the best system, in terms of reduced control workload, available monitoring attention, and lowest levels of g-stress, is the flight director display. In addition, the flight director's P_C-versus-f_C curve is almost flat (i.e. dP_C/df_C=0) for $f_C \in [0.4,0.8]$, which indicates that this design is robust with respect to external attentional demands that might be placed on the pilot. The tunnel, on the other hand, exhibits supprisingly poor results, as it ranks only fourth, after the predictor, and does not meet the performance criterion.

The predictor display requires a very high control attention, $f_{c,req}$=.64, and and the available monitoring attention, $f_{m,avail}$, is only .16. Although the design specifications are met, such a display system may not be acceptable, as it would be too sensitive to a possible degradation in control attention capacity. However, the integrated tunnel/predictor display system gives a satisfactory performance. Next we validate the modeling results experimentally.

EXPERIMENTAL VALIDATION

A primary objective of this study was the validation of the model-based display design procedure described previously. The validation phase consisted of

fixed-base man-in-the-loop simulations of the A10 terrain following scenario for the four synthetic displays. The experiments, conducted largely independently of the analytic effort, were performed at the University of Connecticut.

DISPLAY FORMAT

A precursor to the experimental phase is the design of the display format, i.e. the details of the display panel layout. Clearly, this is largely an art, but can be guided by the results of the display element analysis with regard to threshold values and scale range. Four basic, or status, displays were used in the experiments in addition to the synthetic display. The basic displays were the following.

1. Error Indicator. This showed instantaneous error about the nominal terrain-following path. We used a vertical scale of ± 50 ft range (recall 40 ft is the maximum design error). The distance between scale markings/divisions was chosen as 5 ft, which corresponds to a display threshold of ~ 2.5 ft.*

2. Pitch Indicator. A stylized aircraft pitch indicator was used to display $\theta(t)/q(t)$. A maximum range $\pm 10°$ was allowed. Minimum scale marking was 2.5°.

3. g-Meter. Although vertical acceleration was not shown to be of significance as an observation, our analysis assumed that RMS g-level entered (subjectively) in the pilot's cost functional. Since our simulation was fixed-base, the only possible perception of g-level was via visual stimulus. Thus, the subjects were "aware" of their commanded accelerations.

4. Radar Altimeter. This display is essentially a duplication of the error information, i.e. the difference in altimeter reading from 200 ft is the error. It is included for those cases where the error indicator may be off-scale, $|e(t)| > 50'$. In addition, any realistic display panel will likely contain this information.

Figure 5 shows the display panel layout that we used. The center screen area was set aside for the specific synthetic displays to be investigated. The display in the lower right corner is associated with a side monitoring task

*We generally assume that the display threshold is half the minimum scale marking or 0.05° visual arc, whichever is the larger. The display threshold should be less than the control indifference threshold, y_{max}/θ, for a well-designed display.

which will not be addressed at this time. The entire display was presented to the subject on a VS60 graphics screen; the total display size was 14" x 12".

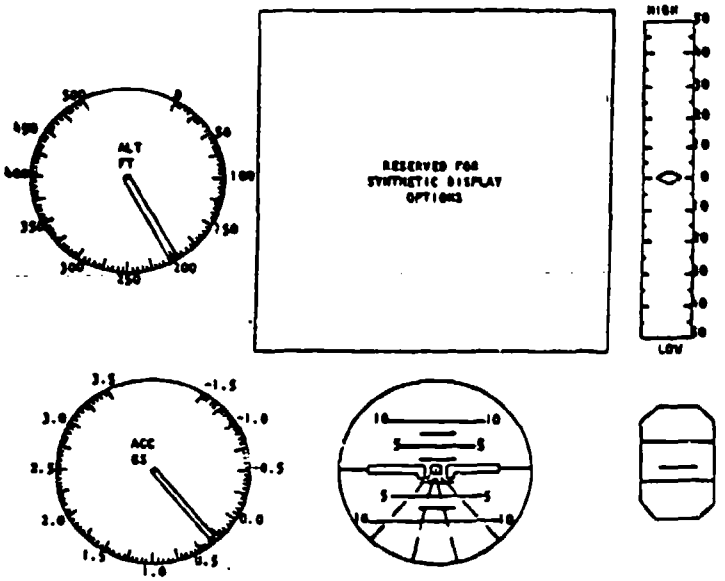


Figure 5. Basic Display Panel Layout for Terrain-Following Simulation.

The basic display format of Figure 5 was the same for all experiments. The only difference among the cases studied was the form of the synthetic display. These are now described.

Predictor Display

In this display we present the future terrain $\Pi(t+T)$ and extrapolated aircraft position $h(t)+Th(t)$ where T=4 sec. The display format used is shown in Figure 6. Here the cross represents the aircraft flight vector. The "terrain-box" represents an 80'(H) x 100'(W) window centered on the terrain path at a distance $D_o=4xU_o\simeq1900'$ ahead of the aircraft. Thus, if the subject kept the cross within the box, the linear prediction of future error would be <40'. For convenience, the cross was fixed at the center of the synthetic display area, i.e., only the terrain-box moved.

Tunnel Display

The stylized tunnel display that was programmed for the experiments is shown in Figure 7. The "tunnel" consists of five "windows", separated in distance by 500'. Thus, with the extension lines, the tunnel presents the future path some 2500' - 3000' ahead of the aircraft. The tunnel is centered on the nominal terrain path and has longitudinal dimensions $\pm$40' (to correspond with e_{max}) and lateral dimensions $\pm$50'.

The tunnel (windows) are fixed in inertial space. Thus, as the aircraft "flies" forward, the tunnel windows move towards the observer. When the leading window reaches a minimal distance of 100' from the aircraft it disappears, and a "new" window appears at the tunnel's end. This gives the illusion of continual forward motion. The perspective view of the tunnel is along the aircraft's flight vector, $\gamma(t)$, i.e., the viewing axes are aircraft centered with the forward z-axis aligned with $\gamma(t)$.

In the present experiments, tunnel variations occur only in the longitudinal axis. However, the computer simulation can treat tunnel/terrain and aircraft motion in both longitudinal and lateral axes. A complete description of the computer simulation and software may be found in [12].

Integrated Tunnel and Predictor Display

This display format is essentially a combination of the two previous displays. The integration has been effected by adding an additional "window" to the tunnel display at a range D_o=1900' ahead of the aircraft. This window does not

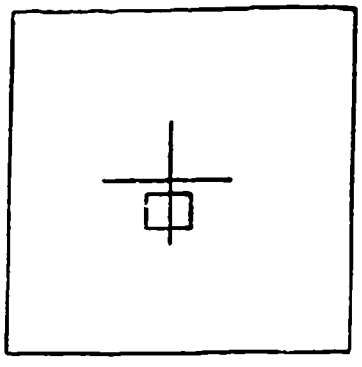


Figure 6. Predictor Display Symbology.

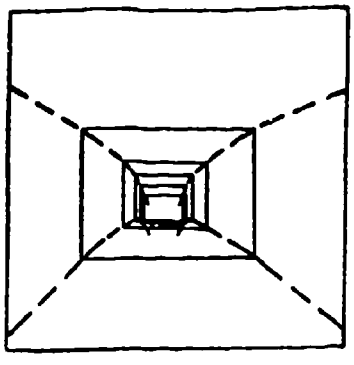


Figure 7. Tunnel Display Format.

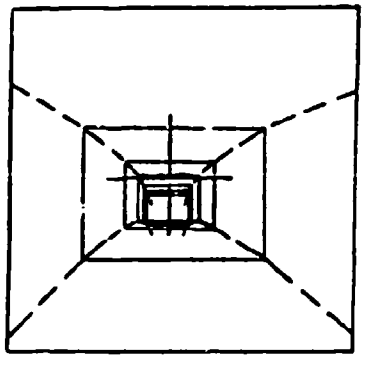


Figure 8. Integrated Tunnel and Predictor Display.

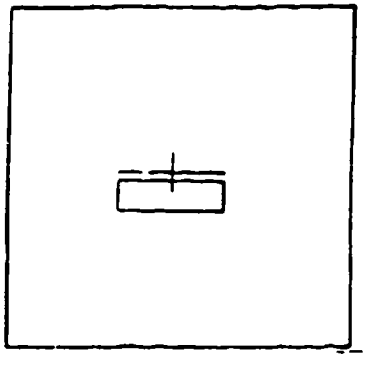


Figure 9. Flight Director Display Symbology.

move towards the observer, but the other windows pass through it. The display format is shown in Figure 8. In order that the predictor window be visually prominent it is shown brighter than the other elements that make up the tunnel display.

The velocity, or flight-path vector is superimposed on the tunnel as a cross. Since the tunnel view is centered on this vector, the cross remains stationary in the center of the viewing area. Of course, the tunnel (which is fixed in inertial space) curves up or down depending on the terrain and aircraft motion. We note that if the tunnel view were centered on the aircraft pitch angle, then the velocity vector projection would per force be different.

Flight Director Display

The flight director signal, $y_{FD}(t)$, as derived via the methodology described earlier is given by

$$y_{FD}(t) = -.003\Pi(t) + .1277\dot{\Pi}(t) + .1095\ddot{\Pi}(t) + .76\alpha(t) - .476q(t) - 1.245\theta(t) + .0594e(t) \quad (32)$$

This has been implemented as

$$y_{FD}(t)=c_1\Pi(t+\bar{T})-c_2\Pi(t)+.76\alpha(t)-.476q(t)-1.245\theta(t)+.0594e(t) \quad (33)$$

where $c_1 = .0745$, $c_2 = .0775$, $\bar{T} = 1.715$ sec. In deriving Eq. (33) we use the approximation $\Pi(t+\bar{T}) = \Pi(t) + \bar{T}\dot{\Pi}(t) + (\bar{T}^2/2)\ddot{\Pi}(t)$ and equate coefficients with Eq. (32). Note that c_1 may be assumed equal to c_2 with little or no observed effects.* In Eq. (33), $\Pi(t+\bar{T})$ is the terrain path T sec. (or $\bar{D}=\bar{T}\cdot U_o \simeq 800$ ft) ahead of the aircraft.

The flight director display is presented to the subject in the form of a compensatory tracking task, as shown in Figure 9. The cross in Figure 9 is stationary with respect to the viewing area; the signal that drives the box is given by Eq. (33). In order to distinguish this display from the predictor the box and cross sizing is different.

EXPERIMENTAL DESIGN

The fixed-base experiments were conducted using a PDP 11/60 to simulate the aircraft equations, terrain, and to update the displays. The display was presented on a VS60 graphics screen and refreshed 30 x per second. Pilot input was via a force stick controller. The specifics of the simulations are given in [6].

*This is convenient as only the path different $\Pi(t+T) - \Pi(t)$ is needed in the flight director signal.

Four subjects for the experiments were selected from the Air Force ROTC student body at the University of Connecticut. The display conditions were presented to them using a Latin square ordering to minimize any transition effects on averaged performance. A data trial lasted 130 sec., the last 0.06 x 2048 = 122.88 sec. of which was used as data. We recorded the 2048 samples of control input δ and error e for each trial. In addition, we computed and recorded RMS values of error, control, pitch and vertical acceleration for each run. Thus, we obtained a total of N = 8 x 4 = 32 trials for each display condition.

It should be noted that none of the subjects had flight experience; two had some fixed-base trainer experience. Thus, it is quite likely that the subjects were (uniformly) not expertly trained on the control task. However, it is quite likely that the relative differences in performance for different displays is not strongly dependent on absolute training level in the present task.

EXPERIMENTAL RESULTS

Table 4 gives the experimental results averaged across the four subjects. The averages were computed first for each subject and then across subjects to obtain the grand averages. The standard deviations in the experimental results, shown in parentheses in Table 4, are the averaged intra-subject variations and not the inter-subject variations. Thus, these numbers are indicative to run-to-run variability that might be associated with a single ("average") human.*

TABLE 4. AVERAGED EXPERIMENTAL RESULTS

Case	N	e_{RMS}	(ft)	g_{RMS}-1	(g)
1	32	22.0	(3.4)	0.44	(0.068)
2	30	25.7	(4.5)	0.53	(0.078)
3	33	17.8	(2.2)	0.44	(0.084)
4	36	8.95	(1.9)	0.36	(0.054)

The results tabulated in Table 4 are quite interesting with regard to the underlying considerations in our display design technique. First note the rank-ordering of displays with respect to e_{RMS} performance. Here we see that the tunnel display (case 2) fares worst. The predictor display (case 1) fares slightly better than the tunnel alone. The combined tunnel plus predictor display (case 3) results in a meaningful performance improvement over that of (1) or (2). The flight director display (case 4) yields significantly better performance than any other display condition -- by a factor of 2. This is highly encouraging validation of our flight director design/synthesis procedure.

The rank-ordering of the different display configurations is in precise agreement with the analytical results of the display design methodology. While the

*If intersubject variability was included in the standard deviations, the values would increase by 20-100%.

absolute levels of control performance between experiment and model disagree slightly, relative performance levels agree well. This is demonstrated by comparing model predictions (at a fixed $f_c \simeq 6$) for cases 1-3 with the experimental results. The e_{RMS} experimental results for cases 1-3 are consistently higher than the model predictions. An explanation for this fact is that the OCM assumes a well-trained subject, whereas the actual subjects -- not being pilots -- were not fully trained with respect to the A10 dynamics. While it is possible to model this effect a posteriori in the OCM by increasing observation/motor noise and/or τ_N, this was not an objective of our efforts. On the other hand, the absolute performance levels for model and data in case 4 are in close agreement. The reason for this is that the "system" dynamics as perceived by the subject are similar to K/s. These dynamics are trivial to learn, so that training effects (after but a few trials) are inconsequential.

CONCLUSIONS

An analytic, pilot model-based, display design methodology has been applied to study workload and performance trade-offs in a high-speed, terrain-following task. The methodology combines pilot limitations, aircraft dynamics and performance requirements in order to determine the requisite information that must be supplied to the pilot.

Man-in-the-loop experiments that evaluated the performance of the four candidate display systems were conducted at the University of Connecticut. The objective of these experiments was to validate the overall display design procedure, including the flight director design synthesis process. Two positive conclusions resulted from this effort.

VALIDATION OF DISPLAY SYSTEM PERFORMANCE

The fixed-base experiments involved a terrain-following control task. (A secondary monitoring task was also included but is discussed elsewhere [6].) The experimental relative rank-ordering of the four display systems, based on control performance, was identical to that predicted analytically at the display element level. It suggests that a large number of potential display (or control augmentation) systems can be evaluated analytically at low expense, and then the most promising options can serve as the candidates for subsequent manned simulation. The time and cost savings of a model-based "front-end" to the complete design process can be substantial.

The spread in absolute levels of performance between the first three display systems and the flight director display was found to be much greater in the experiments than in the model predictions. Our explanation of this result is that the model assumes a well-trained pilot, whereas the subjects were not well-trained on A10 dynamics and so their performance was not at the model-predicted levels. On the other hand, the flight director essentially normalizes out the aircraft dynamics, rendering the control task much simpler and requiring virtually no learning. Here model and data absolute performance levels were commensurate.

VALIDATION OF FLIGHT DIRECTOR DESIGN PROCEDURE

The analytic technique for flight director synthesis is included in the display methodology at the information level. Here, we optimally synthesize or aggregate the information states into a single information variable that could be displayed to the pilot. The flight director signal is designed to relate to the pilot task objectives, i.e. to minimize his workload and/or improve his control performance, and to satisfy the pilot's desired goal of behaving approximately as a gain and time-delay. The man-in-the-loop simulations validated the superiority of the flight director display (over all others considered) with respect to control performance. The ability to analytically design a flight director that is in harmony with pilot control and information processing limitations is a major feature of the methodology. In many situations flight directors are "designed" via extensive simulations and tuning using a pilot in-the-loop. The analytic design, when used as prescribed, can shorten this experimental procedure to a great extent.

REFERENCES

1. Hoffman, W.C., D.L. Kleinman, and L.R. Young, "Display/Control Requirements for Automated VTOL Aircraft," NASA Contractor Report 158905, ASI-TR-76-39.

2. Curry, R.E., D.L. Kleinman, and W.C. Hoffman, "A Design Procedure for Control/Display Systems," Human Factors, 1977, 19(5), pp. 421-436.

3. Kleinman, D.L. and S. Baron, "Manned Vehicle Systems Analysis by Means of Modern Control Theory," NASA CR-1753, June 1971.

4. Kleinman, D.L., S. Baron, and W.H. Levison, "A Control Theoretic Approach to Manned Vehicle Systems Analysis," IEEE Trans. on Automatic Control, Col. AC-16, No. 6, December 1971.

5. Kleinman, D.L. and S. Baron," Analytic Evaluation of Display Requirements for Approach to Landing," NASA CR-1952, November 1971.

6. Korn, J. and D.L. Kleinman,"Advanced Cockpit Pilot Information Modeling," ALPHATECH TR-119, ALPHATECH, Inc., 3 New England Executive Park, Burlington, Massachusetts, 01803.

7. McRuer, D., I. Ashkenas, and D. Grahm, Aircraft Dynamics and Automatic Control, Princeton University Press, 1973.

8. Grunwald, A.J. and S.J. Merhav, "Vehicular Control by Visual Field Cues--Analytical Model and Experimental Validation," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-6, No. 12, December 1976.

9. Grunwald, A.J. and S.J. Merhav, "Effectiveness of Basic Display Augmentation in Vehicular Control by Visual Field Cues," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-8, No. 9, September 1978.

10. Grunwald, A.J., "An Experimental Evaluation of an Electronic Perspective Tunnel Display for 3-D Helicopter Approach-to-Landing," TECHNION, Faculty of Aeronautical Engineering Report, TAE 383, October 1979.

11. Tomizuka, M., "Optimal Continuous Finite Preview Problem," IEEE Trans. on Automatic Control, June 1975.

12. Lovinsky, V.S. and D.L. Kleinman, "Software Documentation for TRNFLW," University of Connecticut, Department of EECS, Technical Report TR-81-10. June 1981.

APPENDIX

C_p AND D_p VALUES (EQ. 26)

T (sec)	$C_p(T)$							$D_p(T)$
1.5 (T*)	-.007	1.43	.83	12.25	0	-12.25	1	0
	-.013	.87	.93	-8.65	0	-8.17	0	-.75
4.0	-.09	3.03	3.43	32.66	0	-32.66	1	0
	-.050	.39	.95	-36.68	0	-8.17	0	-2.0

AD P001632

HUMAN-MACHINE INTERFACE ISSUES IN THE DESIGN OF INCREASINGLY AUTOMATED NASA CONTROL ROOMS[1]

Christine M. Mitchell
Decision Sciences Faculty
George Mason University
Fairfax, VA 22030

ABSTRACT

The paper provides a report of work-in-progress examining the potential effects of introducing increased levels of automation in command and control environments for NASA's near-earth satellites. To date, the work has examined the implications of automation and concluded that there are times when costs outweigh benefits, thus suggesting that automation not be introduced into a system. Assuming that some level of automation is introduced, the roles of the human operator in automated systems is explored. Further research on information displays designed to support the roles of the human in automated systems is described.

INTRODUCTION

NASA-Goddard Space Flight Center in Greenbelt, Maryland provides ground support and direct control for all of NASA's near-earth satellites. The Greenbelt facility has extensive control rooms which permit real time, near real time, and time delayed control of the various components on board near-earth satellites. With the support of a world wide network of ground tracking stations, the mission controllers in Greenbelt establish contact with each satellite several times a day; each satellite contact lasts approximately twenty to thirty minutes.

Traditionally, each NASA mission had its own dedicated mission operations room, i.e., a dedicated control room, with its own set of computers, computer operators, and data controllers to support the mission. Support personnel ensure that telemetry data

[1]This work was supported in part by NASA-Goddard Space Flight Center, under Contract NAS5-26952.

and health and safety data are accurately received from the spacecraft and that commanding data is accurately uplinked and loaded into the spacecraft memory.

In the mid-seventies, NASA-Goddard began to integrate some of these dedicated facilities. The primary motivation was clearly cost. Dedicated computers and personnel were often underutilized and thus very costly. The Multisatellite Operations Control Center (MSOCC) was designed to reduce redundant services and personnel by centralizing computer facilities and support personnel. By requiring some degree of standardization of software and operating procedures, individual satellite missions were modularized, so that they could be easily "plugged" into MSOCC for the life of the mission and detached at missions's end. MSOCC is a support facility which provides such services as scheduling and allocation of data processing resources, coordination and configuration of system hardware, computer operations, and data analysis. At this point in time, MSOCC is a labor intensive operation with approximately ten people per shift, three shifts a day, seven days a week. These personnel are in addition to the mission-specific spacecraft controllers who also provide twenty-four hour coverage, seven days a week.

Because of its labor-intensive characteristics and comparatively obsolete hardware, MSOCC was a natural application for the NASA-Goddard automation program. Currently, substantial portions of MSOCC's hardware is configured manually, mission-specific software and data are stored on tapes or disks which require manual mounting and dismounting for each spacecraft contact, and all communication and scheduling is done manually.

A great deal of effort has gone into the planning for the next two generations of MSOCC (see, for example, Multisatellite Operations Control Center-1 (MSOCC-1) 5-year Transition Plan; Automated MSOCC-1 Data Operations Control System Study Task Summary Report; Operational Requirements Study for Automated MSOCC Data Operations Control System). The first generation is to be a semi-automated facility and the second fully automated. A sophisticated and innovative local area network will link

banks of micro, mini, and large-scale computers, allowing automated schedule-driven allocation and configuration of hardware; the system will be designed to be self-checking and capable of automatically reconfiguring appropriate hardware in cases of transmission failure or degraded conditions. Operating under typical conditions, the system is designed to function fairly autonomously requiring little human interference beyond the schedule which is developed off-line.

NASA, however, has a very conservative attitude (Sagan et al., 1980) and both management and mission-specific personnel require that the MSOCC facility be staffed sufficiently to permit continued operations even in the event of severe failure of the automated system.

In this respect, NASA's experience in consistent with that of many other automated control settings. The result of increased automation in th control room and in other previously manual processes does not imply that the human operator is being replaced but rather that his/her role is changing from that of a direct controller to that of a system supervisor who monitors and directs the computers which carry out the moment to moment control functions (Rouse, 1981; Sheridan and Johannsen, 1976). In several MSOCC studies which outline staffing requirements, for example, the absolute number of personnel is increased.

The planned automation of the MSOCC facility, however, has raised a series of human factors issues about the position and function of the human operator in this highly automated environment. Some of the proposed functions and roles for the human seem less than desirable and fail to fully exploit the capabilities of the human-machine interface.

DESIGN ISSUES IN EMERGING AUTOMATION

There are a number of critical design issues for systems which will include some degree of automation. Fundamentally, the question of whether to automate at all must be addressed; subsequently, if the decision to automate some or all of an existing system is made, the system design must carefully examine the human-machine interface, ensuring an appropriate allocation of tasks between the human and computer as well as devising appropriate mechanisms to allow an efficient and effective human-computer dialogue.

Whether to Automate

The decision about whether or not to automate all or a portion of a a control task is important. Reasons for increased automation of control rooms abound. Rouse (1981) suggests a number. Fundamentally, there is a desire for improved performance. By automating a system, it is hoped that the system will support a higher workload (e.g., an airport will support more aircraft with automated air traffic control equipment; MSOCC will support an increased number of satellites; a data processing facility will support more volume or a wider variety of applications tasks). This is closely linked to economic considerations. Replacing humans or augmenting human capability with computer assistance may allow system efficiency to increase with the same staff or decrease system cost by decreasing personnel. Safety and human dignity also motivate the increasing use of automation. Computers replace human operators in tedious, unpleasant, and hazardous tasks (e.g., file maintenance and other clerical tasks, welding, exploration of deep space). Computers provide warning and alarm systems which build a higher degree of safety into the system than was previously available. There are also some things which computers do better than humans and the shift of such tasks to a computer system will also increase system safety and efficiency.

Finally, it must be admitted that sometimes automation is introduced into a system simply because it's there. The hardware for automation has advanced much more rapidly than design principles to guide its implementation (Mitchell, 1980). It is often unclear when or if to implement some facet of automation. Sometimes automation is introduced for the rather fuzzy and certainly indefensible reason that it will make the system modern or "state of the art".

It is this latter reason that is a cause of concern in many systems. Automation, like most other things, has positive and negative attributes. Before its introduction, designers should be explicit and detailed about precisely what advantages an automated system will yield. To the extent possible, these should be quantitative measures of system performance against which the final system can be evaluated. In addition, the potential problems of automation must be squarely faced, evaluated in terms of the benefits, and minimized to the extent possible, by good design.

NASA-Ames hosted an interesting workshop entitled "Human Factors of Flight-Deck Automation-NASA/Industry Workshop" (Boehm-Davis, et al., 1981). The question of whether or not to automate particular functions was critically addressed by a distinguished panel of participants, representing the Man-Vehicle System Research Division of NASA-Ames, the Federal Aviation Administration, the U.K. Royal Air Force, airline companies, aircraft manufacturers, universities, and consulting firms. The participants generally agreed that technology is now sufficiently advanced so that it is theoretically possible to automate most systems and, that though automation has many benefits, it also has some serious drawbacks. These issues, though discussed in the context of flight deck automation are relevant to a wide variety of systems.

The two issues most relevant to MSOCC automation addressed the effect of automation on the human operator. The first pointed out that while an automated system is functioning properly, the human operator is reduced to a system monitor; this role may leave the human, particularly a highly skilled operator, bored and/or

complacent. The second issue is a direct corollary. Personnel in automated systems are often expected to function in two roles: as a system supervisor when the system is functioning automatically and as a direct or manual controller for emergency or degraded conditions. The workshop participants felt that these two roles are not necessarily compatible or complementary; the roles may require two very different sets of skills, two types of knowledge of the system, making it difficult to transition between roles and to adequately perform the tasks required by each role.

Such issues are rarely discussed in the context of automation yet they are critical. Assuming that a system's reliability depends on its weakest component, it is vital to ensure that the human component and the human-machine interface is as reliable as possible. As a result, in evaluating the costs and benefits of proposed automation, the impact on the human component and the human-machine interface must be clearly and thoroughly addressed. Many times automation-induced problems may not outweigh the benefits of automation and to automate may be the most reasonable decision; however, the decision to automate does not abrogate the issues, it merely shifts the burden to the system designers who must eliminate or ameliorate the adverse effects.

Design Issues for the Human-Machine Interface

Assuming that decision has been made to automate some or all of an existing system, the design issues which affect the human interface can be grouped into three categories: the definition of reasonable and meaningful roles for the human operator, allocation of tasks between computer and human system components, and the creation of interfaces which facilitate the human-computer dialogue.

The first two areas are highly related and are likely to be addressed simultaneously in the design process. Creating a meaningful and reasonable role for the human operator is a result of taking a particular perspective at some point in the design process. The perspective is an operator-centered view of the total system aimed at trying to

understand the set of responsibilities assigned to the operator and the dynamics of his/her interaction with the system. One useful tool for gaining this type of perspective is to conduct a task analysis which carefully analyzes the human operator's sequence of tasks; a task analysis includes identification of the individual tasks, the pace of the operations, and underutilized operator resources.

The traditional design approach is often system-centered with the result that though the overall system, at least theoretically, functions adequately the tasks assigned to the operator are those which are "left over" or unable to be automated - the human operator has traditionally functioned as the flexible component in the control loop. It often happens that no one closely examines the overall operator role which the set of left over tasks implicitly defines.

The problems of adequate allocation of responsibility between the human and computer and of defining a reasonable role for the human are particularly problematic in the proposed automation for NASA-Goddard's MSOCC system (Mitchell, 1981). The proposed configuration for an automated MSOCC is an exciting use of technology and will drastically reduce the amount of direct manual intervention in the DOC (Data Operations Control) and computer operations areas. The staffing plan, however, calls for maintaining or possibly increasing the current staff. It is unclear, however, what the eight to ten people per shift will do as the majority of their current functions will be automated. Currently, computer operators transport, mount, and dismount mission-specific software resident on disks and tapes. Under the proposed automation plan, this activity will be fully automated. The responsibilities of the DOC operator are also unclear. Figure 1 depicts a scenario which was given in an MSOCC-1 Operations Requirements Study (TM-81-6098). The scenario represents the anticipated human-computer dialogue during the preparation for a satellite contact. Examination of the scenario reveals that the only active human input is to type the word "GO" as the second to last step in the sequence. An alternative version of this scenario even eliminates this step, assigning the operator to a completely passive, monitoring role.

Analysis of this situation from an operator-centered perspective, raises a number of questions about the reasonableness of the role assigned to the human. It is now true that the MSOCC personnel feel, and in actuality are, underutilized. Because of the nature of their responsibilities, personnel are generally well-trained and skilled, yet they are used at this point in time for relatively insignificant tasks. With the introduction of automation, their tasks are likely to be further reduced, and under the current plan, not augmented by any additional responsibilities.

The MSOCC-1 scenario raises a number of issues about the place of the human operator in an automated system. Often, there is a tendency to retain the operator as the final redundancy in the control loop, to ensure fail-safe conditions. Sometimes this is indicative of an underlying distrust of automation - a questionable premise in a highly automated environment. The consequences of the misgivings can be severe. The first-order impact is cost. Labor costs constitute a large percentage of a system's operating budget. Building a human backup for every system may be a costly proposition, not offset by benefits received.

A second-order impact directly addresses the anticipated benefits. In many automated systems, the tasks allocated to the operator approach the trivial, yet the operator responsibilities are increased - he/she is responsible for the management and control of a very complex system. In the MSOCC-1 example, the operator performs a perfunctory task, and rarely interacts with the system in a meaningful way. Yet in an emergency, the operator is expected to detect and diagnosis the problem, revert to manual control in order to re-establish equilibrium, and it is questionable whether, in this case, he/she will have the capability should the need arise.

The question then arises what should the human do in an automated system - how should tasks be allocated to optimally use both the system's human and computer resources. The answer is neither perfunctory nor simple. Essentially, the role of the human component must be redefined to include a cross-section of meaningful tasks which

enable the human to function as responsible and important component of the system. This redefinition may require the rethinking of the overall system design, reallocation of tasks, and the expansion of the human's responsibilities. Task allocation requires evaluation of the strengths and weaknesses of human and computer [illegible] (Crawford et al., 1977). In order to ensure a reasonable role for the human [illegible] tasks which are done relatively equally well by both human and computers [illegible] assigned to the human. One interesting possibility receiving some attention is [illegible] rather than static, task allocation scheme wherein tasks are allocated bet[illegible] human and computer by determining who has the most available resources at the time that the need for the task arises (Rouse, 1981).

In MSOCC, there is some thought given to expanding the controller's responsibility to include software development/maintenance as well as the supervisory and occasional manual control activities. Augmenting the role definition in this way is a novel approach but may be a satisfactory solution to the problem of assigning skilled personnel to important but tedious supervisory tasks.

Interfaces for the Human-Computer Dialogue

In addition to defining a reasonable role for the human component, the design process in automated systems must address the problem of providing interfaces between the human and computer which facilitates the human's ability to interact with the system in a rapid and effective manner, with as little effort as possible.

One problem that the dual role of the human in automated systems creates is that the human now has two different and perhaps quite disparate functions, potentially requiring two different sets of skills and two different views of the system. In automatic mode, the human needs a high level, integrated overview of the system; whereas in manual mode the human needs to have an understanding of the system which is detailed and thorough and "nitty-gritty".

One of the difficulties of the multiple functions of the human in complex systems is that the varying sets of responsibilities suggest that the operator needs to build up multiple internal models of the system in order to integrate his knowledge of the system and to guide his control actions. A skilled operator in a highly automated system must build up a hierarchy of internal models which encompass a set of system views which vary from a very general and broad system overview to a variety of very specific and detailed models of particular subsystems. Experimental and theoretical research suggests that human understanding of a complex system is guided by an internal or "mental" model of the system built up by the operator over time. The adequacy of the internal model will govern the timeliness and appropriateness of an operator's responses.

One way to facilitate the development of appropriate internal models is through information displays which assist in organizing information, presenting it in modes which facilitate assimilation and integration, thereby reducing the cognitive load on the human operator.

In traditional, hardwired dedicated displays, there was little choice about information display design. Each hardware device, data channel or sensor generated a data item which was individually displayed to the operator (e.g., the battery, the voltage regulator). Control room designers could choose how to display the data (dials, bar graphs, needles, etc.) and could arrange the set of displays on control panels but had no opportunity to selectively display data, to group or aggregate it into higher level summaries. In essence, the displays, due to limitations of technology, were directly tied to the lowest level hardware subsystems (Figure 2). Traditional displays placed a tremendous burden on the human operator. The human was responsible for monitoring, sometimes vast amounts of, displayed data, selecting out relevant items, combining and integrating the low level data into meaningful forms compatible with his higher level information needs.

The advent of computer-based displays eliminated the need for this type of display but not necessarily the practice. Computer-based displays allow data to be filtered, summarized, or aggregated, and displayed in forms only limited by the imagination of the designer. Unfortunately, perhaps because it is easier, many computer-based displays simply use the CRT as a new medium on which to display "the same old data in the same old mode" (see for example, Figure 3). As early as 1975, Braid warned "...there is an alarming tendency . . . to propose replacement of the dedicated conventional instruments by a few dedicated electronic displays . . . Such proposals ignore the flexibility that electronic displays offer."

This is particularly a problem with MSOCC in which multiple display pages are used to display great amounts of low level, hardware specific data (see for example, Figure 3). The controller must monitor these displays, abstract out relevant data, integrate it into forms necessary for his high level decision making. The current displays are very detailed, and completely lack any decision aiding features.

In an automated MSOCC environment, such displays would be even more inappropriate; the low level items would require a good deal of human information processing resources to integrate them into forms needed to support the supervisory control activities of the human operator. The human-computer dialogue issues for MSOCC and other automated systems are really ones of design: how do you use the flexibility of the computer to best create information displays and control. One strategy is to use the flexibility to present information in forms which are compatible with the user's mental model of the system and current role. In highly automated systems, assuming that the operator has at least two sets of internal models: one which allows him/her to function as a monitor and system supervisor, and a second which allows him/her to function as a manual controller, a reasonable suggestion is that perhaps, at the very least, the control room of an automated system ought to have two sets of displays which the operator can choose between: one set giving a high level system

overview, the other, detailed views of individual subsystems. When acting in a supervisory capacity, the high level, overview displays would be used. If more detail is desired, if some problems are suspected, lower level, detailed displays may be accessed.

Computer-based information displays which explicitly support the dual roles of a human in an automated system by presenting hierarchically organized information will be evaluated in the MSOCC control environment. At this time, no sample MSOCC displays are available. However, in order to illustrate some of the concepts which will be used to design the MSOCC displays, an experiment with displays designed using these principles will be described in the next section. The system and displays were developed at The Ohio State University (Mitchell, 1980).

PROTOTYPE HIERARCHIC DISPLAYS

The laboratory system simulated a conveyor system in which engines were routed in and out of various check points. Depicted in Figure 4, the system had engines arriving at Station 1, the diamond labelled "1", which the controller either routed into Buffer Storage or on to Station 2. Once at Station 2, the engine needed to go to the Test Station, Station 6, passing through Stations 3 or 4. Once tested, an engine either was routed out the system throught Station 3 or into the Repair Station, depending on the outcome of the test. The system was highly constrained, allowing no more than one engine at a station or on a conveyor belt at any given time.

Figure 5 contains an information display for this system. This display might correspond to a traditional hardwired display, or to a fairly primitive CRT display in which there was no attempt to exploit the capabilities of the computer-driven display.

In trying to create displays that were more attuned to the human's information needs, a discrete control modelling methodology was used to structure information needs by control functions (Miller, 1979). Individual control activities were grouped together into meaningful control functions with a strategy that might be similar to the way in

which a controller thinks about them. Next, information needed to evaluate the feasibility of actions for particular control functions was identified, examined and structured.

An example may help to illustrate the point. Entry control is a vital control function in the system. It concerns routing engines from the Buffer portion of the system to the Test-Repair loop. Poor entry control strategy will result in poor overall system performance. The fundamental entry control decision is whether or not to release an engine waiting at Station 2 into the Test-Repair loop. Given the system constraints, in order to release an engine, a number of conditions must be met: an engine must be waiting on Station 2, Stations 3 and 4 must both be clear, and both the Test-Repair Feed and Test-Repair Exit conveyors must be clear. There is a natural structure and hierarchy to these conditions which is represented in Figure 6. The presence or absence of an engine at Station 2 is of primary importance. One can not route an engine which is not there. Given the presence of an engine at the Station, the states of the other related system components must be examined. The rules of the system constrain the operator so that an engine may be released from Station 2 only if each of the other four components are in a clear or idle state. Thus, at one level, all the controller is concerned about is whether or not these four components are in the required configuration. This suggests that a higher level information system might be appropriate, one that summarizes the respective statuses of these components and presents it in a form compatible with the human's model of the system. Figure 7 depicts this situation by means of the Test-Repair Feed System. This system is in fact not a system component at all - it is a pseudo-component, an artifact developed to enhance the human-computer dialogue. This system is in the available state only when the four real system components it summarizes are in the appropriate states for a entry control action, i.e., all available, otherwise the Test-Repair Feed System is unavailable.

Figures 8 and 9 depict displayed information in response to an operator's request for entry control information. Figure 8 gives the status of Station 2 but indicates that the Test-Repair Feed System is not in an appropriate state for entry control activity.

Figure 9 illustrates another aspect of the information display system. In this case, there is no engine at Station 2, thus, regardless of the state of the other system components, an entry control action can not be undertaken. As a result, the state of Station 2 is all the information that is displayed. The empty status of Station 2 is a sufficient condition for terminating further consideration of a control action at that point. This type of selective display of information is an attempt to match the processing strategy of an operator and to reduce cognitive load by dynamically limiting and filtering displayed information. If the controller desired to see the more detailed status of the individual system components, he could summon that information (Figure 10). This figure depicts the status of all the system components potentially affecting an entry control decision.

As a third option the controller could request to see all the information, unorganized and unfiltered. This display is given in Figure 5. It is interesting to note that controllers who had access to these three levels of displays rarely took advantage of the lower level displays, preferring the aggregated, summarized, and filtered displays.

This method of information display assumes that system components and information about system components can be structured in a hierarchic form. For a supervisory or monitoring role, the human controller views the system at the higher levels or possibly drops down one or two levels, viewing the systems as a collection of hierarchical subsystems. The switch to a manual mode requires that the controller descend down the hierarchy viewing the system at lower and more detailed levels, likely requiring lower and more detailed information. This notion of hierarchic structure is compatible with much of the system approach to design as well as models of human cognition. The problem, of course, is that displays of this type are not easy to design, the computer is a new display medium and there is little prior experience.

SUMMARY

This hierarchical approach to information display has several advantages. It explicitly forces the system designers to develop a set of human-oriented system models which will guide the design of the displays. By designing the displays around the operator's decision making needs, the displays are likely to become more human-oriented and less hardware-oriented. By trying to provide the appropriate information at the appropriate time, there is likely to be less information displayed at any given time and the quality of the displayed information will require less operator effort to integrate into an assimilatable form. A very pressing problem with contemporary control rooms is that there is just too much information for an operator to be able to quickly, easily and accurately assimilate. Humans are easily overloaded, particularly by the displays of great amounts of irrelevant information (Ackoff, 1967; Seminara et al., 1979). Moreover, human ability to integrate multiple pieces of displayed data into meaningful information is very limited (Rouse, 1973; 1974; 1975). As a result, a reasonable and perhaps vitally necessary direction for research in the area of automated control room design is to develop displays which provide active decision aiding for the modern controller, displays which provide information compatible with the operators' current internal model, which filter out irrelevant information, and which summarize and condense lower level information so as to be in a form suitable for the operator's high level information needs.

Over the next several years, NASA-Goddard will undertake to further develop, implement, and evaluate these display design concepts. They will be implemented on some portion of the automated MSOCC system and actual MSOCC personnel will be used as subjects in an experimental evaluation.

REFERENCES

Ackoff, Russel L., "Management Misinformation System", Management Science, Vol. 14, No. 4, December, 1967.

Boehm-Davis, Deborah A., Renwick E. Curry, Earl L. Wiener, and R. Leon Harrison, Human Factors of Flight - Deck Automation - NASA/Industry Workshop, NASA Technical Memorandum 81260, January, 1981.

Braid, J.M., "Integrated Multi-Function Cockpit Display Systems", AGARD Conference Proceedings No. 167 on Electronic Airborne Displays, NATO Advisory Group for Aerospace Research and Development, AGARD-CP-167, 1975.

Chu, Y.Y., W.B. Rouse, "Adaptive Allocation of Decision Making Responsibility between Human and Computer in Multitask Situations", IEEE Trans. Syst., Man, Cybern., Vol. SMC-9, No. 12, Dec., 1979.

Crawford, Billy M., Donald A. Topmiller, and George A. Kuck, Man-Machine Design Considerations in Satellite Data Management, NTIS ADA041287, April, 1977.

Kelley, Charles R., Manual and Automatic Control, John Wiley and Sons, Inc., NY, NY, 1968.

Miller, R.A., "A Finite State Systems Model of Coordination in Multi-Person Teams", Proceedings of the International Conference on Cybernetics and Society, Denver, Colorado, 1979.

Mitchell, Christine M., The Design of Computer-Based Integrated Information Displays, Doctoral Thesis, The Ohio State University, 1980.

Mitchell, Christine M., Hu... ...achine Interface Issues in the Multisatellite Operations Control Center-1 (MSOCC-1), NASA Technical Memorandum 83826, August, 1981.

NASA, Operations Requirements Study for the Automated Multisatellite Operations Control Center - 1 (MSOCC-1) Data Operations Control System (DOCS), NAS5-24300 (TM-81-6098), 1981.

NASA, Multisatellite Operations Control Center-1 (MSOCC-1) 5 Year Transition Plan, NAS5-24300 (TM-81-6090), 1981.

NASA, Automated MSOCC-1 Data Operations Control System Study Summary Report, NAS5-24300 (TM-81-6128), 1981.

Rasmussen, Jens and William B. Rouse (Eds.), Human Detection and Diagnosis of System Failures, Plenum Press, NY, NY, 1981.

Rouse, W.B., "A Model of the Human in a Cognitive Production Task", IEEE Trans. System, Man, and Cybernetics, SMC-3, Sept. 1973.

Rouse, W.B., "A Model of the Human as a Suboptimal Smoother", Proceedings of 1974 IEEE Decision and Control Conference, Nov. 1974.

Rouse, W.B., "Design of Man-Computer Interfaces for On-Line Interactive Systems", Proceedings of the IEEE, Vol. 63, No. 6, June 1975.

Rouse, W.B., "Adaptive Allocation of Decision Making Responsibility between Supervisor and Computer", in Monitoring Behavior and Supervisory Control, T.B. Sheridan and G. Johannsen (Eds.), Plenum Press, New York, 1976.

Rouse, W.B., "Human-Computer Interaction in Multi-Task Situations", IEEE Trans. System, Man, and Cybernetics, SMC-7, No. 5, May 1977.

Rouse, W.B., "Human-Computer Interaction in the Control of Dynamic Systems", Computing Surveys, Vol. 13, No. 1, March 1981.

Rouse, W.B., "Models of Human Problem Solving: Detection, Diagnosis and Compensation for System Failures", submitted for publication, 1982.

Sagan, Carl and Members of the NASA Study Group, Machine Intelligence and Robotics: Report of the NASA Study Group, March, 1980.

Seminara, J.L., S. Eckert, S. Seedenstein, Wayne Gonzalez, R. Stempson and S. Parsons, Human Factors Methods for Nuclear Control Room Design, EPRI NO-1118-54, June 1979.

Sheridan, T.B. and G. Johannsen (Eds.), Monitoring Behavior and Supervisory Control, Plenum Press, New York, 1976.

HUMAN - COMPUTER DIALOGUE FRAGMENT

FROM THE PROPOSED AUTOMATED MSOCC-1

Statement Source	Time Tag	Item Number	Control Statements	Comments
SCHEDULE	185:20:16:10	1	DI SAMA	Display PROC SAMA
SCHEDULE		1.1	CO LN1 TO TAC6	These items displayed from PROC
SCHEDULE		1.2	CO TAC6 TO AP5	
SCHEDULE		1.3	CO AP5 TO KCRT (MOR1)	
SCHEDULE		1.4	CO AP5 TO SCR (MOR1)	
SCHEDULE		1.5	DLL SAMASYS TO AP5	Downline load SAM-A software to AP5
SCHEDULE		2	WAIT	Wait for operator intervention
KEYBOARD		3	GO	Operator key-in
SCHEDULE		4	S SAMA	EXEC SAMA

DOCS Operator-Computer Interaction Scenario for Automated MSOCC-1 Operations

FIGURE 1

PORTION OF A 747 PILOT PANEL

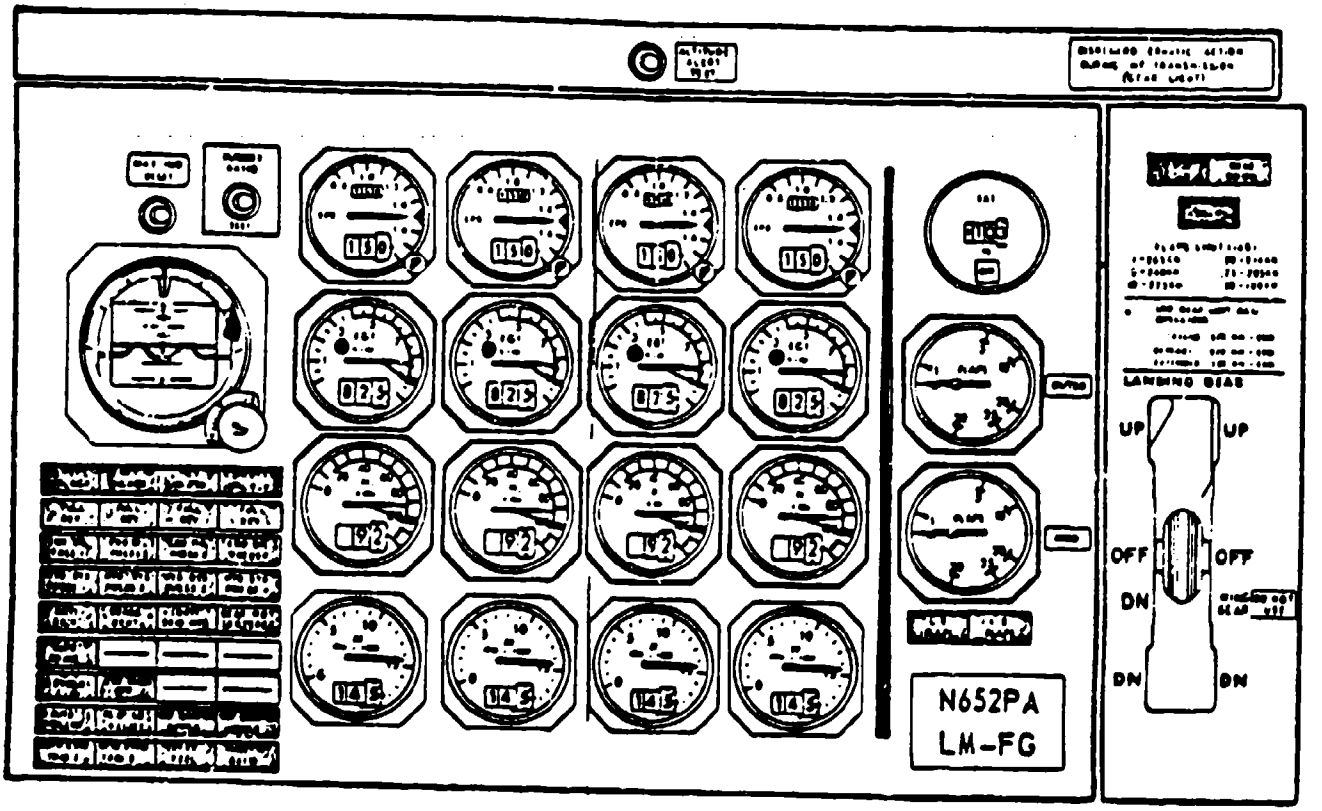
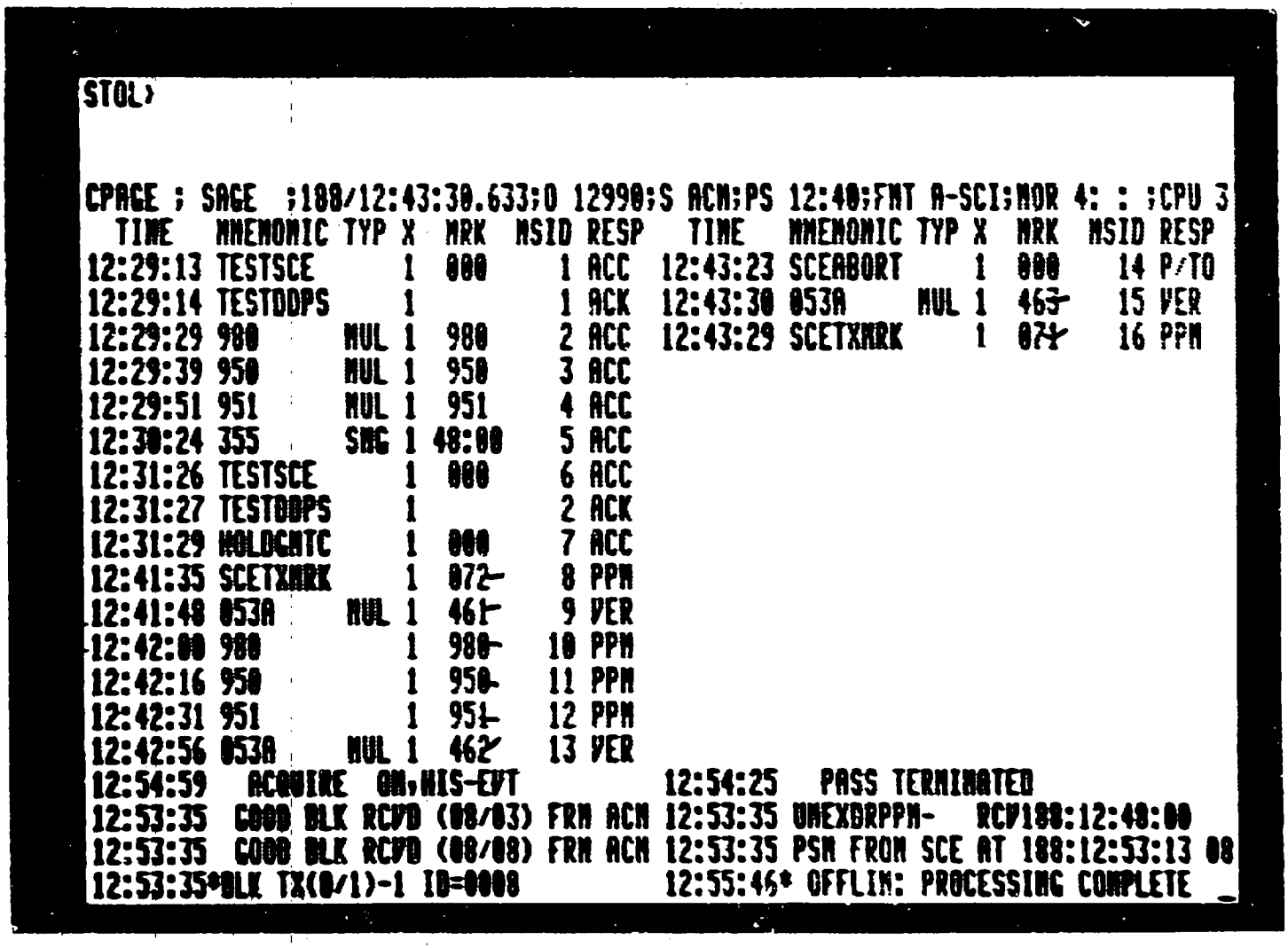


FIGURE 2



```
STOL>


CPAGE ; SAGE  ;188/12:43:30.633;0 12990;S ACN;PS 12:40;FNT A-SCI;MOR 4: : ;CPU 3
  TIME   MNEMONIC TYP X  MRK  MSID RESP   TIME   MNEMONIC TYP X  MRK  MSID RESP
12:29:13 TESTSCE      1  000     1 ACC  12:43:23 SCEABORT     1  000    14 P/TO
12:29:14 TESTDDPS     1             1 ACK  12:43:30 053A     MUL 1  463    15 VER
12:29:29 980      MUL 1  980     2 ACC  12:43:29 SCETXMRK     1  072    16 PPM
12:29:39 950      MUL 1  950     3 ACC
12:29:51 951      MUL 1  951     4 ACC
12:30:24 355      SNG 1 48:00    5 ACC
12:31:26 TESTSCE      1  000     6 ACC
12:31:27 TESTDDPS     1             2 ACK
12:31:29 HOLDCNTC     1  000     7 ACC
12:41:35 SCETXMRK     1  072     8 PPM
12:41:48 053A     MUL 1  461     9 VER
12:42:00 980          1  980    10 PPM
12:42:16 950          1  950    11 PPM
12:42:31 951          1  951    12 PPM
12:42:56 053A     MUL 1  462    13 VER
12:54:59   ACQUIRE  ON,HIS-EVT          12:54:25   PASS TERMINATED
12:53:35  C000 BLK RCVD (00/03) FRM ACN 12:53:35 UNEXDRPPM-   RCV188:12:48:00
12:53:35  C000 BLK RCVD (00/08) FRM ACN 12:53:35 PSN FROM SCE AT 188:12:53:13 00
12:53:35*BLK TX(0/1)-1 ID=0008          12:55:46* OFFLIN: PROCESSING COMPLETE
```

FIGURE 3

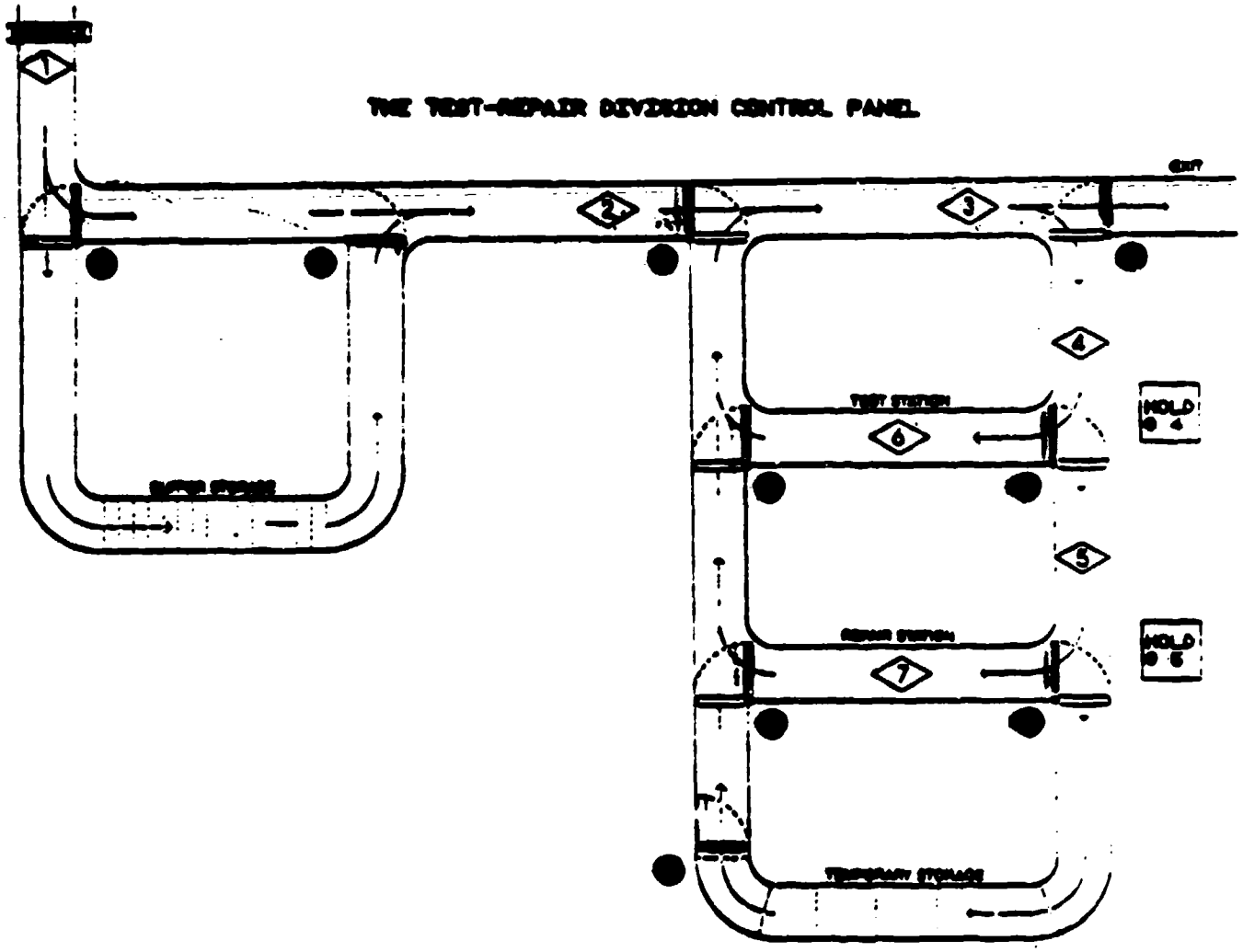


FIGURE 4

TEST-REPAIR DIVISION INFORMATION SYSTEMS

EXPEDITE: TYPE 1 ENGINES

STATION 1:	CLEAR	PRODUCTION BUFFER TRACK:	CLEAR
STATION 2:	TYPE 1 ENGINE	TEST-REPAIR FEED TRACK:	CLEAR
STATION 3:	CLEAR	TEST-REPAIR EXIT TRACK:	CLEAR
STATION 4:	ENGINE HELD (1,T)	HOLD AT 4:	ON
STATION 5:	ENGINE HELD (2,R)	HOLD AT 5:	ON
STATION 6:	BUSY(TYPE 1)	TEMPORARY STORAGE LOCK:	OFF
STATION 7:	LOCKED(1,C)		

BUFFER STORAGE: 1,2,2,2,1,1,2,2,2,2,2,2

TEMPORARY STORAGE: (1,C),(1,R),(1,C),(1,C)

SWITCH 1: CLOSED TO STATION 2

FIGURE 5

FIGURE 6

ENTRY CONTROL INFORMATION

INPUTS: STATION 2, TEST-REPAIR FEED TRACK, TEST-REPAIR EXIT TRACK, STATION 4, STATION 3.

STATES: ENTER, STANDBY

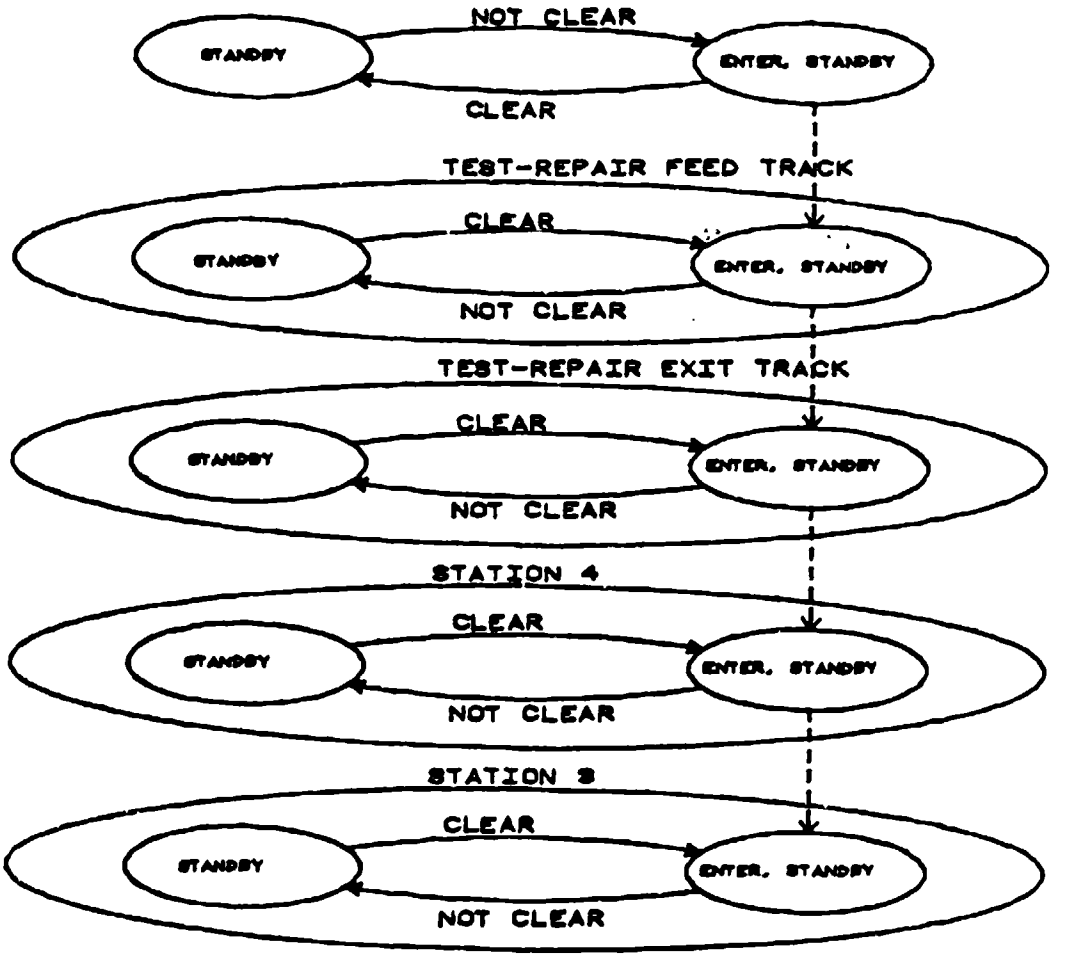
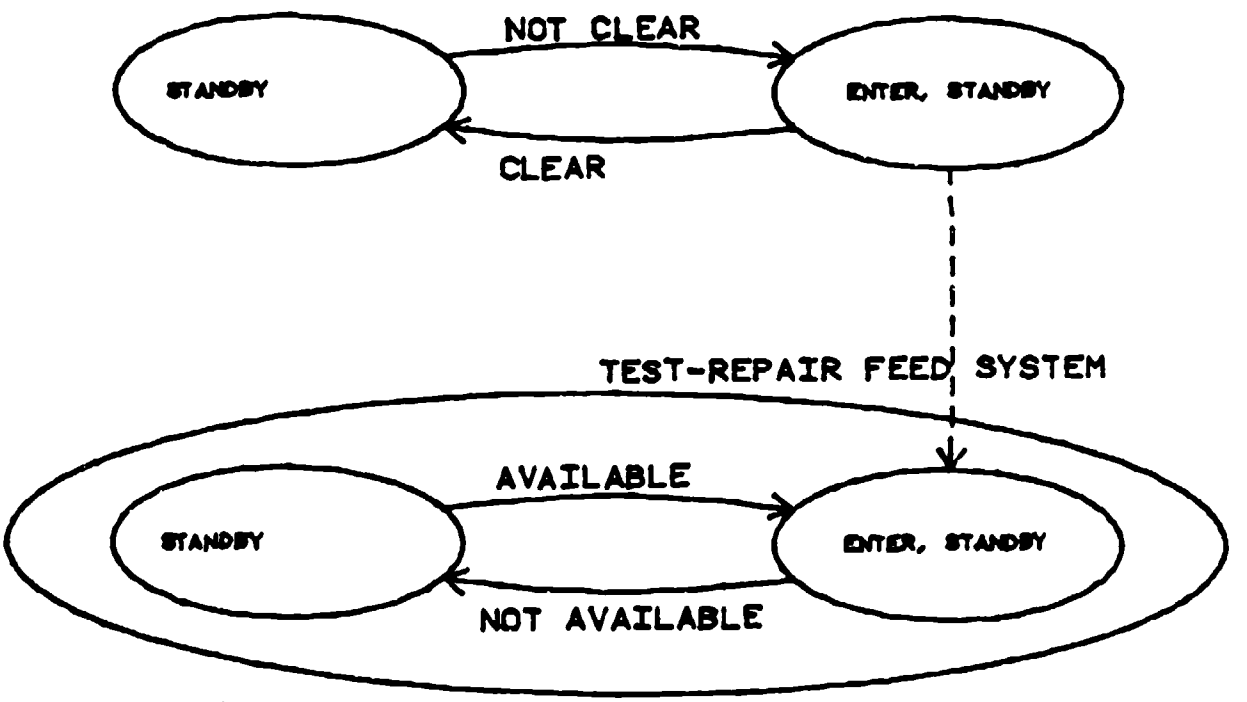


FIGURE 7

ENTRY CONTROL INFORMATION

INPUTS: STATION 2, TEST-REPAIR FEED SYSTEM.

STATES: ENTER, STANDBY



HIERARCHIC INTEGRATED INFORMATION DISPLAY

EXPEDITE: TYPE 1 ENGINES

ENTRY CONTROL INFORMATION

STATION 2: TYPE 2 ENGINE

TEST-REPAIR FEED SYSTEM: NOT AVAILABLE

FIGURE 8

HIERARCHIC INTEGRATED INFORMATION DISPLAY

EXPEDITE: TYPE 1 ENGINES

ENTRY CONTROL INFORMATION

STATION 2: CLEAR

FIGURE 9

GROUPED PRIMITIVES INTEGRATED DISPLAY

EXPEDITE: TYPE 1 ENGINES

ENTRY CONTROL INFORMATION

STATION 2: CLEAR

STATION 3: CLEAR

STATION 4: CLEAR

TEST-REPAIR FEED TRACK: BUSY

TEST-REPAIR EXIT TRACK: BUSY

FIGURE 10

AD P001633

KEYBOARD DESIGN VARIABLES IN DUAL-TASK MODE SELECTION

by
Mark D. Hansen
Texas A&M University
College Station, Texas 77843

ABSTRACT

Research was performed to determine the optimal and desirable qualities of a keyboard interfaced with a remote display. The keyboards were two 3 x 3 keyboards, one with buttons 5.16 mm in diameter and the other 12.7 mm square. The experimental task involved subjects centering a set of crosshairs on a cursor on a CRT (with their right hand), while simultaneously selecting modes (with their left hand) on the keyboards without observing the keyboards. The experimental design was as follows: (1) two keyboards; (2) three orientations of the keyboard (0, 15 and 35 degrees) from the horizontal and (3) with and without gloves (approved Air Force flight gloves). This design was conducted as a 3 x 2 x 2 repeated measures factorial design. Performance measures were response time on the keyboards, selection errors on the keyboard and tracking scores.

Results of this study indicate that mean response time was found to be significantly lower on the large keyboard when examining both keyboards as a whole and with and without gloves. However, error rate was not found to be a significant factor across conditions. In addition, tracking scores were found to be significant, the tracking only versus the combined task condition. No specific angle was found to be significant, but pilots had a preference for the 15 degree (75%) over the other two keyboards. Thus the results of this study indicate the use of the enlarged keyboard in the 15 degree orientation when using gloves.

INTRODUCTION

There are numerous problems involved in aircraft cockpit design. Two problems of experimental interest are display location and panel design. With the advent of single function electronic displays and digital computers, console space has become a distinct problem. These problems are as follows: (1) there is a limit to the space available for displays on aircraft consoles; (2) there is a limit to the number

of functions to be displayed; (3) these functions displayed are far below the actual number of functions a pilot must perform; (4) in a high G setting, the display would most likely be located outside the pilot's reach and vision envelope.

Various military contractors have designed numerous cockpit input devices. These input devices include keyboards, light pens, joysticks, selection buttons and touchwires. Many of these input devices have inherent problems when incorporated in a high G environment. These problems in many cases cause the system to become inoperable in a crucial high G situation.

Armstrong and Poock (1981) at the Naval Postgraduate School at Monterey, California investigated voice recognition system performance. They identified task duration and increased concurrent operator mental or motor workload to adversely affect performance of voice recognition systems even in simple phrases such as "VHF" and "UHF". Lovesey recommended an interactive display for flight functions to be controlled. An interactive display incorporates numerous functions into one display, thus allowing the pilot to use additional information by storing and retrieving it whenever necessary. However, Newman and Welde (1981) delineated numerous problem areas with these devices (i.e., the field of view is too limited). Researchers at Wright Air Base studied several different multi-function keyboard designs in the F-111 and A-7D fighter aircraft cockpits. Their designs were only minimally acceptable for the situation involving a "look but can't touch and the touch but can't see" due to the design needed in a high G tilt-back seat (Bateman, Reising, Herron, and Calhoun, 1978).

Hottman (1980) developed an acceptable solution to this problem by evaluating three multifunction keyboards, interacting with a digital computer which selected and operated many of the controls for non-flying tasks on modern military aircraft. Of the three input devices evaluated (a 3x3 keyboard, a 2x4 keyboard, and an eight-position joystick), the 3x3 keyboard was determined as optimal. In arriving at this decision he neglected to examine inclination angle of the keyboard (from the horizontal) or the effect of gloves.

Research was performed concerning the design of controls in gloved conditions. Huchingson (1972) examined the use of gloves and various pressure suits (ventilated and pressurized) and found a decrement in dexterity, gripping strength and arm speed as compared to the bare hand condition. Chapanis (1972) also examined the design of controls, however none of his design recommendations included specifications when using gloves.

The keyboard design was based on the classic study conducted by Fitts and Seeger (1953). Fitts and Seeger exhibited matched stimulus (SR) sets, 8 lights every 45° on a coordinate circle, to determine which resulted in the best overall effectiveness. The best SR set consisted of the above described stimulus set, and the response set consisted of eight pathways radiating from a central point like spokes

on a wheel. Another principle utilized in the design of these keyboards was determined by Alden, Daniels and Kanarick (1972). These principles involved standards such as: key dimensions: diameter-1.27 cm, center to center spacing-1.81; force and displacement: force-0.9 to 5.3 oz., displacement-0.13 to 0.64 cm, keyboard size: a miniaturized keyboard for touch operations causes a decrement in performance; and entry method: chord entry is faster than sequential entry for tasks like mail sorting.

In measuring performance the most widely used objective measure is by secondary tasks (Wierwille and Williges, 1978; Brown, 1978; Ogden, Levine and Eisner, 1979; and Chiles and Alluisi, 1979). The use of secondary tasks is based on the assumptions that an upper bound exists on the ability of the human operator to process information and that the mental resources which form this limited capacity can be shared among tasks. The methodology requires the operator to perform an extra task along with a primary task of interest. Performance measures are obtained by comparing single and dual task performance. However, a second type of workload theory (Allport, Antonis and Reynolds, 1972; Damos and Wickens, 1980; and Hirst, Spelke, Reaves, Caharack and Neisser, 1980) assumes that the information processing system has several structure-specific capacities. This model is used to argue that the secondary task methodology is of limited value because obtained workload measures would be dependent upon the degree to which particular primary and secondary tasks share common mental functions, and therefore, common capacities.

SUBJECTS

Twenty-four pilots located in the Bryan-College Station Area participated as subjects in this study. Pilots' ages ranged from twenty-two to fifty-two years of age. Pilot experience ranged from one-hundred to ten thousand hours. In addition, pilot background ranged from private to commercial to military.

DESIGN

This study was conducted as a 3x2x2 repeated measures factorial and presentation of treatments followed a 12x12 latin square arrangement. The independent variables were: (1) two keyboards in a 3x3 arrangement (one with 5.16 mm diameter (small) keys and one with 12.4 mm square (large) keys); (2) the angle of orientation of the keyboard from the horizontal (0°, 15° and 35°); and (3) with and without USAF approved flight gloves. Dependent variables were tracking score on a flight simulation task, response time in selecting flight modes, and the error rate in selecting flight modes.

APPARATUS

The mode selection task was performed on a Pet/Commodore computer which was electronically interfaced to two keyboards with momentary-on pushbuttons.

The tracking task was performed on an Apple II computer which was electronically interfaced to an Everest and Jennings joystick.

PROCEDURE

A mode was defined as a three letter acronym of a non-flying task. Modes were located in the same position throughout the study and were as follows: VHF (radio), FLT (flight), NAV (navigation), FPR (fuel pressure), UHF (radio), OPR (oil pressure), RPM (revolutions per minute) and ALT (altitude). A sequence was defined as a group of three modes. Response time was defined as the amount of time to key in a sequence of modes presented aurally to the subject. Selection errors was defined as selecting the incorrect mode or selecting modes in the incorrect order.

Each presentation of treatments was divided into sub-sessions: (1) tracking alone and, (2) tracking and performing the mode selection task simultaneously. Each was randomly assigned first and lasted for two minutes.

Each subject made five different selections of sequences on both keyboards in each sub-session. The positions on the displays were a spatial analog of the keyboard the subject was using. An example of a mode selection sequence was as follows: "select VHF, RPM, NAV." The pattern would appear and the subject would look at the CRT for each mode and select them accordingly. After selecting the modes (the computer gave them feedback by beeping when each mode was selected), the CRT went balnk. The computer then generated a new sequence (each sequence was generated such that each sequence did not constitute one complete row or column) which was then transmitted to the subject at the appropriate time.

The operation of both tasks simultaneously was identified as a dual-time sharing task. The tracking task was identified as the most important, however, the mode selection was identified as nearly as important.

RESULTS

Using a univariate analysis of variance for repeated measures indi-

cated that response time using the large keyboard was significantly briefer than for the small keyboard (pr > .05).

A summary of significant statistics can be seen in Table 1. Further investigation using a Duncan's Multiple Range Test (p = .05) using the large keyboard without gloves was significantly briefer than for the large keyboard with gloves and the small keyboard with and without gloves. Surprisingly, neither keyboard size nor gloves affected the number of keying errors. Tracking accuracy was found to be significantly higher in the tracking only condition than in the combined task condition. Angle of orientation of the keyboard did not significantly affect response time, keying errors or tracking scores at the .05 level of significance. However, pilots preferred the 15° angle (75%) as assessed by a post-test questionnaire. They also preferred the large keys (75%).

Table 1. Summary of Significant Statistics

Mean Response Time as a Function of Keyboard Size	
	Response Time (sec)
Large Keyboard	2.871
Small Keyboard	3.112
Mean Response Time as a Function of Keyboard Size With and Without Gloves	
	Response Time (sec)
Large Keyboard With Gloves	2.6612
Large Keyboard Without Gloves	3.0282
Small Keyboard With Gloves	3.1458
Small Keyboard Without Gloves	3.0774
Mean Tracking Scores for Tracking Only and Tracking During Mode Selection	
	Tracking Score
Tracking Only	4688.79
Tracking and Mode Selection	4028.66

CONCLUSIONS AND RECOMMENDATIONS

The results of this study indicate that the keyboard with the large keys (as assessed by response time and questionnaire) in the 15° angle (also assessed by questionnaire) should be used in selecting modes with USAF approved flight gloves from a remote display.

Recommendations for future research are:

1. To provide for a higher simulation with CRT screens being placed one on top of the other.
2. To evaluate another 3x3 keyboard to determine optimality (i.e., moving the keys closer).
3. To evaluate another 3x3 keyboard with the middle button to determine if tactual feedback aids in performance.
4. To further investigate angle of inclination to determine if there is an optimal angle.

REFERENCES

Alden, D.G., Daniels, R.W., Kanarick, A.F., Keyboard design and operation. A review of the major issues. Human Factors, 1972, 275-294.

Allport, D.A., Antonis, B., Reynolds, P. On the division of attention: a disproof of the single channel hypothesis. Quarterly Journal of Experimental Psychology, 1972, 24, 225-235.

Armstrong, J.W. and Poock, G.K. Effect of Task Duration on Voice Recognition System Performance. Naval Postgraduate School, Monterey, California, NPS55-81-017, 1981.

Bateman, R.P., Reising, J.M., Herron, E.L., Calhoun, G.L. Multifunction keyboard implementation study. Wright-Patterson AFB, Ohio, AFFDL-TR-78-197, 1978.

Brown, D.I. Dual Task Methods of Assessing Workload. Ergonomics, 1978, 21, 221-224.

Chapanis, A. Design of controls, in H.P. Van Cott and R.G. Kinkade (eds.), Human engineering guide to equipment design, U.S. Government Printing Office, Washington, D.C., 1972, chap. 8.

Chiles, W.D. and Alluisi, E.A. On the specification of Operator or Occupational Workload with performance-measurement methods. Human Factors, 1979, 21, 515-528.

Damos, D.L. and Wickens, C.D. The identification and transfer of training in timesharing skills. Acta Psychologia, 1980, 46, 15-39.

Fitts, P.M. and Seeger, C.M. S-R compatibility: spatial characteristics of stimulus and response codes. Journal of Experimental Psychology., 1953, 46, 199-210.

Hirst, W., Spelke, E.S., Reaves, C.C., Caharack, G. and Neisser, U. Dividing attention without alternation or automaticity. Journal of Experimental Psychology: General, 1980, 109, 98-117.

Hottman, S.B. Comparison of discrete control devices to select menus on an interactive display. Unpublished masters thesis, Texas A&M University, College Station, Texas, 1980.

Huchingson, R.D. Effects of Pressure Suits on Seven Psychomotor Skills. Perceptual and Motor Skills, 1972, 34, 87-92.

Lovesey, E.J. The instrument explosion - a study of aircraft cockpit instruments. Applied Ergonomics, 1977, 8, 23-30.

Newman, R.L. and Welde, W.L. Head-Up Displays in Operation: Some Unanswered Questions. First Symposium on Aviation Psychology, 1981, 119-120.

Ogden, D.G., Levine, J.M., and Eisner, E.J. Measurement of Workload by Secondary Tasks. Human Factors, 1979, 21, 529-548.

Williges, R.C. and Wierwille, W.W. Behavioral measures of Aircrew mental workload. Human Factors, 1979, 21, 549-574.

ACKNOWLEDGEMENTS

I wish to express my thanks to Dr. R. Dale Huchingson for his guidance and professional advice in formulating this research study. I also wish to express my thanks to Dr. Charles M. Stoup for his assistance in formulating the computer data analysis programs and to Dr. James K. Hennigan for his helpful suggestions throughout this study. A very special thanks to Stephen B. Hottman for his timely and invaluable assistance in the set-up of the experimental equipment, computer programs and professional advice, without which this study could not have been performed. I would also like to thank Dr. Robert P. Bateman for his advice and suggestions throughout this study, and my graduate career at Texas A&M University. I must also thank Miss Denise Clugey for her many long hours of typing and revising this thesis.

AD P001634

AN ANALYSIS OF CONTROL RESPONSES AS A FUNCTION OF DISPLAY DYNAMICS IN PERSPECTIVE AIRCRAFT DISPLAYS WITH PREDICTION

Jeffery L. Maresh
Department of Aeronautical and Astronautical Engineering

Richard S. Jensen
Department of Aviation

The Ohio State University
Columbus, Ohio

ABSTRACT

This paper describes plans to analyze flight control response data in conjunction with RMS error information in an attempt to determine how pilots process information in a simulated curved approach and landing task using predictive information in a perspective display. An investigation of predictive algorithms in which the parameters are varied across the levels of the system control hierarchy was conducted. Preliminary empirical data indicating a relatively long response period for flight controls tends to suggest that the nature of the mental processes involved are cognitive rather than perceptual. Results of this experiment also indicate that pilot control response is directly related to the hierarchical order of the predictor algorithm. Additional analyses are needed to determine how those results may impact the information processes used in these tasks. Because the data were recorded at one-second intervals, an additional experiment is planned in which data will be recorded at more frequent intervals to validate the findings.

A SIMULATOR STUDY

In a flight simulator study Jensen (1981) demonstrated an effective application of various predictor algorithms to visual guidance and flight-path prediction in forward-looking true-perspective flight displays using a computer generated CRT display as shown in Figure 1. In this study 18 professional pilots each flew four curved landing approaches under varying simulated windshear conditions with each of 18 display co 'gurations. The displays represented parametric combinations of fi second-, and third-order predictive flight path algorithms and four proportions of pursuit prediction versus compensatory quickening, plus five hybrid display configurations and a conventional crosspointer display.

-----Figure 1 here-----

Figure 1. Experimental perspective predictor display.

RMS lateral error results, Figure 2, indicate a significant improvement with increasing proportions of pursuit prediction. Second- and third-order predictors were significantly better than first-order predictors but were not significantly different from each other. Best vertical control, Figure 3, was achieved with an intermediate combination of prediction and quickening. Compared with the coventional crosspointer, the best experimental displays reduced RMS lateral error by a factor of ten. Vertical error was also reduced but by a lesser amount.

-----Figure 2 here-----

Figure 2. RMS lateral error as a function of percent prediction and computational order.

-----Figure 3 here-----

Figure 3. RMS vertical error as a function of percent prediction and computational order.

However, the most interesting and important result was the large statistically significant differences in aileron, rudder, and elevator control activity between displays with first-, second-, and third-order prediction algorithms. Figure 4 shows a power spectral analysis of aileron control movements from a preliminary study of six of the displays tested in the experiment. It can be seen from these data that the control strategies vary to a considerable extent from one display to another. The conventional crosspointer (CP) for example shows large amplitude movements at about .075 HZ. The third-order predictor (3rdP) and the third-order predictor with 33 percent quickening (3rdPQ) result in very small amplitude movements throughout the frequency spectrum. Pilots apparently were overcontrolling in the first case and undercontrolling in the second. Similar results were found for elevator and rudder control responses. These results may be considered surprising since the pilots were all highly skilled professionals whose response patterns should have been well established for aircraft control prior to the experiment.

-----Figure 4 here-----

Figure 4. Power spectral analysis of aileron control responses.

The data suggest the need for rethinking the task of continuous control of higher-order systems. In all cases the major period of aileron control activity was on the order of 8-12 seconds. Data for elevator control (Figure 5) indicate a major period on the order of 3-10 seconds depending on the display dynamics used. It may be that these major control decisions are made discretely at the cognitive level, whereas the trial and error searching for more information as seen in the lower-order predictor displays and in the conventional crosspointer display is made in a continuous fashion. Clearly, predictive display dynamics have a considerable effect on the amount and types of cognitive processing (and control movements required for perceptual processing) required in the control of an aircraft on a curved approach.

-----Figure 5 here-----

Figure 5. Power spectral analysis of elevator control responses.

Flight Task Analysis

In an analysis of the pilot's control task, a significant component that can be identified is cognitive prediction resulting from the mental transformations required by the aircraft control hierarchy. For example, to correct an error (due to disturbance inputs, pilot control errors, or course curve adjustments) in the horizontal direction, the pilot estimates the amount of heading change that is necessary to effect an optimum correction. Given this estimate of the necessary heading adjustment, the pilot estimates the optimum bank for that heading change. Finally, given an estimate of the necessary bank and pitch, the pilot estimates the amount of aileron control movement appropriate for that amount of bank change. Considering the vertical component that must be controlled simultaneously, the cognitive prediction task is very demanding. This task is the major element to be learned initially in instrument approach instruction and it is often the first to be forgotten in the loss of instrument "currency." If the amount of control power is proportional to the amount of cognitive processing involved, then the most effective display would be that which causes pilots to produce the least amount of control power and error in both the lateral and vertical modes.

The Perception-Action Cycle

The research work by Owen and his associates (Owen, Warren, Jensen, Mangold, and Hettinger, 1979) suggests that perception is a active process. They have advanced the idea that a considerable amount of the control activity input by the operator of a vehicle is made for the purpose of obtaining information rather than correcting error. This may explain the phenomena that we see in the lower-order display where a high level of control activity is evidenced. These results suggest that a detailed analysis of control style is essential for an understanding of

the reasons for the advantage of one display design over another in higher order systems.

Additional Data Analyses

To determine the effect of these factors additional analyses are needed of the control response data. Power spectral analyses of the data obtained from each of the 1296 approaches of the Jensen (1981) experiment are being made in an attempt to determine the relationship between the hierarchical levels of information in the predictor algorithm and control response strategies.

A major obstacle to these analyses is the fact that the sampling rate for the data was 1 Hz. Most control theorists believe that, because humans can make control responses at the rate of 5 Hz, data should be sampled at a rate not less than 10 Hz to avoid losing detail in the waveform. The reason data were not recorded at this high rate in the Jensen experiment (and many others as well) was that the limited storage capacity of the computer made it prohibitive.

On the other hand, a number of factors suggest that this high sampling rate may not be needed. First, if the primary interest is the analysis of discrete cognitive responses with a period between decisions of several seconds, a sampling rate of 10 Hz amounts to overkill. Second, consideration should be given to the frequency response of the pilot-aircraft combination. During the approach phase of flight, aircraft dynamic response characteristics are typically a fraction of what they are capable of during high speed flight. In addition, on an ILS approach the pilot usually limits the bank angle of the aircraft to a fraction of the what he normally uses for normal bracketing procedures. The experienced professional pilot will usually make his course corrections only after making estimates of the outcome of each control input. Other factors that may cause variability in control response strategies are pilot experience, currency, and atmospheric turbulence. These factors considered together could produce the low frequency of control activity found in the preliminary experiments described earlier.

Therefore, a validation experiment is planned to determine the required sampling rate acceptable for the analysis of aircraft approach and landing tasks. The point at which the accuracy becomes degraded due to insufficient sampling rate can be determined experimentally. A T-40 twin-jet flight simulator will be used to fly several ils approaches using a conventional electro-mechanical cross-pointer display. Data will be sampled at 10Hz using a Harris/6 Computer. These data will be analyzed using power-spectral analysis initially at 10 Hz, then at 5, 2 and 1 Hz and compared with each other to determine the amount of loss of accuracy as the sampling rate is decreased.

Interpreting the results of the power spectral analysis, one must consider the magnitude of control power in the context of error produced. Lower control power is indicative of lower cognitive processing (and, perhaps lower mental workload), but it may also be indicative of undercontrolling if accompanied by high levels of steering error. A correlation of control power and RMS error would be useful in this regard.

The ideal flight display is one that allows the pilot to deliver the aircraft to the runway with the least amount of error and mental workload. To evaluate these parameters, one must be able to correlate control power with error at various positions of the approach. One can observe from the lateral mode spectral and error analysis for example, that the lowest power does not necessarily produce the lowest error. It is also evident that superiority in the lateral mode does not necessarily produce similar results in the vertical mode. To establish an order of merit for the display factors of interest each of these parameters must be used.

REFERENCES

Jensen, R.S. Prediction and quicking in perspective flight displays for curved landing approaches. Human Factors, 1981, 23(3), 355-363.

Owen, D. H., Warren, R., Jensen, R. S., Mangold, S. J., and Hettinger, L. J. Optical information for detecting loss in one's own forward speed. Acta Psychologica, 1981, 48, 203-213.

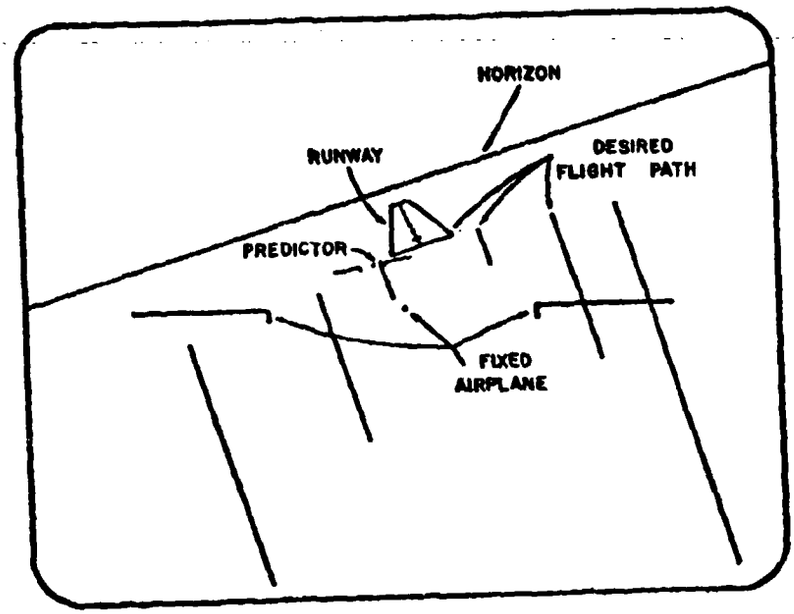


Figure 1. Experimental perspective predictor display.

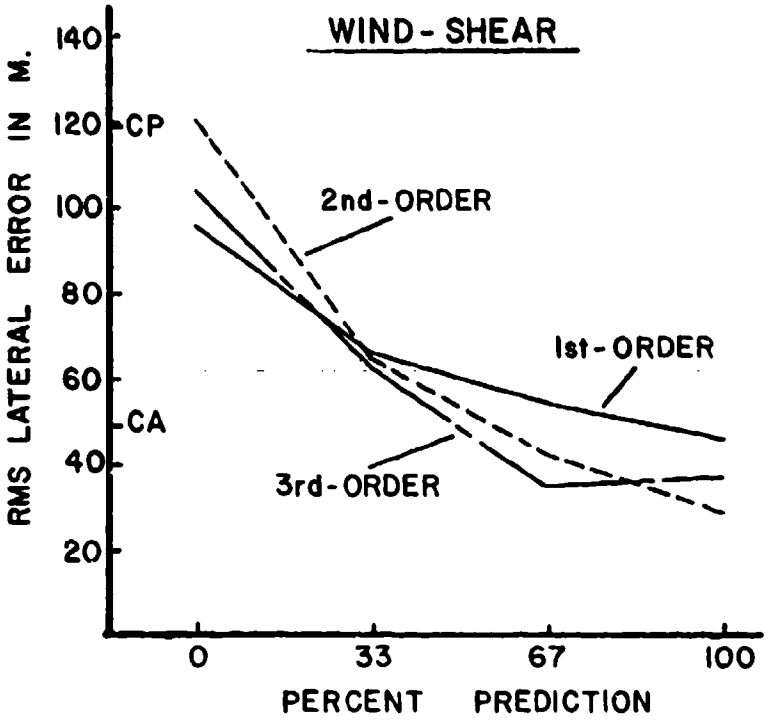


Figure 2. RMS lateral error as a function of percent prediction and computational order.

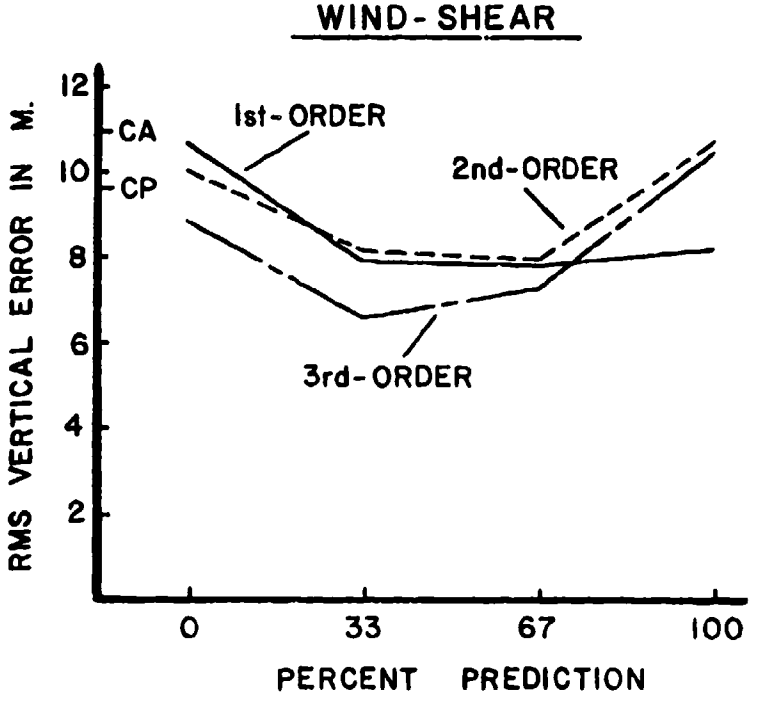


Figure 3. RMS vertical error as a function of percent prediction and computational order.

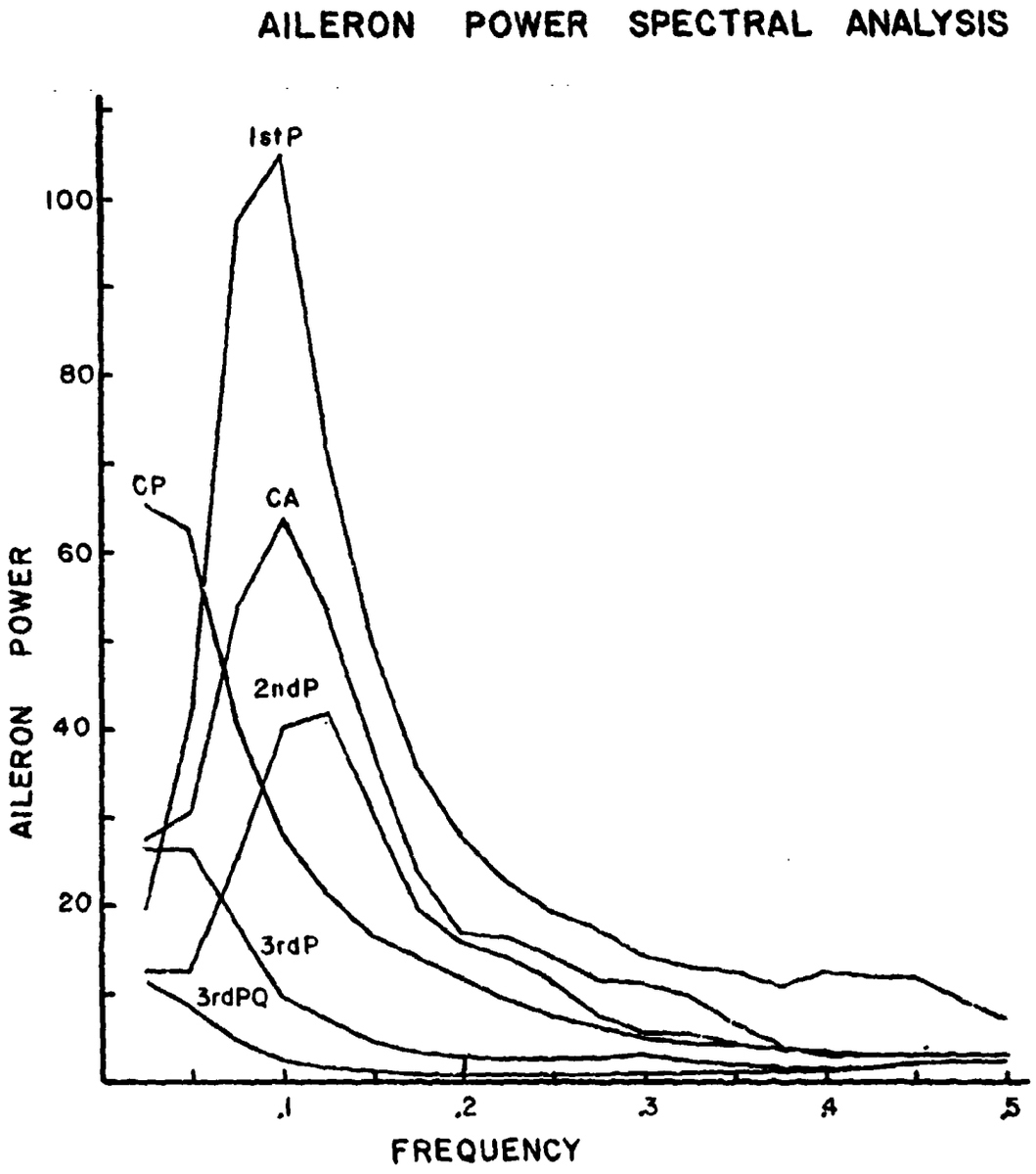


Figure 4. Power spectral analysis of aileron control responses.

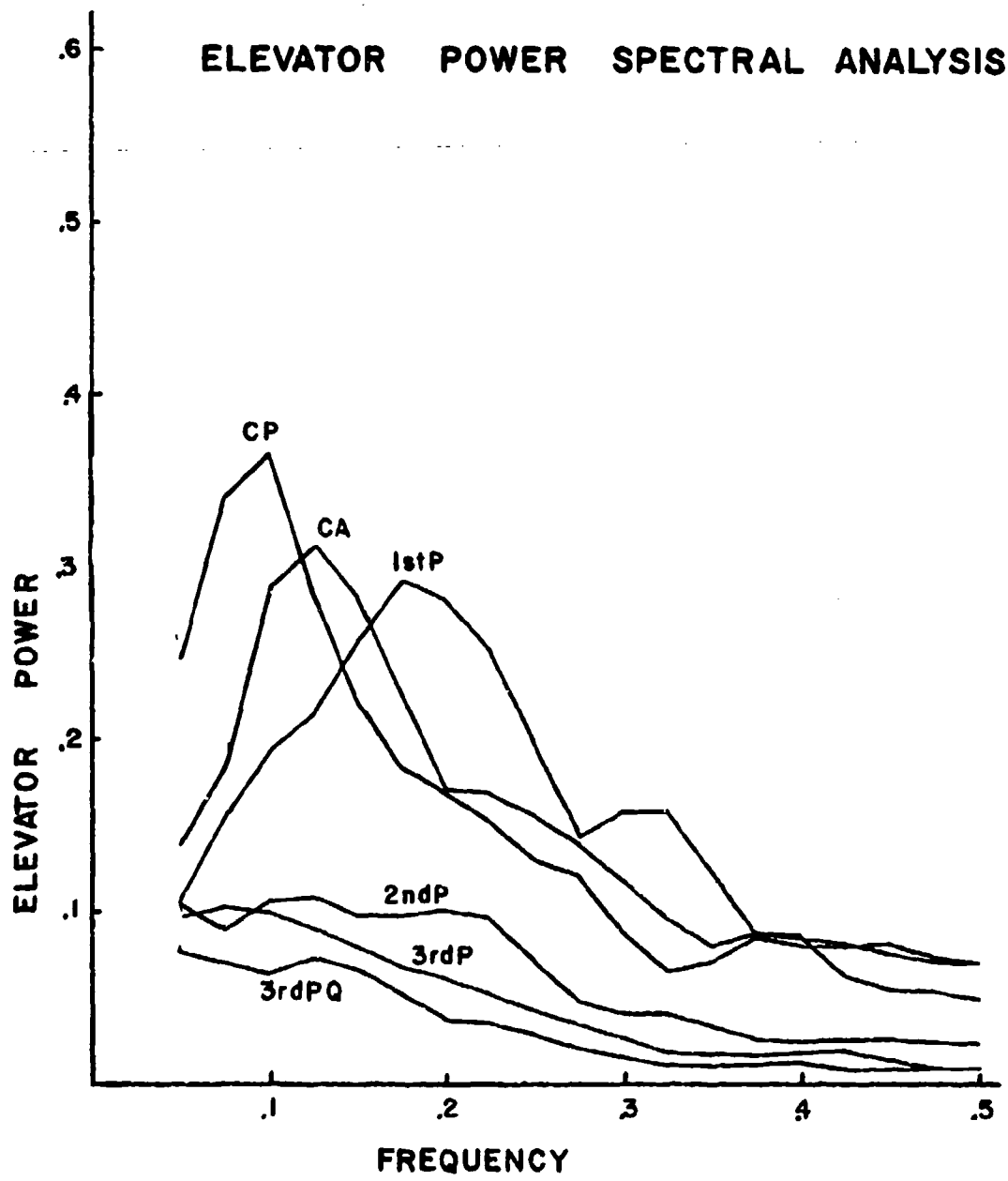


Figure 5. Power spectral analysis of elevator control responses.

AD P001635

Instrument Position-Control Manipulation (IPCOM II): Helicopter Pilot Response as a Function of Control-Display Compatibility

Kathleen M. Craig
San Jose State University Foundation
San Jose, Ca. 95192

Abstract

The proposed research will examine differences in reaction time and task completion measures as a function of control-display configuration in a helicopter cockpit environment. The two configurations to be compared are ipsilateral and contralateral, with the ipsilateral configuration defined as one having each control (cyclic/velocity; collective/altitude) and its corresponding instrument placed on the same side. The contralateral, conventional helicopter configuration, is the converse. Simultaneous manipulation of altitude and velocity controls will be required. Fitts' law will be used to provide levels of task difficulty.

Introduction

The relationship between man and the machine he operates is particularly important when the "machine" is a helicopter. In special conditions, such as night or foul weather flying, the pilot is almost completely dependent upon quick and reliable feedback from the aircraft's display and control systems. In U. S. Army NOE maneuvers (nap-of-the-earth), the aircraft is flying as close to the earth's surface as possible, utilizing trees and vegetation as cover from enemy radar, and is often only split seconds away from becoming part of the scenery. These situations demand a man-machine system that provides a manageable level of pilot workload--and critical to this is the rational combination of control position and corresponding instrument placement.

In order to fly a helicopter the pilot must use both hands and both feet. The altitude control, called the collective, is a left-hand control. The longitudinal and lateral velocity control, the cyclic, is manipulated with the right hand. The pilot's feet are used to control the tail rotor by means of rudder pedals. In most American-built helicopters the instrument panel is similar to that of its fixed-wing counterpart. The traditional arrangement places the gauge for altitude (controlled with the

left hand) on the right side of the panel, and the gauge for airspeed (controlled with the right hand) on the left side of the panel. This display arrangement has even been carried over into the U. S. Army's new helmet-mounted display of flight instruments in the Pilot Night Vision System (PNVS). In the PNVS sight, altitude and airspeed information are displayed as vertical scales on the right and left respectively of the pilot's line of sight (see Figure 1). The conventional helicopter control-display configuration is not in line with stimulus-response compatibility theory--i.e., each control (cyclic/velocity; collective/altitude) and its corresponding instrument are on opposite sides. In fact, the "incompatibility" of this configuration may add to the amount of workload and cost the pilot valuable extra time.

Stimulus-response compatibility research has indicated that reaction time is faster in a psychomotor task when the required response is in the same direction as the stimulus. A previous study done at NASA-Ames (Hartzell, Dunbar, Beveridge, & Boothe, elsewhere in these proceedings) compared an ipsilateral control-display configuration with the conventional (contralateral, as described above) helicopter configuration. This study involved sequential, discrete tasks--changing altitude or airspeed--and reaction time was found to be significantly faster with the ipsilateral arrangement. The proposed research will bring the experimental situation more in line with the real-world helicopter environment (simultaneous manipulation of cyclic and collective and addition of rudder control) to test whether an ipsilateral control-display configuration is superior in a multi-element control task.

The mathematical paradigm employed to provide various levels of task difficulty will be Fitts' law (Jagacinski, Hartzell, Ward, & Bishop, 1978). In the analysis of discrete movements, Fitts' law refers to the log linear relationship between the difficulty of a task and the time required to complete it. The first part of this law states that the time required to move a pointer from starting position over a given distance to the center of a target increases in a log linear fashion as the index of difficulty increases (see Figure 2). The index of difficulty is defined by the formula [$ID = \log_2 (2A/W)$ where A = distance from the home position to the center of the target (amplitude); W = width of target]. The amplitudes and widths can be manipulated so that there are several combinations for each index of difficulty. This is a demonstration of the reciprocity between speed and accuracy--more accurate movements take longer to accomplish.

According to the second part of Fitts' law, reaction time will remain constant (slope of Ø) throughout difficulty levels; Fitts states that the initial response to

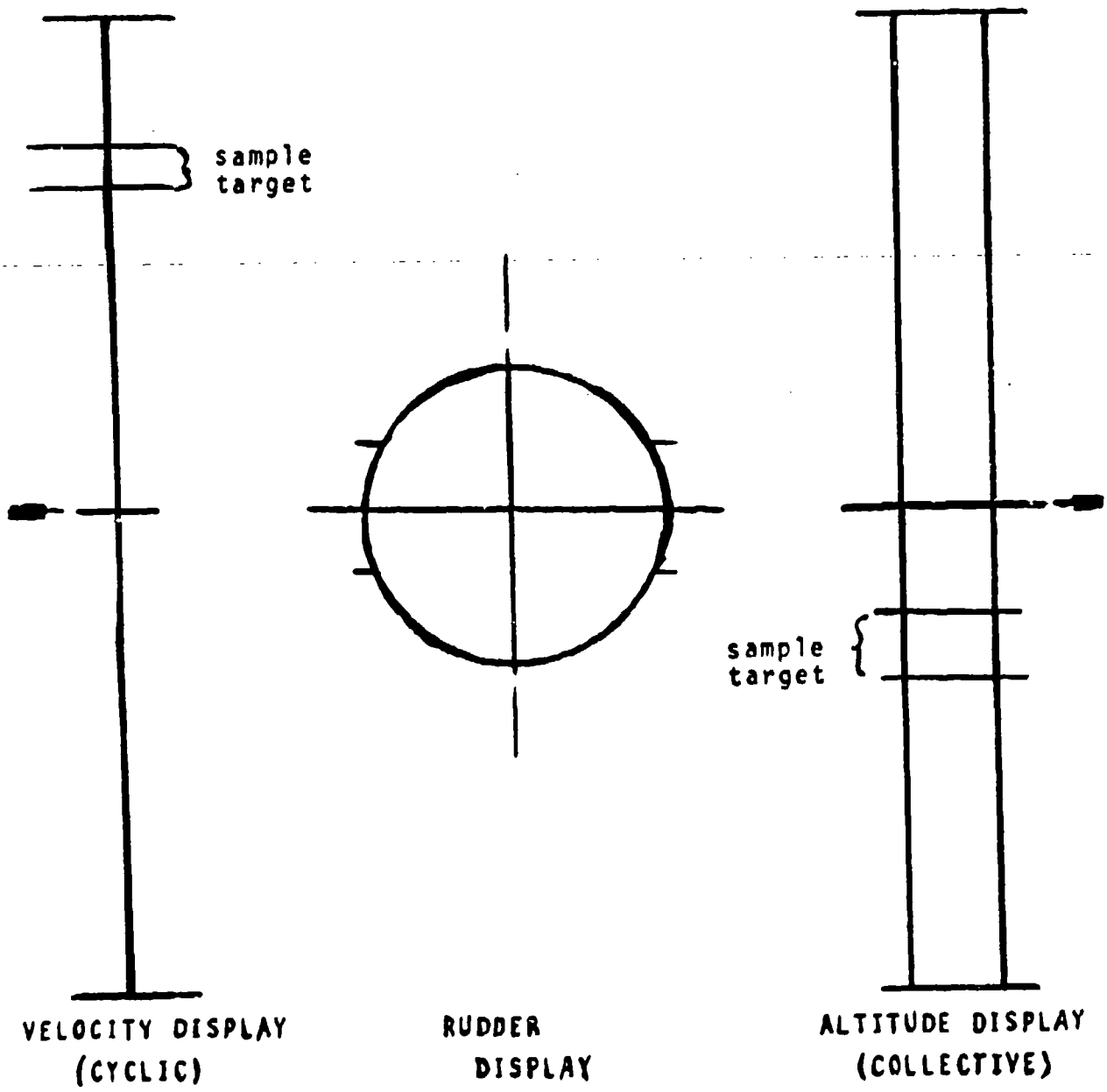


Figure 1. Modified Pilot Night Vision System (PNVS) display, contralateral configuration.

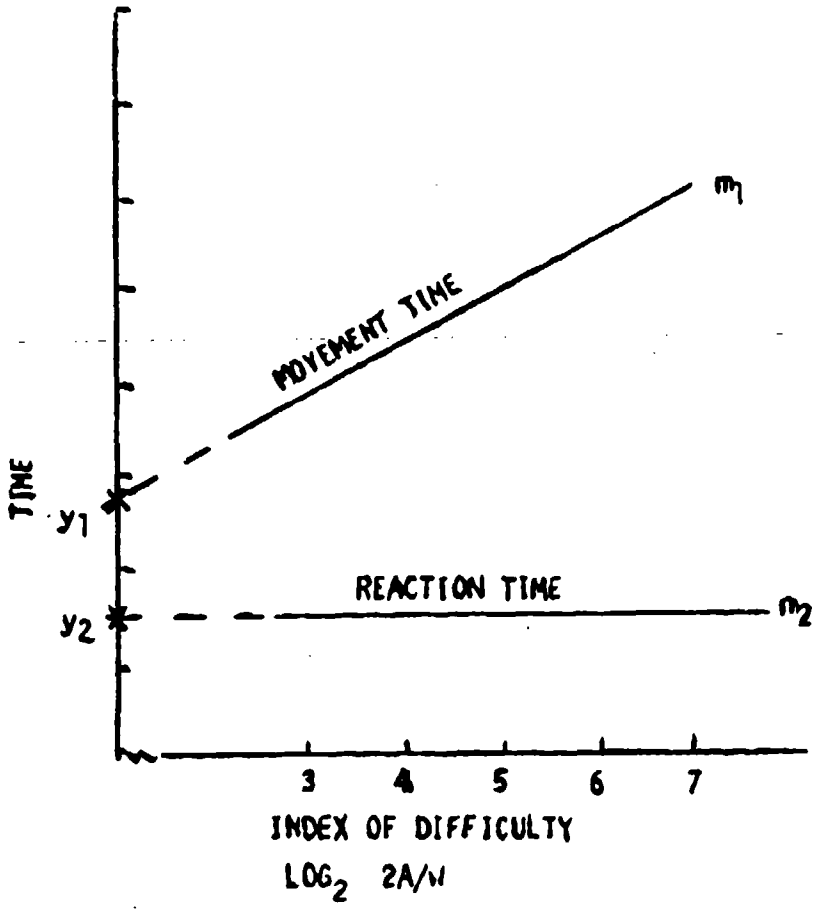


Figure 2. Illustration of Fitts' law relationship between index of task difficulty and time. The slopes and y-intercepts of the regression lines are represented by m_1 and y_1 (movement time) and m_2 and y_2 (reaction time). Reaction time is shown to be independent of difficulty level.

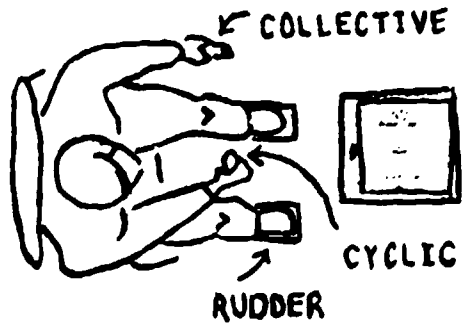


Figure 3. Subject station

the presentation of the stimulus will not be affected by the index of difficulty. Therefore, variations in reaction time can be analyzed in terms of differences in control-display configuration.

Differences in reaction times, capture times, and task completion time as a function of control-display configuration will be examined. Two hypotheses will be tested:

1) Fitts' law is a valid paradigm in a multi-element control task; and

2) in a multi-task situation, the ipsilateral control-display configuration will result in significantly better performance than the contralateral control-display configuration.

Demonstration of the superiority of the ipsilateral control-display configuration in the experimental situation will suggest the desirability of possible cockpit modifications to improve pilot performance in actual helicopter maneuvers.

Method

Subjects. Sixteen non-helicopter pilot subjects with normal or corrected to normal vision will be used--eight for each control-display configuration. Subjects will be tested initially on Jex's (1966) "Critical Tracking Task" which is designed to measure psychomotor ability. Subjects will be classified by performance on this task according to level of tracking ability (moderate or high), and experimental groups will be matched. Subjects will be randomly assigned to two experimental conditions, resulting in two groups of eight subjects each.

Apparatus. The apparatus will be similar to that used in the previous experiment (Hartzell et al., elsewhere in these proceedings). The subject station, as seen in Figure 3, is modeled after the control station of a helicopter. A seat, and cyclic and collective controls are mounted on a platform; in front of the seat is a 53.34-cm. display oscilloscope, which is sloped back from the vertical by four degrees. The rudder control will be added for this study, and will consist of foot pedals mounted on the platform. The control dynamics for the cyclic, collective, and rudder will be independent of each other, and similar to those of an AH-64 helicopter. The rudder will operate a simple k/s plant. Presently (subject to modification), the dynamics of the collective and cyclic are as follows:

COLLECTIVE $\frac{h}{\delta_c} = \frac{6.66}{s(s+0.3)} \quad \left(\frac{ft./sec.}{in}\right)$

CYCLIC $\frac{u}{\delta_e} \cong \frac{-22.89\ (s+1.24)}{[s^2+2(.75)(1.62)s+(1.62)^2](s+.15)^2} \frac{-9}{s} \cdot \Phi/\delta_e(s)$

The position of each channel (control) will be recorded every 10 msec.

The display (contralateral configuration) that will appear on the oscilloscope is as shown in Figure 1. For the ipsilateral condition, the cyclic and collective displays will be reversed. This display is comparable to the electronically-generated symbology that is displayed with the PNVS. The contralateral arrangement is what is conventionally displayed to the pilot. Figure 1 also shows pointers and sample targets for the cyclic and collective; it should be noted that the target sizes and locations may be different for each control. The rudder display will be dynamic, and present throughout the block of trials. The compensatory tracking task will require the subject to null out the random-appearing error to keep the vertical cursor centered, and is intended to be a secondary loading task.

Procedure. Each of the two experimental groups will be trained and tested on only one control configuration--ipsilateral or contralateral. The secondary task (rudder control) included in both configurations will be the same for the two groups. All subjects will be trained to an asymptotic criterion before data measurements are taken. They will be considered to have reached asymptote when mean performance does not vary more than 5%, and standard deviations do not vary more than 10% for a period of eight sessions.

For each primary task control (cyclic and collective), there will be 24 possible targets. These will be factorial combinations of three amplitudes, four target widths, and two directions (up or down from the center). The same combinations of amplitudes and widths will be used for both the cyclic and collective targets, and will yield six different indices of difficulty (as calculated by the formula $ID = \log_2 2A/W$). The sequences of target presentation for each of the controls will be independent; since these sequences will be randomly generated, the resulting mixtures of cyclic and collective targets will also be random.

At the beginning of each experimental session, prior to the first trial of the session, the rudder task will be activated. Subjects will be instructed to use the foot pedals to keep the vertical cursor (see Figure 1) as centered as possible for the duration of the session. A red light appearing 2 seconds prior to the trial and located

at the bottom/center of the oscilloscope will be a warning of the beginning of each trial. The primary display (see Figure 1) will then be activated. the altitude and velocity displays will appear simultaneously at the beginning of each trial, and each will disappear following attainment of criterion as described below. If criterion is not reached after 30 seconds, both pointers and targets will disappear simultaneously. The vertical scales will remain 2 seconds longer, signalling to the subject that the task was not successfully completed.

The subjects' primary task will be to manipulate the cyclic and collective so as to bring the pointers to the areas between the respective target lines as quickly as possible and keep them steady for 350 msec. The trial will be over when this criterion has been attained for each primary target and both pointers are within the target area, or 30 seconds after appearance of the display if the criterion is not met. There will be a 7-second interval between trials.

Subjects will have two blocks of trials per day. All of the trials will be randomly generated. The first block will consist of 10 practice trials, and 36 data trials (6 trials for each index of difficulty). The combinations of targets and indices of difficulty for the cyclic and collective will be randomly generated. After a 10-minute rest period, the subject will receive another session of 36 data trials. Subjects will be given performance feedback after each session to encourage rapid improvement. When asymptotic performance for each subject is maintained for a period of 4 days, the experiment will terminate. Data for these 4 days (eight sessions) will be collected and analyzed as outlined below.

Design and Analysis. The experimental design is a 2 (control configuration) x 2 (control) x 8 (sessions) factorial with repeated measures on the last two factors. Data from the last eight sessions will provide the raw dependent measures. Three different types of measures will be taken: reaction time (RT) for cyclic and collective (from appearance of the target to 2% deflection of each control stick); capture time (CT) for cyclic and for collective (from end of RT to achievement of capture criterion); and total time (TT-- from appearance of the display to end of trial). Median reaction times and capture times for each control, and total times for each trial will be calculated for each subject's performance on each of the six indices of difficulty for each of the last eight sessions.

It is hypothesized that a regression line fitted to each subject's data over sessions on capture time and total time will demonstrate the log linear relationship posited by Fitts (between the difficulty of a task and the

time required to complete it). This demonstration will be based on high correlation coefficients for the regression analyses. It is also hypothesized, in accordance with the second part of Fitts' law (the initial response to the presentation of the stimulus will not be affected by the index of difficulty), that the slopes of the regression lines generated by RT measures on each of the indices of difficulty over sessions will not be significantly different from zero.

For each display configuration, each subject's median reaction times for each control for each difficulty level will be averaged, and a t test comparing the resulting regression line slopes with a slope of zero will be performed to insure that there is not a significant difference. Following this, the slopes and intercepts of the regression lines generated by data on RT, CT, and TT as described above will be used as dependent measures in six separate 2 x 2 x 8 analyses of variance, one for the slope and one for the y-intercept of each time measure.

The number of control reversals (initial movement in the wrong direction) will also be recorded. An examination of the data should show less reversals for the ipsilateral condition than for the contralateral condition. Analysis of the data generated by the rudder control will be an RMS error analysis. It is hypothesized the error should be greater with the contralateral control- display configuration.

The above analyses should confirm the research hypotheses in the following way:

1) The regression lines that will be fit to each subject's data according to index of difficulty on each of the three time measures will demonstrate the applicability of the Fitts' law paradigm to a multi-task situation.

2) The y-intercepts of the reaction time regression lines will be significantly lower for the ipsilateral control-display configuration than for the contralateral configuration. Regression line slopes for capture time and total time measures will be significantly lower for the ipsilateral configuration.

Confirmation of these expectations would indicate that the existing helicopter instrument arrangement may be neither as compatible nor as efficient as is desirable in time-critical flight situations.

References

Hartzell, E. J., Dunbar, S., Beveridge, R., & Boothe, R. Helicopter pilot response latency as a function of the spatial arrangement of instruments and controls. NASA-Ames Research Center, in these proceedings.

Jagacinski, R. J., Hartzell, E. J., Ward, S., & Bishop, K. Fitts' law as a function of system dynamics and target uncertainty. Journal of Motor Behavior, 1978, 10, 123-131.

Jex, H. R., McDonnell, J. D., & Phatak, A. V. A "critical" tracking task for manual control research. IEEE Transactions on Human Factors in Electronics, 1966, 7, 138-145.

AD P001636

HELICOPTER PILOT RESPONSE LATENCY AS A FUNCTION OF THE SPATIAL ARRANGEMENT OF INSTRUMENTS AND CONTROLS

E. James Hartzell
US Army Aeromechanics Lab.
NASA Ames Research Center
Moffett Field, CA 94035

Sherry Dunbar
San Jose State U.
Psychology Department
San Jose, CA

R. Beveridge
Informatics General Inc.
Palo Alto, CA 94303

R. Cortilla
US Army Aeromechanics Lab.
NASA Ames Research Center
Moffett Field, CA 94035

ABSTRACT

The reported study addresses the question of the spatial arrangement of helicopter instruments and controls in terms of stimulus-response compatibility. The results indicate that in airspeed and altitude adjustment tasks the compatible placement of controls and displayed information may result in a significant time savings and reduced workload and therefore increased mission performance. Fitts' Law is used as a dependent measure to assess the performance of subjects in a discrete manual control task.

INTRODUCTION

Military helicopter pilots are now required to fly day/night, all weather missions at low levels and high speeds in order to avoid detection by radar. This scenario introduces conditions requiring rapid and correct responses by the pilot in order to avoid obstacles and complete the mission without mishap. Low-level, high-speed flight is the most demanding in terms of performance and workload of all the missions in aviation. Every effort must be made to improve the interface between the pilot and aircraft in order to provide the widest margin of safety possible. This study was undertaken to investigate one aspect of the pilot interface problem. The problem is one of stimulus-response (S-R) compatibility between instrument placement and control position in current and near-future helicopters.

The altitude indicator in most helicopter instrument panels is located to the right of the pilot's line of sight (LOS) while the airspeed indicator is located to the left. Both instruments are

typically located near the top of the instrument console indicating their importance in the flight control task. The initial decision to place these instruments in this arrangement was based on fixed-wing tradition rather than the unique requirements of modern rotocraft operations and flight scenarios. Specifically, changes in altitude (depicted to the right of the pilot's LOS) are effected by manipulation of the collective in the left hand. The airspeed (displayed to the left of the pilot's LOS) is controlled by the cyclic stick in the right hand.

The above suggests a condition of S-R incompatibility between the instruments and their respective controls. That is, there is a lack of simple spatial correspondence between salient perceptual dimensions and the organization of responses for the performance of a skilled task. Conflicts that might result from this incompatible arrangement are not only intuitively apparent, but have been associated with slower response times and higher error rates in choice reaction time studies (ref. 1) when there was no simple geometric correspondence between the stimulus (light) and the required response (stylus movement). Thus, it is possible that in the control of a helicopter in precision maneuvers, valuable time is lost and confusion-induced errors made due to a lack of S-R compatibility. This may be, in turn, a contributing factor to the high level of pilot workload associated with low alt ıde, high speed helicopter operations.

This lack of S-R compatibility has, through tradition, been carried forward into the cockpit of the US Armys' advanced design helicopters. These rotocraft are specifically designed to fly all weather, day/night, nap-of-the-earth (NOE) missions. To accomplish these missions, the pilot is equipped with a helmet-mounted monocular CRT display of the outside-the-cockpit scene. The scene is produced by a Forward Looking Infrared (FLIR) camera which is gimbal-mounted in the nose of the helicopter. The movement of the camera is slaved to the head movements of the pilot. Sensors located on the helmet and referenced to cabin mounted emitters detect pilot head movements and convert these movements into commands to direct the FLIR camera in azimuth and elevation. Superimposed on the display of the outside scene of the passing world is a collection of flight-related symbology elements which are intended to take the place of the standard panel instruments (see Fig. 1). The total system of FLIR and helmet-mounted display is known as the pilot night vision system (PNVS). Note that the airspeed is shown on the left side of the helmet-mounted display and the altitude indicator is located on the right side of the display. Thus, even in the most modern of helicopter designs incorporating the highest levels of technology, the tradition of S-R incompatibility prevails. With the rapid advances in technology and the increase in demands on the pilot this may be the time to reconsider the positioning of displayed information in the cockpit of modern helicopters.

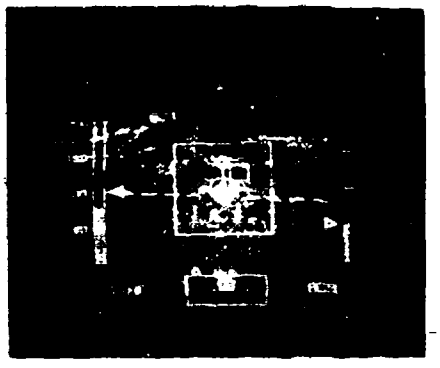


Figure 1. Flight related symbology superimposed on the display of the outside scene as seen by the pilot in the Pilot Night Vision System (PNVS).

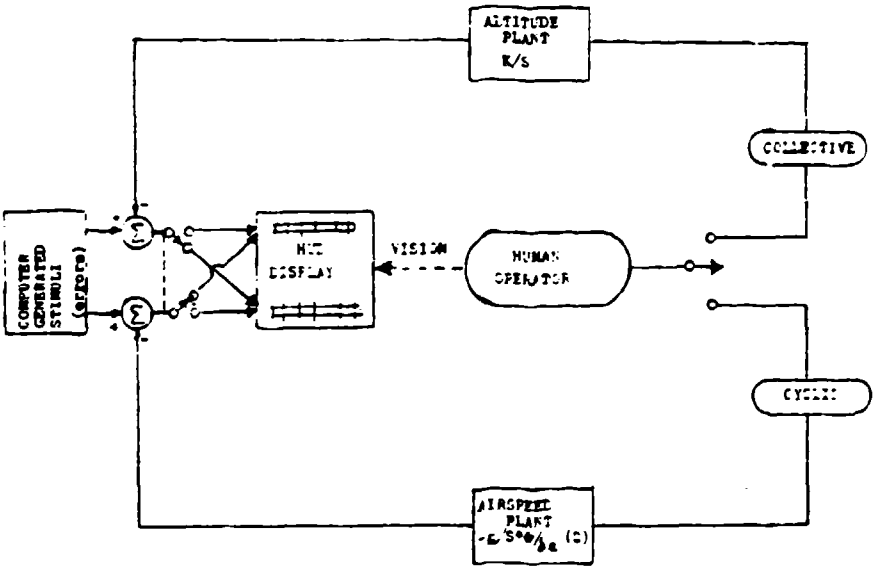


Figure 2. The task requires that the pilot act as a regulator to close two control loops, one each for altitude and airspeed.

It can be argued that training, experience, and learned scan patterns may outweigh any adverse effects on performance or mental workload due to S-R incompatibility. However, the impact that the position of these two primary instruments has on the pilot's control task should be determined by objective research rather than by tradition. To this end, a PNVS-type display was simulated and subjects were required to effect changes in airspeed and altitude in one of two conditions. The first was the control/display compatible, or ipsilateral arrangement where the control and corresponding display were on the same side. The other condition was the control/display incompatible, or contralateral arrangement where the control and corresponding display were on opposite sides of the subject's LOS.

The research question of interest may be stated as : "Is there any difference in the time required to effect control when the controlled element is contralateral (current arrangement) or ipsilateral (stimulus response compatible arrangement) to the control manipulator?".

Experimental Considerations

The research question introduced above involves several considerations which must be taken into account when formulating the experimental objectives and procedures in order to project the results to the real world. These considerations are: (1) stimulus input; (2) response output; and (3) measurement of performance. The pilot may be viewed as a regulator whose task is to close two control loops as depicted in Fig. 2. The perception of the visual stimulus serves as the input and the motor control reflects the output. In the current study, the control position, instrument spatial and visual angle relationship, is modeled after the PNVS.

Stimulus Input:

As mentioned above, in night NOE flight, electronically generated symbology is presented on the helmet-mounted display to provide the pilot with flight instrument information. The PNVS airspeed and altitude information indicators are located approximately +/- 15 degrees about the vertical center reference line. The same visual angles were maintained on a fixed CRT display mounted above an instrument panel to simulate the PNVS arrangement. A schematic representation of the two displays used can be seen in Fig. 3.

Response Output:

The pilot is viewed as a regulator that responds to a detected error and is required to bring a system back into balance. Simplified airspeed and altitude responses to the cyclic (right hand) and collective (left hand) controls respectively were simulated. The airspeed and height control plant dynamics are quite different and the regulator, the pilot, uses different muscle bundles to

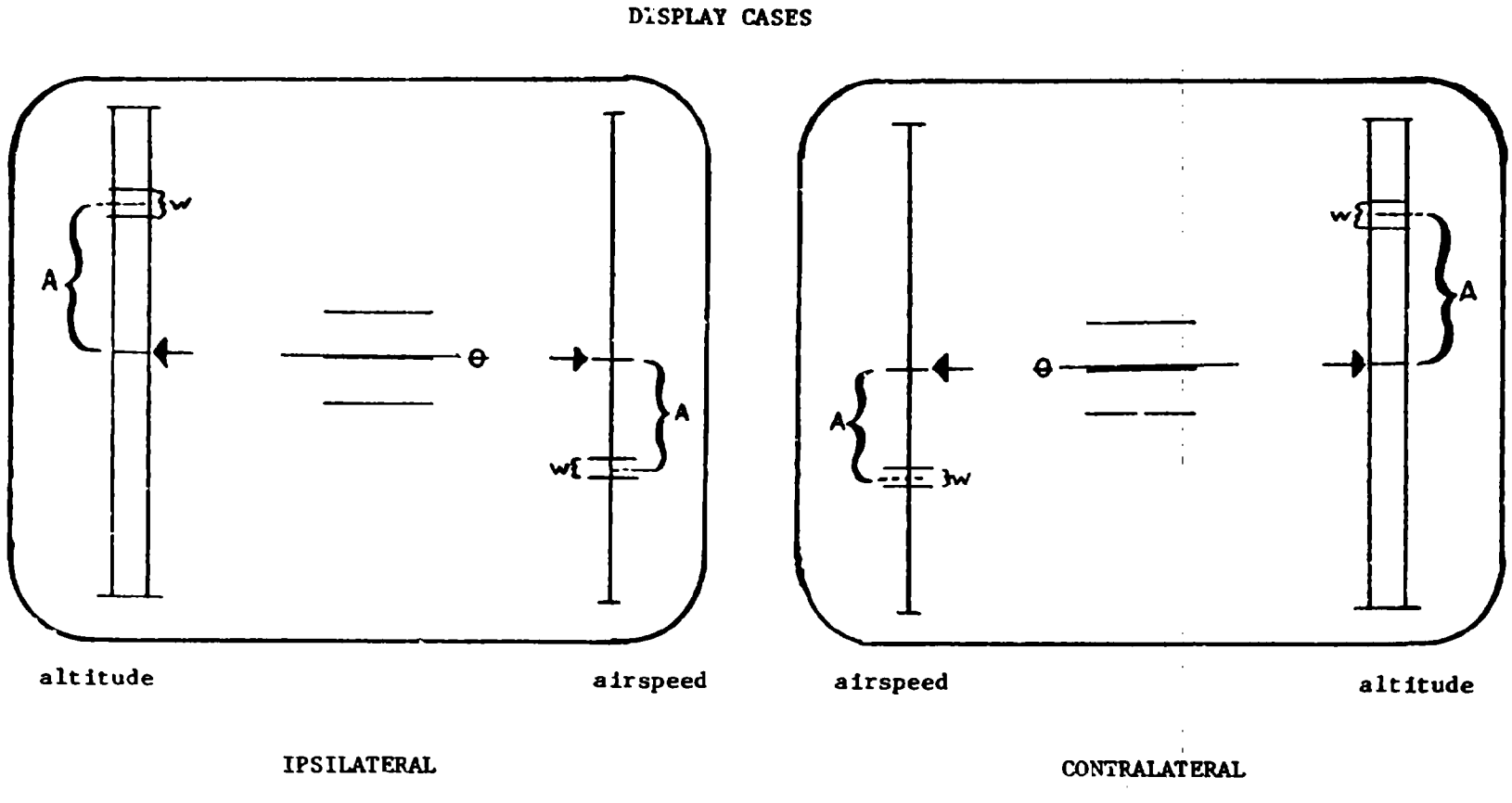


Figure 3. Schematic representation of the two display cases showing amplitude (A) and width (W) of a sample target, as well as pitch attitude (θ) used in the early part of the experiment.

effect control. The two hands are placed in different orientations to the body and the collective (height control) requires an up/down motion, whereas the cyclic requires a "stirring" motion in a more or less horizontal plane.

Performance Measurement:

Simple reaction time measurements could be used as the primary performance metric; however reaction is only the first segment of the overall control task. Since the pilot is acting as a regulator in a closed loop control system, the task is to reduce perceived errors to an acceptable level. Therefore, the control movement time to achieve a criterion is a major variable of interest. The total time required to achieve control includes the pilot reaction time and the manipulator movement time, thus all of these variables were measured in this study. The task will be defined in more detail in the following section. In brief, the subjects were asked to bring a pointer to a specific position on the simulated PNVS thermometer scales using one of the described helicopter controls, while measurements of the above times were recorded. The task was analagous to a pilot reaching a new commanded altitude or airspeed by manipulation of the flight controls.

ANALYTIC APPROACH

A suitable model to employ in structuring the analysis and design of this experiment was proposed by Fitts and Peterson (ref. 2); the resulting mathematical relationship became known as Fitts' Law. They found that there exists a log linear relationship between speed and accuracy of response movements to tasks of varying difficulty. The relationship is expressed by:

$$ID = a + b \log_2 2A/W \qquad (1)$$

Where: A = the amplitude of the initial error
W = the width of the target area
ID = the index of difficulty expressed in bits
a and b are constants

The findings of Fitts and others indicate that the time to effect a reduction in an initial error of amplitude (A) to a given criterion target width (W) varies in a linear fashion with the index of difficulty (ID). Regression lines are fit to the data produced from experiments in which the time to respond and the time to reach the target area are measured as a function of the difficulty of the task. These two regression lines may be seen in Fig. 4. Fitts' Law predicts that the reaction time regression line will have a slope (m_r) which is very near zero. The intercept (y_r) of this function will be an estimate of the response time for all difficulty levels. The movement time regression line (m_m) will have a slope not equal to zero. This indicates that as the task becomes more difficult (established by the ID), more time is required to move to the target. Thus, the task with the greatest

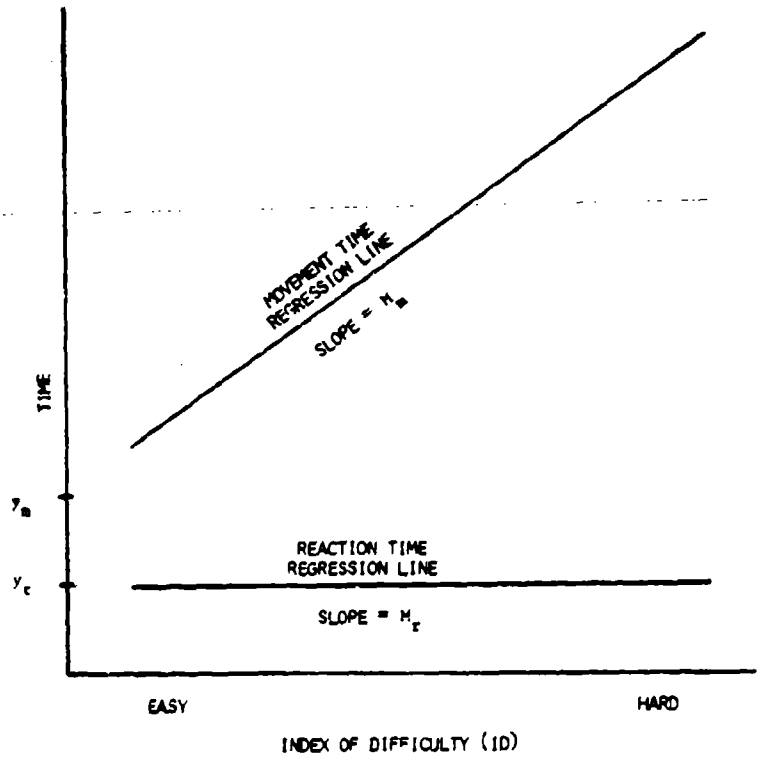


Figure 4. Regression lines fit to the data where the time to respond and the time to reach the target area are measured as a function of the difficulty of the task. Task difficulty is shown on the abscissa as a continuim from easy to hard. Note that the slope of the reaction time regression line (m_r) is plotted as zero, and the slope of the movement time regression line (m_m) is greater than zero. The y-intercepts (y_r and y_m) are indicated on the ordinate.

amplitude (A) and the narrowest target width (W) will result in the longest movement time.

Jagacinski, et al. (ref. 3) replicated Fitts' work using zero- and first-order control dynamics as well as a variation in the capture criteria. Their task involved a joystick manipulator and a CRT-displayed task very much like the current experimental display. Although different slopes were obtained for the zero- and first-order systems, the linear relationship predicted by Fitts held. Others, (refs. 4,5) have also extended Fitts' Law to include up to first-order dynamics. All have found the relationship described by Fitts for a pure manual position reponse task to be robust for higher order dynamics implemented with CRT or visually coupled tasks.

The Fitts' Law model structure, used in a control theoretic sense, lends itself well to answering the current research question. If the spatial relationship between the controller and the controlled element makes no difference, then the slope of the movement time to index of difficulty function will not differ significantly, no matter where the indicator is placed relative to the control device. Fitts' Law also predicts that the index of difficulty will have no impact on the slope of the reaction time regression line (m_r), thus, a slope of zero was predicted. Differences in the y-intercept of the reaction time function will be a predictor of any differences in response time between the contralateral or ipsilateral control/display arrangement. Differences between the slopes of the altitude and airspeed or collective and cyclic movement time regression lines were predicted because the plant dynamics are first- and second-order respectively. Differences in the slopes of the movement time regression lines as a function of the ipsilateral or contralateral arrangement will indicate that the effects have propagated beyond the initial response and into the control phase of the task.

METHOD

SUBJECTS:

Sixteen non-helicopter-rated subjects with normal or corrected to normal vision were employed as members of the subject pool. A screening procedure was used in selecting subjects to insure that all possessed at least a minimum amount of psycho-motor ability. Subjects performed the Jex Critical Tracking Task (ref. 6) and were selected to participate if their score on the Critical Tracking Task was equal to or above one standard deviation below the mean of a distribution of scores produced by helicopter pilots and other persons known to have good motor skills. The subjects were then randomly assigned to two groups: eight to the contralateral and eight to the ipsilateral conditions. In each group, four were high performers and four were moderate as determined by their

scores on the criterion task to balance the tracking skills of the two groups.

APPARATUS:

Cyclic and collective controls and a seat typical of a helicopter were mounted on a platform with a suitable rack on which a 53 cm oscilloscope, sloped back from the vertical by four degrees, was mounted. This arrangement was referred to as the subject station and is seen in Fig. 5. The cyclic control stick had breakout forces of 317.5g forward and 408.2g aft and force gradients of 153.6g/cm and 108.9g/cm respectively. The cyclic stick was spring loaded for center return with a travel of +/- 14 degrees about the center position. The left side, floor-mounted collective control stick was 63.5cm long with an angular motion of +/- 10 degrees about a detent which was provided such that the subjects could return the control to a zero position by feel between trials. The friction of the collective control was adjustable and was left to the subject to establish. However, too much friction masked the feel of the detent and too little did not restrain the collective control from falling of its own weight. As a result only a narrow band of friction adjustments was selected by the subjects.

PLANT DYNAMICS:

The plant dynamics were implemented on a digital computer. The transfer function for the collective controlled altitude rate plant was,

$$\frac{\dot{h}}{\delta_c} \cong \frac{6.66}{(S+.3)} \quad \frac{ft}{sec\ in} \tag{2}$$

and the cyclic-controlled pitch attitude/airspeed plant, a reduced order function representative of a "SCAS on" (stability and control augmentation system) system, was implemented to reflect changes in airspeed as a function of changes in deflection in the cyclic stick. The transfer function is expressed in two parts as follows:

$$\frac{\theta}{\delta_e} \cong \frac{.7133(S+1.24)}{(S+.13)^2\,(S^2+2(.75)(1.62)S+(1.62)^2} \quad \frac{rad}{in} \tag{3}$$

combined with the gravitational constant yields,

$$\frac{u}{\delta_e} = \frac{-g}{S} \cdot \frac{\theta}{\delta_e}(S) \quad \frac{ft}{sec\ in} \tag{4}$$

by expansion, the resultant transfer function is,

$$\frac{u}{\delta_e} \cong \frac{-22.9(S+1.24)}{[S^2+2(.75)(1.62)S+(1.62)^2](S+.15)^2} \quad \frac{ft}{sec\ in} \tag{5}$$

This function represents a simplified small perturbance model of a

Figure 5. The subject station with cyclic, collective, and seat typical of a helicopter showing also, the CRT and a sample task display.

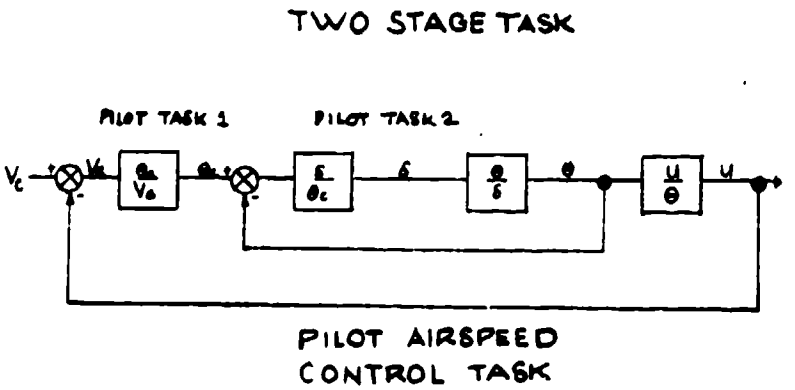


Figure 6. In manipulation of the cyclic, the subject performs a two stage task requiring a change in pitch attitude before a change in airspeed can occur.

high performance helicopter, trimmed at approximately 80 kts and out of ground effect.

The subjects were required to establish a commanded rate of change in height with the collective control, in accordance with equation (2) as represented by the amplitude (A) on the display to an indifference threshold indicated by the target width (W) as is depicted in Fig. 3. In the case of the cyclic stick, the subjects were required to establish a new commanded airspeed by controlling the plant expressed in equation (5), above or below the trim condition. This required a two stage task in that the pitch attitude must first be changed in order for a change in airspeed to occur. This two stage task may be seen in Fig. 6. When a velocity change is commanded (Vc), an error is detected between the current state and the commanded airspeed represented by (Ve). This defines the outer loop control task of the pilot or pilot task 1. However, pilot task 2 must be accomplished first before the desired airspeed (u) is obtained. The inner loop, pilot task 2, consists of adjusting the pitch attitude (θ) by means of the deflection (δ_a) of the cyclic control such that the desired relationship between pitch attitude and resultant airspeed results.

Early in the learning phase of the experiment the subjects were unable to accomplish the task without pitch attitude information presented on the task display. However, as each subject became proficient in the task, the pitch attitude information, seen in the center of each display case in Fig. 3 as an artifical horizon, was removed. The withdrawal of the display resulted in a small setback in the learning curve which was overcome by all subjects in a short time.

PROCEDURES:

Subjects were seated in the subject station in a darkened room so that their eyes were approximately 60 cm from the oscilloscope. They communicated with the experimenter through a microphone and headset. At the beginning of each trial, a red light appeared in a photographic dark field on the center LOS of the subject, and declined by 17 degrees from the horizontal.

The subjects were instructed to fixate this light for two seconds. At the end of two seconds, the display to the left or right was activated with one of the instrument display cases seen in Fig. 3 and dependent on a random and balanced schedule . The displayed cases were centered about the LOS by 15 degrees resulting in a total visual angle of 30 degrees in the horizontal.

The subjects' task was to manipulate the appropriate control so as to bring the pointer between the two target lines as quickly as possible. When the pointer had remained "steady" within the target area for 350 msec, the trial was over and the display disappeared from the screen. The steadiness criterion was defined as the smoothed estimate of the pointer velocity remaining less than or

equal to 0.5 degrees per second over a sampling interval of 350 msec. Following the end of each successful trial, the screen remained blank for two seconds and the next trial began with the illumination of the fixation light. If the subject failed to achieve the 350 msec capture criterion within 30 sec of the appearance of the target and pointer, the trial automatically ended. The vertical lines in the display remained for two seconds, but the target and pointer disappeared as a signal to the subject of failure to meet the criterion.

DESIGN

A repeated measures group design was used for the structure of the experiment and subsequent analysis (ref. 7). Each repeated experimental session consisted of 72 trials, six data trials for the six indices of difficulty for each of the two controls (cyclic and collective). The difficulty levels were generated from a combination of three amplitudes (A) and four target widths (W) resulting in difficulty indices of 2.3, 3.1, 3.9, 4.7, 5.5, and 6.3 bits, and two directions (target above or below the center). The order of presentation followed a two-part schedule. During the first experimental session of each day a subject received ten practice trials of randomly mixed indices of difficulty cases. Practice trials were followed by 72 data trials consisting of a random ordering of cases with the constraint that each of the six cases appeared twelve times (six cyclic and six collective). The second session proceeded with the same format (omitting the practice trials) after a ten minute rest period.

The design of the experiment called for a period of four days (8 sessions of 72 trials each) during which asymptotic performance was maintained by each subject. Asymptotic performance was defined as a change in variability (as measured by the standard deviation) of no more than 10% over a four day period and no more than a 5% change in the mean of reaction and capture times over the same period. This requirement was the factor which determined the number of days each subject participated and thus the total number of days of the experiment. The average number of days to asymptotic performance per subject was 18 days or 40 hrs of training.

Data were sampled at a rate of 100 samples per second per channel (10 msec sampling rate) for the collective and the cyclic. The occurrence of control reversals (selecting the right control but with movement in the wrong direction) and blunders (selecting the wrong control) were also recorded.

RESULTS

Mean and median reaction and movement times were calculated for each subject's performance on each control for the six indices of difficulty over the last eight sessions. The median was chosen as the proper measure of central tendency due to a mild skewness of the time measures. Further, the initial ANOVA indicated that there was no days effect, consequently, the data were collapsed over

eight replications. The nonsignificant results of this ANOVA support the choice of the criteria for the definition of asymptotic performance in this experiment.

A separate linear regression line was fit to each subject's reaction and movement times for each of the two controls, as a function of the index of difficulty. The results of this regression analysis demonstrated that the correlation coefficients for the reaction time functions were relatively low. This was to be expected as the slopes of the reaction time regression lines approached zero. However, the correlation coefficients for the individual subjects regression functions for movement times ranged from approximatly 0.94 to 0.99 for the collective control and 0.96 to 0.99 for the cyclic stick.

Given the relatively high correlation coefficients obtained for the individual subjects, the slopes and intercepts of the regression lines were used as dependent measures in four separate ANOVA's, one each for the slopes and for the intercepts of the reaction and movement time measures. The results of the ANOVA procedures are summarized in Table 1.

The analysis of the reaction time (RT) measures indicated no effect for the slope (m) of the regression lines between cyclic and collective controls, but did show a significant effect for the intercept (y) between the contralateral and ipsilateral groups. However, there was no difference between the reaction time intercepts of the cyclic and collective controls. This indicates that no matter what the control or controlled plant dynamics were, it took significantly longer to respond to the contralateral than the ipsilateral display case.

The ANOVA results for movement time (MT) slopes (m) indicated that there was the expected significant difference, between the cyclic and collective controls. This result was predicted because of the difference in the order of the plant dynamics of the two controls. A significant main effect was also revealed for the slopes of the MT for condition (ipsilateral/contralateral) for the cyclic control. This effect was not found with the collective control. The analysis showed no differences between conditions or controls for the intercepts (y) of the movement time regression lines. A summary of the results of the MT analysis may be seen in Fig. 7. Note that though there is no significant difference between the slopes of the collective control as a function of display condition, it took more time to meet the capture criterion with the contralateral arrangement.

The results of the control reversal and blunder analysis can be seen in Fig. 8. These data show that on the average there were more control reversals in the ipsilateral condition for each control than for the corresponding control in the contralateral group. A similar comparison of the occurrence of blunders between conditions produced the opposite result. There were more blunders

			IPSI	CONT		
CYCLIC	MT	m	1.155	1.387	*	**
		y	.988	.748	ns	
	RT	m	.0001	-.0008	ns	
		y	.343	.420	*	
COLLECTIVE	MT	m	.491	.498	ns	
		y	.655	.738	ns	
	RT	m	-.0013	-.0016	ns	
		y	.360	.417	*	

ANOVA RESULTS

Table 1. Results of four separate ANOVA's, one each for the slope and y-intercept of the reaction and movement time measures. (*, $p < .05$ **, $p < .01$)

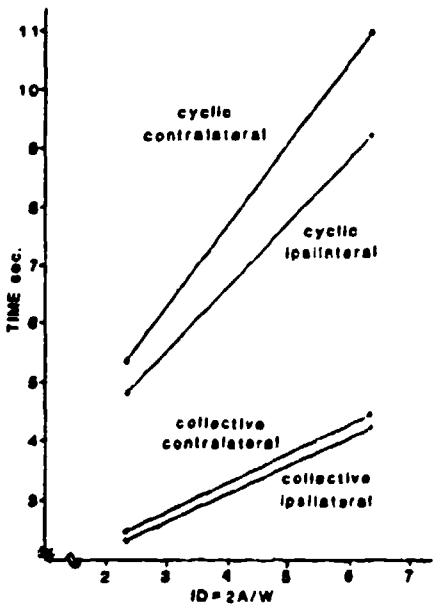


Figure 7. Results of the MT analysis showing the regression lines for the two controls in the two conditions.

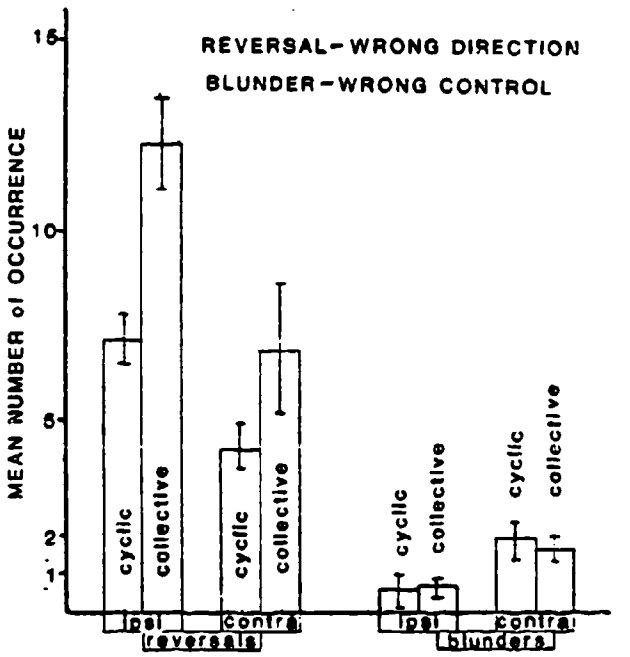


Figure 8. Results of the reversal and blunder analysis showing the more frequent occurrence of reversals in the ipsilateral condition, and the greater number of blunders in the contralateral condition.

in the contralateral condition for both controls than in the ipsilateral arrangement. There were far less blunders than reversals on the average across subjects.

DISCUSSION

REACTION TIME RESULTS:
The analysis of reaction times for the two controls and the two display conditions supported the superiority of the compatible S-R arrangement (ipsilateral) over the contralateral condition. This is born out by the intercept of the regression lines but may be expressed in terms of the mean reaction times. In a sense, pure response time is a more accurate measure than the intercept of a regression line with a slope of zero. Table 2 presents the mean response times in msec for the eight subjects in each of the display conditions and for both controls. The average subject standard deviation is also included and inspection of these data reveals rather low variability, indicating that the subjects had reached a high level of proficiency with the task.

	IPSILATERAL		CONTRALATERAL	
	MEAN	S.D	MEAN	S.D
CYCLIC	348	15	430	19
COLLECTIVE	354	14	419	17

Table 2. Mean response time and average standard deviation (msec) for each display group and control.

The average difference in response time between the ipsilateral and contralateral groups is approximately 74 msec when both controls are considered together. An example of the speed, altitude and flight profile of a helicopter in a typical NOE mission will help to bring these results into focus.

Consider a helicopter flying at an altitude of three meters and an airspeed of 15.7 m/sec with a rotor blade clearance on either side of two meters. Under these conditions, and with the existing PNVS contralateral display arrangement, a pilot would be expected to respond to a commanded change in airspeed or altitude in 425 msec. In this interval of time, the helicopter would travel about 6.65 meters (before the pilot responded with a control action). The ipsilateral or S-R compatible arrangement saves the pilot 74 msec in response time and a distance traveled of 1.2 meters. At great-

er altitudes and in the absence of obstacles, this advantage is of little consequense other than perhaps reducing workload. At night in a NOE flight, advantages associated with the use of S-R compatible instrument-control arrangements may be critical.

A plausible explanation for the 74 msec superiority of the ipsilateral over the contralateral arrangement of controls and instruments in this study is the additional information processing time required to respond to the incompatible S-R condition. The literature in support of the superiority of reaction time performance in S-R compatible systems is extensive and offers many explanations. The purpose of this study was to determine if there was a difference between the compatible and incompatible instrument-control relationship. A practical as well as a statistical difference has been revealed by the results of this study.

MOVEMENT TIME RESULTS:
The second aspect of the overall task, moving the cursor to the target area, provides an opportunity not only to assess the question of S-R compatibilty but also to investigate the Fitts log-accuracy scale in a task involving the complex and higher order dynamics implemented in this study. The Fitts paradigm was used in the design of this experiment in order to provide a systematic set of increasingly difficult tasks which the subjects were required to perform. No attempt was made to use the information theoretic properties of Fitts Law in assessing performance with the two display groups. Instead, the speed/accuracy trade off imposed by the index of difficulty of a given task was used to produce a measureable time course of control movement. A complete analysis of Fitts' Law is beyond the scope of the current paper. However, the rather high correlation coefficients associated with the regression analysis suggest that even for the dynamics as expressed in equation (5), the well trained subjects generated a set of movement times which obey Fitts' Law.

The significant difference between the slopes of the ipsilateral and contralateral groups for the cyclic control was unexpected. Had the slopes been the same but the lines offset in time, then a pure time delay could have been posited much the same as was suggested in the case of response time measures. A significant difference would have been found in the intercepts of the regression lines between the two conditions. The analysis however, suggested that there is both a time and a slope difference. Close inspection of Fig. 7 will lend some meaning to the problem statement. For an ID of 2.3 there is a 600 msec difference between conditions whereas an ID of 6.3 produces a difference of about 1.7 seconds. Note that the collective control produces two parallel lines, offset in time, as a function of the display conditio.. This offset is not a significant result but it is rank ordered as would be expected from the reaction time findings. Also, since the plant dynamics associated with the two controls are of different order, one would expect the cyclic (the higher order and more difficult plant) to have a steeper slope. The fact that the

slopes for the two conditions are significantly different for the cyclic control suggests that there is, in a cognitive sense, a "different plant" being controlled. There is no difference in the dynamics. Only the displays are different. The effect seen might be due to a forced change in control strategy between the two display conditions. An alternate explanation is that there is an increase in the time required to effect control as the difficulty of the task (ID) increases as a result of the S-R incompatibility. That is, the slope differential is about 0.23. This is not an unreasonable explanation in that increasing difficulty may require increasing information processing time.

The results of this study make it clear that the effects of S-R incompatibility are not limited to the initial response phase of a control activity. They seem to propogate into the control phase as well. For the highest ID (6.3) there is a 1.7 sec diference between the ipsilateral and contralateral conditions. Given the conditions of NOE flight described above, in 1.7 seconds the helicopter would have traveled 26.6 meters farther with the contralateral condition before reaching the commanded control state than with the ipsilateral display of flight information. Again, the S-R compatibility issue has both practical and statistical significance.

REVERSAL AND BLUNDER RESULTS:

If one rank orders the difficulty of the tasks within the experiment, the contralateral display of airspeed controlled by the cyclic is the most difficult. This is due to the combination of the higher order cyclic plant and the contralateral display case. The easier task, by the same argument, would be the collective control with the ipsilateral display case. Based on easier dynamics and a less confusing display case, one would predict fewer control reversals and blunders with the collective-ipsilateral tasks. Although this proved to be true for control blunders it was not found in the reversal results. The highest number of reversals was associated with the easiest task; collective with the ipsilateral display case. In fact, about 34 % of the ipsilateral-collective trials were initated with control reversals. Control reversals were detected on 19% of the contralateral-collective trials averaged over all subjects in this group. For the cyclic control, the ipsilateral and contralateral groups produced 19 and 12% reversals respectively averaged over all trials per session and over subjects in each display group. Since it was not possible to measure the difference between response time in the right direction from response time in the wrong direction, the reaction times reported above are contaminated with the reversal events. Reaction times were measured from the presentation of the display until a 2% deflection of the proper control in the correct direction was detected. Thus, if the reversal times were excluded, the response times for the ipsilateral display cases would be even less than was reported. Because there were fewer reverals for the contrala-

teral group, there would be an even greater difference in reponse times between the two display conditions if the reversal times were removed from the response times.

A possible explanation can be structured from the increased information processing time associated with the contralateral display case as posited above for the response time results. If the contralateral condition requires more time for the subjects to "sort-out" the S-R incompatible arrangement and response, then the subject also has more time to decrease the likelihood of responding in the correct direction. With the easier task, the strategy of the subjects may have been to respond and then correct for direction. However, even in the ipsilateral-collective task there was a better-than-chance probability (66%) that the subjects would respond in the correct direction.

CONCLUSION

The results of this experiment suggest the need to reconsider the S-R compatibility of the instrument information displayed in the PNVS in particular and in the cockpits of advanced technology aircraft in general. The results of the reaction time analysis are in keeping with predictions based on the theory of the superiority of compatible S-R arrangements. The propogation of the adverse contralateral effect into the control movement phase of the cyclic task is a compelling reason to attend to the possible problems of information transfer early in the design of new systems.

References

1. Fitts, P.M. & Seeger, C.M. SR compatibility: Spatial characteristics of stimulus codes Journal of Experimental Psychology, 1953, 46, 199-210.

2. Fitts, P.M. & Peterson, Jr. R. Information capacity of discrete motor responses. Journal of Experimental Psychology, 1964, 67, 103-112.

3. Jagacinski, R.J., Hartzell, E.J., Ward, S., & Bishop, K. Fitts' law as a function of system dynamics and target uncertainty. Journal of Motor Behavior, 1978, 10, 123-131.

4. Langolf, G.D., Chaffin, D.B., & Foulke, J.A. An investigation of Fitts' law using a wide range of movement amplitudes. Journal of Motor Behavior, 1976, 8, 113-128.

5. Sheridan, T.B. & Ferrell, W.R. Remote manipulative control with transmission delay. IEEE Transactions on Human Factors in Electronics, 1963, HFE-4, 25-29.

6. Jex, H.R., McDonnell, J.D., & Phatak, A.V. A "critical" tracking task for manual control research. IEEE Transactions on Human Factors in Electronics, 1966, 7, 138-145.

7. Winer, B.J. Statistical Principles in Experimental Design: New York, McGraw Hill, 2nd Ed., 1971.

CONTROLS, DISPLAYS, AND SITUATION AWARENESS ON THE FINAL APPROACH

By

Lee H. Person, Jr.

18th ANNUAL CONFERENCE ON MANUAL CONTROL

June 8 - 10, 1982

Dayton, Ohio

A prime research objective of Langley's Advanced Transport Operating Systems (ATOPS) program has been the development of electronic control and display concepts which enhance a pilot's situation awareness during landing approaches in reduced weather minima. A basic philosophy has been to provide the pilot situation and predictive information, in a simple integrated format, and de-emphasize the command type information associated with current flight directors. Additionally, a philosophy which keeps the pilot in the decision making and action role has been strongly maintained.

This talk traces a segment of the ATOPS research which retains the pilot in the control loop as an active component of an electronic aircraft system where precise navigation solutions, control augmentation, and electronic display concepts have been integrated to increase situation awareness and reduce workload. It focuses on the development of information requirements for a Primary Flight Display (PFD) to complement an augmented manual flight control mode which allows the pilot direct control of the orientation of the aircraft velocity vector.

AD P001637

CHARACTERISTICS OF A COMPENSATORY MANUAL CONTROL SYSTEM WITH ELECTROCUTANEOUS FEEDBACK

by
Kazuo Tanie,*) Susumu Tachi, Kiyoshi Komoriya, Hideo Iguchi,
Masao Takabe and Minoru Abe
Mechanical Engineering Laboratory
Ministry of International Trade and Industry
Tsukuba Science City, Ibaraki, 305 Japan
and
Yoshihiro Sakai
University of Electro-Communication
1-5-1, Chofugaoka, Chofu-Shi, Tokyo, 182 Japan

ABSTRACT

Electrocutaneous stimulation using electric pulse trains is one of the most effective cutaneous displays for tactual communication systems. This paper deals with the human operator's characteristics in compensatory manual tracking system including an electrocutaneous display. Firstly, an electrocutaneous manual compensatory tracking system which consists of a random noise generator, a low-pass filter, a multichannel stimulator, a stimulus pulse energy controller, an isolator and a computer was constructed. Secondly, human operator's characteristics under several parameter situations in a compensatory tracking system (i.e., display transformation method, display gain, and pulse frequency) were evaluated from Root Mean Square values (RMS) of control errors and transfer functions of human operators. Thirdly, differences between human operator's control performance in various cutaneous and visual tracking displays were discussed. As a result, the optimal condition for an electrocutaneous display in a tracking control system was presented and human operator's control performance in electrocutaneous manual tracking system was found to be as good as that in vibro-tactile though inferior to that in visual.

INTRODUCTION

Several previous investigations have been concerned with human performance in compensatory manual tracking tasks [1],[2],[3]. Compensatory tracking has been proved a useful tool for research into the characteristics of the human operator in man-machine systems. Most experiments have used a continuous visual display for indicating the magnitude and direction of the tracking

*) He is a Visiting Scholar at the Biotechnology Laboratory of UCLA, L.A., CA 90024 until August 13, 1982.

error[1]. Tactile displays offer the possibility of analogous one dimensional and spatial displays. Several investigators have proved that tactile displays can be used successfully to construct an effective sensory substitution system for prosthetic limbs, sensory prostheses and so on[4],[5],[6],[7].

Stimulation of tactile sense can be accomplished with electrical, mechanical, thermal, and chemical stimuli. It has been established that the most practical method for communication purposes is the mechanical and electrical. Investigators that have experimented with tactile tracking include Durr[8], Geldard[10], Hirsch and Kadushin[11], Howell and Briggs[12], Weissenberger and Sheridan[13], Seely and Bliss[14], Alles[15], and Hill[16]. Most of this research is concerned with tactile tracking with mechanical or vibro-tactile stimuli. Investigations concerning with manual tracking with electrocutaneous feedback has been carried out by Szeto[17], Schmid and Bekey [18], and Anani and Körner[19]. Szeto proposed seven kinds of electrode tactual codes and evaluated the effectiveness of them from RMS value of control error. Schmid et al. used a two dimensional electrocutaneous display to estimate the behavior of the spinal cord injury handicapped from the transfer function of the human operator in tracking tasks. Anani et al. studied human operator performance in tracking system including an afferent electrical nerve stimulation. Though researchers indicated above have actively investigated manual tracking tasks with electrocutaneous feedback, there has been little research on compensatory tracking tasks.

In general, human operator behavior in manual tracking tasks depends on various tracking system parameters - the type of command signal, command signal bandwidth, display gain, closed loop gain, the type of display transformation and controlled element dynamics[20]. Manual tracking experiments in the past involve varying these parameters. The present research emphasized varying the type of display transformation and stimulus pulse parameters as well as other classical parameters, display gain and command signal bandwidth. From this approach, it should be possible to find the optimal stimulus condition in manual tracking tasks with electrocutaneous feedback. The difference between electrocutaneous stimulation and other display modalities will also be investigated.

INSTRUMENTATION

Electrocutaneous communication can be accomplished with several stimulus waves and information transmission codes. However, pulse stimulation is usually chosen with the information codes, magnitude, pitch and location of the perceived pulse stimulus. In this series of experiments, magnitude and pitch of pulse stimulus were chosen as information codes. The use of those codes are important for constructing a simple electrocutaneous communication system since it allows to build the system under the use of less electrodes, though location code can transmit a larger amount of information.

Fig. 1 shows the experimental setup used. The experiment system includes a computer(PDP 11/40), multichannel simultaneous stimulator, stimulus energy controller, isolator, electrodes, low-pass filter with cutoff frequency adjustable, random noise generator and joystick. The output of stimulator was isolated by the photo-coupling type isolator and was presented on to the skin

of a subject via two sets of wet electrodes. The two sets of electrodes were placed on the skin above the biceps brachii(Fig. 1). Two electrode sites were located 100 mm apart along the longitudinal axis of the biceps brachii. At each site were three electrodes placed 20 mm apart perpendicularly to the muscle's length. The two outer electrodes at each site were connected and used as a common, and negative pulses were applied to each central electrode.

In previous paper the authors reported that the magnitude sensation of electrocutaneous stimulation was strongly related to the stimulus pulse energy[21]. For this reason, stimulus pulse energy and frequency were used as stimulus parameters relevant to magnitude and pitch codes in electrocutaneous information, respectively.

The task of the subject in this experiment is to eliminate tracking error by manually adjusting a joystick in response to command signals. The command signal to the system is white noise through a low-pass filter. Movement of joystick, which consists of a dial and potentiometer, produced an analog voltage output proportional to the displacement from the center position. Error signals, the voltage difference between command and joystick output, were calculated using operational amplifiers. After the detection of the error signals, they were sampled at the rate equal to pulse rate of stimulation used. These sampled error signals were put into computer via analog-to-digital converter and used to modify pulse parameters. Pulse energy and pulse frequency of stimulus were varied according to the sampled tracking error signals. The electrode site near the shoulder was stimulated when tracking error was positive, and electrode site near the elbow was stimulated when tracking error was negative. The selection of stimulus sites according to the direction of tracking error was performed using computer controlled relay device. A computer controlled multichannel simultaneous stimulator was used to produce stimulus pulse trains with appropriate parameters corresponding to sampled tracking error signals. The energy of each pulse was regulated by a stimulus energy controller associated with constant current isolator. The details of multichannel stimulator and stimulus energy controller have been described in other papers[22],[23],[24].

Throughout the experiment, the dynamics of the joystick were kept constant. The stroke of joystick movement was about 90 degrees in each of clockwise and counterclockwise directions from the center position. As a controlled element, position vechicle was used.

Subjects were researchers of experience. One of them performed all of experimental tasks. The results for the subject are shown in the following sessions.

EVALUATION OF THE MAGNITUDE CODE

The magnitude code information will be transmitted to a subject intermittently, when pulse stimulation is used. Human operator performance in this manual control system with magnitude code electrocutaneous feedback is considered to depend on the pulse frequency of stimulation used. The use of pulse stimulation with long pulse interval may cause the decrease of human operator controllability because of the decrease of transmitted feedback information per unit time.

The first experiment was conducted to investigate the relation between human operator controllability and pulse frequency. In the second experiment, human operator behavior in a manual tracking task with several system parameters were measured under the use of the optimum stimulus frequency estimated from the first experiment.

EXPERIMENT I ---- VARYING STIMULUS PULSE FREQUENCY

METHOD

Pulse frequency and bandwidth of the command signal were varied. Six or seven kinds of pulse frequencies in the range from 10 pps to 100 pps, and white noise command band-limited at 0.1, 0.2, 0.4, 0.5 and 0.8 Hz, which show cutoff frequencies of low-pass filter(F_C), were used. A subject was required to perform compensatory tracking tasks to each of those command signals, perceiving electrocutaneous feedback with each of those pulse frequency stimulations. Pulse energy of stimulus was varied proportionally according to the increase of tracking errors in the range from minimum threshold to maximum threshold. The display transformation is shown in Fig. 3(a).

Each trial in this tracking experiment was 3.5 minuite long. The trials were performed after the subject had familiarized himself with experimental apparatus and procedures. Tracking errors were recorded into a data recorder (TEAC 260). The data playbacked were transferred into the memory of computer. Tracking errors were squared and integrated in the computer in order to get the RMS values. A RMS value of tracking error was normalized by a RMS Value of command signal. The normalized RMS values of tracking errors were used as a final human operator performance measure in this experiment.

RESULTS AND DISCUSSIONS

Fig. 2 shows the subject's normalized RMS value as a function of stimulus pulse interval. Each curve in the figure corresponds to the experimental results with each bandwidth of command signal used. As shown in Fig. 2, the trial for F_C=0.8 Hz was performed only under the use of stimulations with pulse intervals less than 50 ms, because it was difficult for a subject to track the command signal when stimulations with pulse intervals larger than 50 ms were used. From the inspection of the experimental results, it is found out that (a) the smaller both of pulse interval of stimulation and bandwidth of command signal are, the better the compensatory tracking scores are, and (b) the decreasing rates of RMS scores increase according to the increase of bandwidth of command signal, especially when the stimulations with pulse intervals larger than 50 ms are used.

One possible interpretation of those results may be related to how long the human operator can persist in each of the magnitude sensations evoked intermittently by pulse stimulation. A second possible interpretation may be related to the decrease of directional information of tracking errors under the use of stimulations with long pulse intervals. The subject suggested that the latter interpretation was valid.

To investigate this problem, brief experiments were performed. The experimental procedure was the same as experiment I except that the

experimental system contained an additional visual display for exact presentation of directional information of tracking error by means of LED (light emitting diode). The results of the experiment indicated that the tracking scores under the use of stimulation with 100 ms pulse interval in this experiment became approximately equal to those with 10 ms pulse intervals in the previous experiment, while those under the use of stimulation with 10 ms pulse interval were not improved. From the results, it can be concluded that in order to keep good operator performance stimulations with pulse intervals less than 20 ms should be used, and an establishment of exact directional information display is important in a tracking experiment with the use of long pulse interval.

In the following experiments associated with magnitude code stimulation, 10 ms pulse interval was used.

EXPERIMENT II ---- VARYING SYSTEM PARAMETERS

METHOD

Human performance in the tracking system with different transformations and display gains were compared in each experiment. In this experiment display transformation was varied to compare the linear and threshold transformation shown in Fig. 3. In the linear transformation, pulse energy of stimulation varied proportionally according to instantaneous tracking error. In the threshold transformation, pulse energy of stimulation was turned to an appropriate energy level from the minimum sensation threshold, when an instantaneous tracking error became more than a specified value. The linear transformation was tested at several different display gains, and the threshold transformation was done at several dead bands. Each of those experiments was performed at three different bandwidths of command signals. Tracking tasks were same as those in the previous experiments. A normalized RMS value of the tracking error and a human operator transfer function were used as human performance measures. Firstly, the normalized RMS values were computed for each trial, and a gain which indicated the best RMS score was chosen for each bandwidth of command signal in each transformation. After those investigation, another experiment was performed at the best gain in order to estimate the human operator performance from the transfer function. In the experiments associated with RMS score and transfer function, a subject had one and three 3.5 minuite tracking trials for each parameter situation, respectively. Three kinds of variable, tracking errors(e), band-limited command signal(r) and operator output signals(u), were recorded into a data recorder for each trial. It should be noted that the operator output was equal to control value because tracking system used in the experiments had an unity controlled element.

Performance measures were evaluated from the recorded data using a digital computer as mentioned in the above session.

RESULTS AND DISCUSSIONS

Fig. 4(a) shows the relation between the normalized RMS of the tracking error and the display gain for the linear transformation. Fig. 4(b) shows the relation between the normalized RMS of the tracking error and the normalized dead band for the threshold transformation. Each curve in both figures

corresponds to the result of each bandwidth of command signal used. The principal features of the results are as follows: (1) For the linear transformation, each set of trials shows the RMS of tracking error decreases when display gain increases. The RMS score has a tendency to keep relatively constant at higher display gain. (2) For the threshold transformation, each set of trials shows the RMS of tracking error has U-shaped characteristics with varying dead band. (3) RMS scores of tracking tasks for the linear transformation are superior to those for the threshold transformation.

The above result(1) suggests that human operator performance in tracking tasks with magnitude code electrocutaneous feedback will not be effected by adaption of the receptor under the skin. This is usually observed in tactual communication systems and causes temporal decrease of display gain during the tracking task. The difference between the results for the linear and threshold transformation indicates that it is important to display the magnitude information of tracking error as well as the directional information in order to keep good operator perrormance. The decrease of RMS score at lower dead band in the threshold transformation has much to do with instability of the system caused by a large equivalent closed loop gain. The decrease at larger dead band is considered to relate to the decrease of display information accuracy.

Transfer functions are estimated from the following equation,

$$Y_H = F_{RU}(jw)/F_{RE}(jw)$$

where Y_H: human operator transfer function, $F_{RU}(jw)$: cross-power spectral density of closed system input and system output, $F_{RE}(jw)$: cross-power spectral density of closed system input and tracking error.

The two cross-power spectral densities were calculated, utilizing a standard Fast Fourier Transformation program(FFT) in the computer. For FFT analysis, analog data in data recorder were digitized at 10 Hz rate. Proper precautions were taken to reduce aliasing due to sampling by appropriately filtering the data at each step of the operation. FFT were performed for 500 sections of 512 points overlapped by one point with a sampling frequency of 10 Hz. Each cross-power spectral density for each parameter situation was obtained from the average across all of the FFT results. Fig. 5 shows human operator transfer functions. These are results for bandwidths 0.3 Hz, 0.5 Hz and 0.8 Hz at display gain 53 erg/volts and normalized dead band 0.5, which indicated optimum operator performance in previous experiments. Both operator gain and phase shift are plotted. The main difference between the curves is their varying gain. The difference is greater under the use of a command signal which includes higher frequency components. The increase of command signal bandwidth causes the reduction of operator gain. In general, it is well-known that human operator gain decreases in a manual tracking control system with visual feedback when he has command signals which include higher frequency components. This is because he must attend to the whole frequency components which are included in the command signal. Hence, he cannot keep the operator gain constant. The experimental results are interprered in the same way.

The difference between the linear and threshold transformation is not so significant under the use of command signal band-limited at 0.3 Hz. The operator gain characteristics for the linear transformation is evidently superior to that for the threshold transformation under the use of command

signals which are band-limited at larger frequency, that is, 0.5 Hz and 0.8 Hz.

EVALUATION OF THE PITCH CODE

METHOD

One type of display transformation was used in this experiment (Fig. 3(c)). The pulse interval of stimulation was proportionally varied according to tracking errors as shown in the figure. The range of pulse interval was from 10 ms to 100 ms, because it is well-known that the operator has excellent frequency discriminability in this range[25]. Minimum level of pulse interval was fixed at 10 ms. The experiments were broken into two parts as well as the previous experiments. In the first experiment, the RMS score of tracking task was determined with varied display gain. In the next experiment, human operator performance was evaluated using human operator transfer functions with the use of the display gain at which an operator showed an optimum RMS score in the tracking task. The experimental procedures and the processing method of experimental results were identical to the previous experiment. Energy of stimulation pulse was set at a comfortable level for a subject.

RESULTS AND DISCUSSIONS

Fig. 6 shows the normalized RMS scores of tracking errors for various display gains. Bandwidth of command signal employed was from DC to 0.3 Hz. Energy of stimulation pulse was 42.4 erg.

The results indicate that (i) the RMS scores vs. display gain shows U-shaped characteristics and (ii) the RMS scores are inferior to those for the magnitude code. The reason for the decrease of RMS score at lower display gains may be explained by the fact that the subjects must recognize the tracking error through stimulation with longer pulse intervals at those gains. In a tracking task with higher display gain, operator will use frequency ranges with a larger just noticeable difference more often to recognize the tracking error. It is considered to be reason why RMS scores at higher display gain decrease.

Fig. 7 shows human operator transfer functions, which includes the results for command signals band-limited at 0.3 Hz, 0.5 Hz and 0.8 Hz. The display gain was 300 ms/volts, which was the optimum gain in above experimental results. From the inspection of the results, it is found that the operator gain with pitch code communication has a tendency to decrease much more than with magnitude code communication, when command signal with higher frequency component is used. The human operator behavior for pitch code communication, however, is approximately same to that for magnitude code communication. The facts imply that a subject has an inadequate adaptability for the change of command signal bandwidth in a tracking task with pitch code stimulation. The reduction of operator gain under the use of command signal with higher frequency components may be related to information display mechanisms of pitch code communication. When using pitch code communication, information display rate will be varied according to the magnitude of tracking error. With display transformation in Fig. 5, information display rates decrease when tracking error go to zero. Such a decrease of information rate may cause a reduction

of operator controllability, because it made it difficult for a subject to track the wide bandwidth command signal with high accuracy.

COMPARISON OF ELECTROCUTANEOUS DISPLAYS

Considering the electrocutaneous displays tested in these experiments, the best is estimated to be the linear transformation using the magnitude code. The next best is the threshold transformation using the magnitude code or the linear transformation using the pitch code. In order to compare these electrocutaneous displays with each other, and with typical visual and vibrotactile displays performance, the equivalent gain and time-delay parameters for the displays were computed, employing the method which was proposed by Hill[16]. The following equation, which is a kind of extended crossover model was used in order to determine the parameters.

$$Y_H = (K/jw)e^{-(jw\tau-(\pi/2))}$$

where K is the equivalent gain, τ is the equivalent time-delay, and $\pi/2$ is phase-lag constant coefficient, which is needed to get the better fit to data at low frequency when a corssover model is applied to a manual tracking system with position vehicle. For vibrotactile displays, the data of Bliss[9], which were obtained from tracking trials using tactile contact and tactile air jet displays, were used. The data of Elkind, McRuer, Bliss and the authors were used for visual displays. All of data were obtained with a position vehicle. The authors' visual and electrocutaneous display data, and Bliss's visual and tactile contact data were the results from one subject. To compare two parameters K, τ, associated with the electrocutaneous display, the results which were best fitted with the crossover model parameters were chosen for each display transformation. Those results associated with F_c = 0.5 Hz (linear), F_c = 0.3 Hz (threshold) and F_c = 0.3 Hz (frequency). Fig. 8 shows the equivalent gain and time delay parameters for position vehicle plotted on the K-τ plane with equal phase margin curves, calculated using the following equation.

$$\pi-[(K\tau+(\pi/2))-(\pi/2)] = a$$

where a is phase margin.

It can be seen in Fig. 8 that performance obtained with each transformation of electrocutaneous displays is consistently ordered from the worst to the best -- threshold, frequency and linear. In the recent practical application of electrocutaneous communication, only the pitch code has been utilized. It should be noticed that the results indicated that the magnitude code was most important in a compensatory manual tracking system. The electrocutaneous displays yield less gain and greater time-delays than the other displays except tactile air jet. Some care, however, must be exercised when the results of Bliss are interpreted, as Hill suggested[16]. The subject is a man of quite excellent tracking ability and his visual gain is much higher than McRuer and Elkind's average data. The gains on both his tactile and visual trials should be scaled down about 30 percent to match that of the average subject. The results in the present research is considered to provide the data of the average subject, because the subject's visual gain and time-delay approximately

are equal to McRuer and Elkind's average data, as shown in Fig. 8. If the above mentioned modification for the data of Bliss is done, the electrocutaneous linear transformation with magnitude code offer performance fairly comparable to that obtained with tactile contact display of Bliss, though it is still poorer than visual display.

CONCLUSIONS

The feasibility of electrocutaneous displays in compensatory manual tracking has been demonstrated. Three kinds of electrocutaneous displays, linear transformation, threshold transformation using magnitude codes, and linear transformation using pitch code, have been investigated under the best stimulus conditions.

The results of transfer function analysis indicated that the electrocutaneous display was most effective with linear transformation using magnitude code and least effective with threshold transformation.

When compared with other displays, performance for linear transformation using magnitude code was approximately equivalent to that for tactile contact display. However, equivalent gain was less and equivalent time delay was longer for electrocutaneous display than for visual display.

ACKNOWLEDGEMENTS

Authors would like to acknowledge the help and encouragement of Professor Ichiro Kato of the Waseda University, Professor Hitoshi Mochizuki of the University of Electro-Communication and Professor Sadao Fujimura of the University of Tokyo.

Although this research was carried out in Japan, this paper was documented when one of the authors(KT) was as a Visiting Scholar at the Biotechnology Laboratory of UCLA sponsored by the Japanese Government. The authors would like to thank Professor John Lyman, Messrs. David Mackinnon and Joseph Guokas of the Biotechnology Laboratory for their kind support during the preparation of this paper.

REFERENCES

1. A. Tustin: The Nature of the Operator's Response in Manual Control and Its Implication for Controller Design, Journal of IEE, Vol. 94, Part IIA, No. 2, 190/202 (1947).
2. D. McRuer, D. Graham, E. Krendel and W. Reisener, Jr.: Human Pilot Dynamics in Compensatory Systems--Theory, Models, and Experiments with Controlled Element and Forcing Function Variables, Air Force Flight Dynamics Laboratory, Air Force Systems Command, Wright-Patterson AFB, Ohio, Tech. Rept. 115-1, AFFDL-TR-66-15, Contract AF33(616)-7501 (1965).
3. J. I. Elkind: Characteristics of Simple Manual Control Systems, M.I.T. Lincoln Laboratory TR-11 (1956).

4. T. W. Beeker, J. During and A. Den Hertog: Artificial Touch in a Hand Prosthesis, Medical and Biological Engineering, Vol. 5, 47/49 (1967).
5. I. Kato, S. Yamakawa, K. Ichikawa and M. Sano: Multifunctional Myoelectric Hand Prosthesis with Pressure Sensory Feedback -- Waseda Hand 4P, Advances in External Control of Human Extremities, 155/170 (1970).
6. R. E. Prior and J. Lyman: Electrocutaneous Feedback for Artificial Limb - Summary Progress Report, Bulletin of Prosthetics Research, Fall, 3/37 (1975).
7. R. N. Scott, R. H. Brittain, R. R. Caldwell, A. B. Cameron and V. A. Dunfield: Sensory-Feedback System Compatible with Myoelectric Control, Medical Biological Engineering & Computing, Vol. 18, 65/69 (1980).
8. L. B. Durr: Effect of Error Amplitude Information, M. A. Thesis, University of Virginia, Charlottesville (1961).
9. J. C. Bliss: Tactile Perception: Experiments and Models, Stanford Research Institute, Menlo Park, California, Final Report, Contract NAS-3649, SRI Project 6070 (1967).
10. F. A. Geldard: Cutaneous Channel of Communication, Sensory Communication (Ed. W. A. Rosenblith), The M.I.T. Press, 73/87 (1961).
11. J. Hirsh and I. Kadushin: Experiments in Tactile Control of Flying Vehicles, 6th Annual Conference on Aviation and Astronautics, 64/68 (1964).
12. W. C. Howell and G. E. Briggs: An Initial Evaluation of a Vibrotactile Display in Complex Control Tasks: Ohio State University, Columbus, Technical Report, (813)-5, AD-320 (1959).
13. S. Weissenberger and T. B. Sheridan: Dynamics of Human Operator Control Systems Using Tactile Feedback, Journal of Basic Engineering, 297/301 (1962).
14. H. F. Seely and J. C. Bliss: Compensatory Tracking with Visual and Tactile Displays, IEEE trans., Human Factors in Electronics, Vol. HFE-7, 84/90 (1966).
15. D. S. Alles: Information Transmission by Phantom Sensation, IEEE Trans., Man-Machine Systems, Vol. MMS-11, 85/91 (1970).
16. J. W. Hill: Describing Function Analysis of Tracking Performance Using Two Tactile Displays, IEEE Trans., Man-Machine Systems, Vol. MMS-11, 92/100 (1970).
17. A. Y. Szeto: Electrocutaneous Codes for Sensory Feedback in Prostheses and Orthoses, Ph.D. Dissertation, University of California, Los Angeles, (1977).
18. H. P. Schmid and G. A. Bekey: Tactile Information Processing by Human Operators in Control Systems, IEEE Trans. on System, Man and Cybernetics, Vol. SMC-8, 860/866 (1978).
19. A. B. Anani and L. M. Körner: Afferent Electrical Nerve Stimulation: Human Tracking Performance Relevant to Prosthesis Sensory Feedback, Medical Biological Engineering & Computing, Vol. 17, 425/434 (1979).
20. T. B. Sheridan and W. R. Ferrel: Man-Machine Systems, The M.I.T. Press (1974).
21. S. Tachi, K. Tanie and M. Abe: Experiments on the Magnitude Sensation of Electrocutaneous Stimuli, Bulletin of Mechanical Engineering Laboratory, No. 30 (1978).
22. K. Tanie, S. Tachi, K. Tani, Y. Maeda, T. Ohno, A. Fujikawa and M. Abe: Development of Multichannel Simultaneous Stimulator, Journal of Mechanical Engineering Laboratory, Vol. 31-2, 104/116 (1977).

23. K. Tanie, S. Tachi, K. Komoriya and M. Abe: Study of Electrocutaneous Parameters for Application to Dynamic Tactual Communication Systems, First Mediterranian Conference on Medical and Biological Engineering, Digest of Papers, Vol. II, 17-65/17-68 (1977).
24. K. Tanie, S. Tachi, K. Komoriya, M. Abe and M. Miura: Electrocutaneous Information Transmission System Using Stimulus Energy Control Device, Japanese Journal of Medical Electronics and Biological Engineering, Vol. 18-1, 40/42 (1980).
25. K. Tanie, S. Tachi, K. Komoriya and M. Abe: Study on Frequency Dimension in Electrocutaneous Stimulation, Journal of Mechanical Engineering Laboratory, Vol. 33-4, 159/170 (1979).

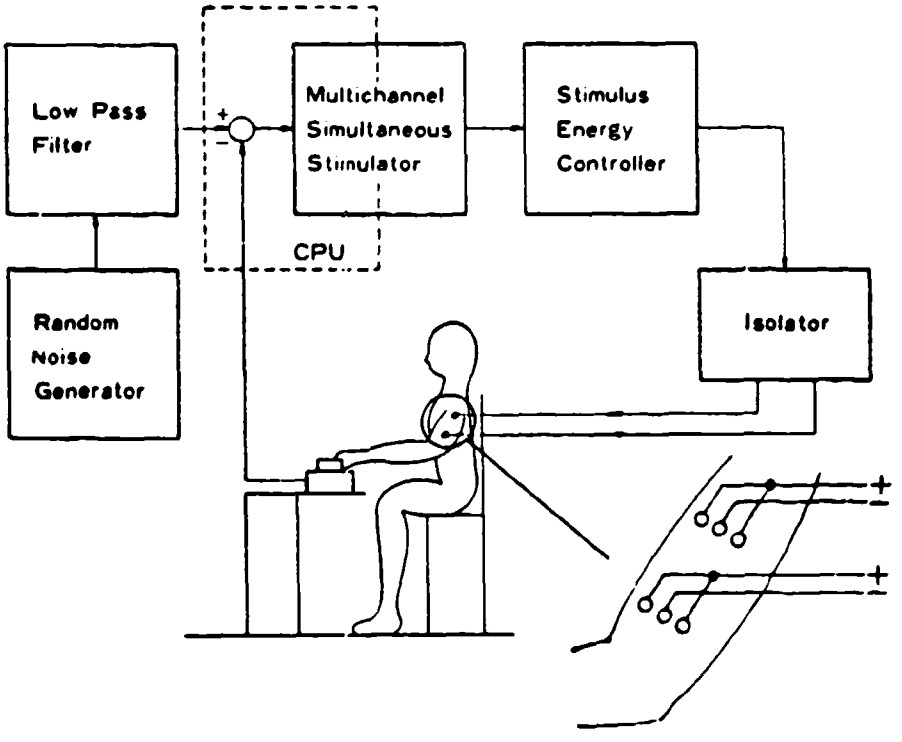


Fig. 1 Experimental Setup.

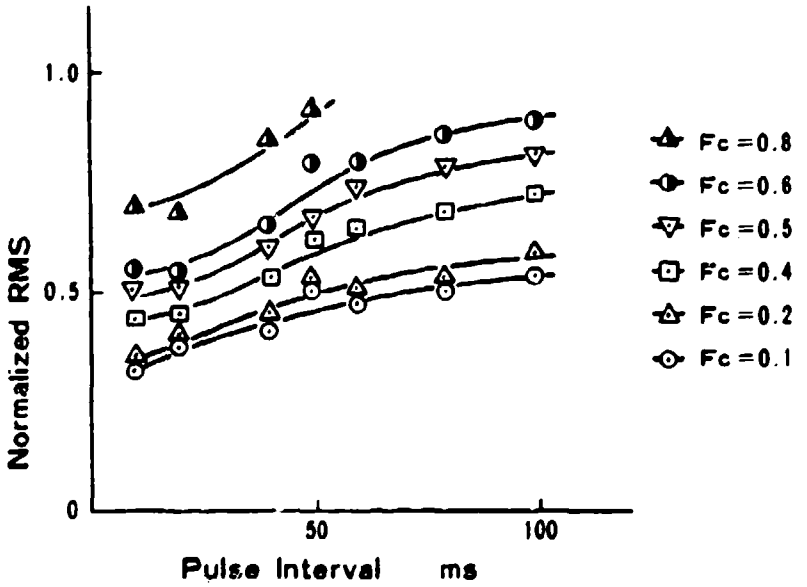
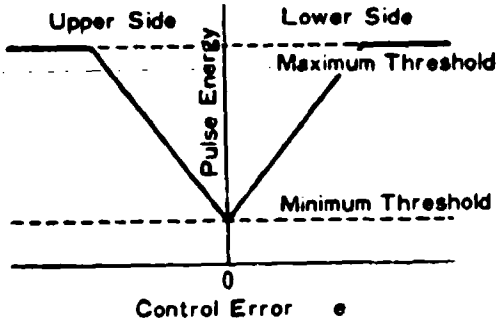
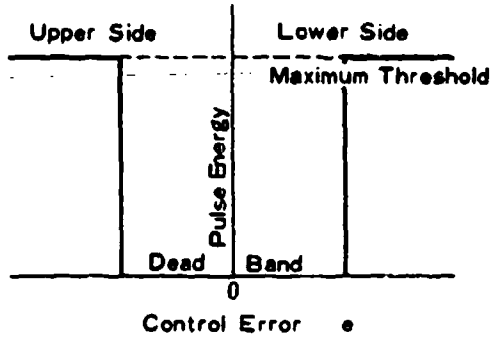


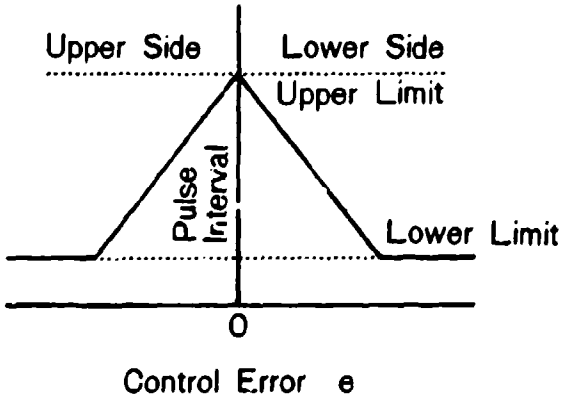
Fig. 2 Relation between Pulse Interval and Normalized RMS Value of Control Error (F_c: Cutoff Frequency of Low-Pass Filter Used to Limit the Bandwidth of Command Signal).



(a) Linear Transformation.

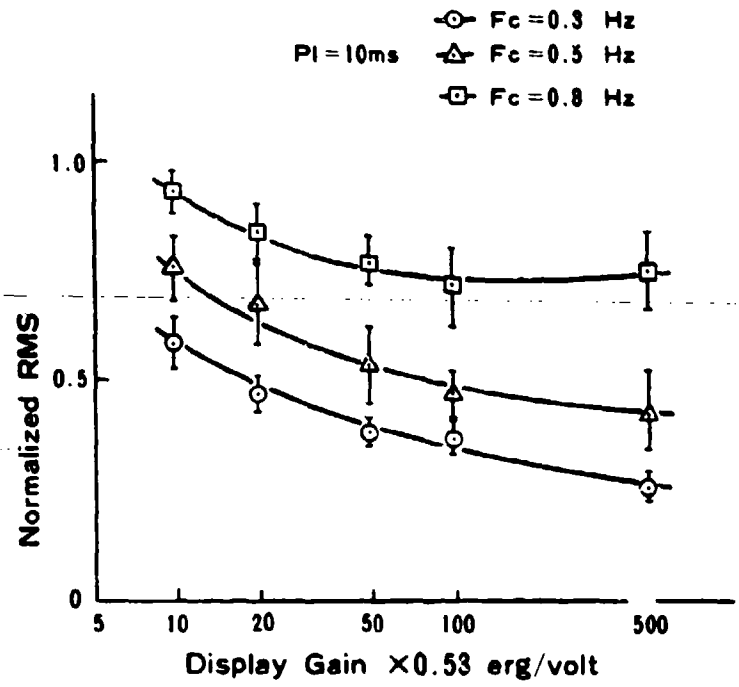


(b) Threshold Transformation.



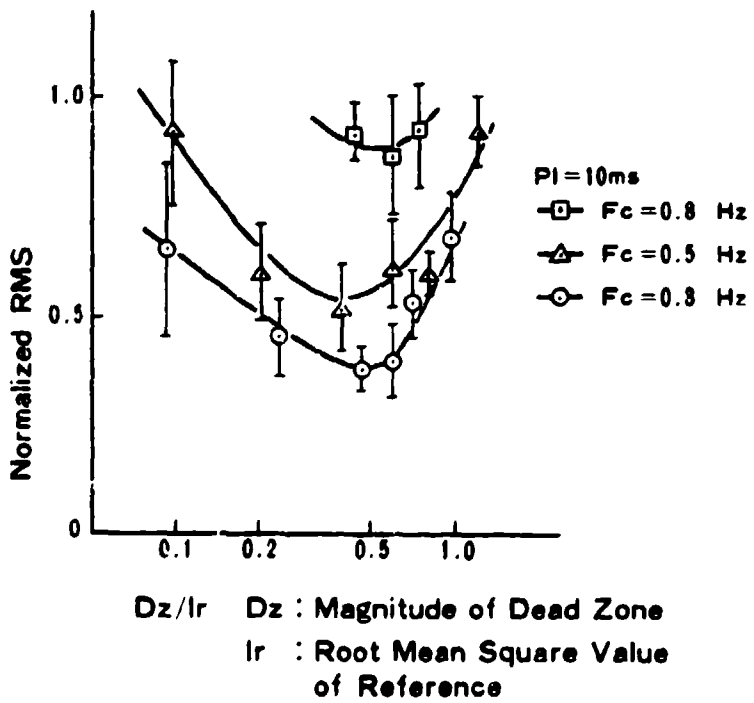
(c) Frequency Transformation.

Fig. 3 Display Transformations



(a) Linear Transformation.

Fig. 4 Relation between Display Gain or Dead Band and Normalized RMS Value of Control Error (PI: Pulse Interval, Fc: Cutoff Frequency of Low-Pass Filter Used to Limit the Bandwidth of Command Signals).



(b) Threshold Transformation.

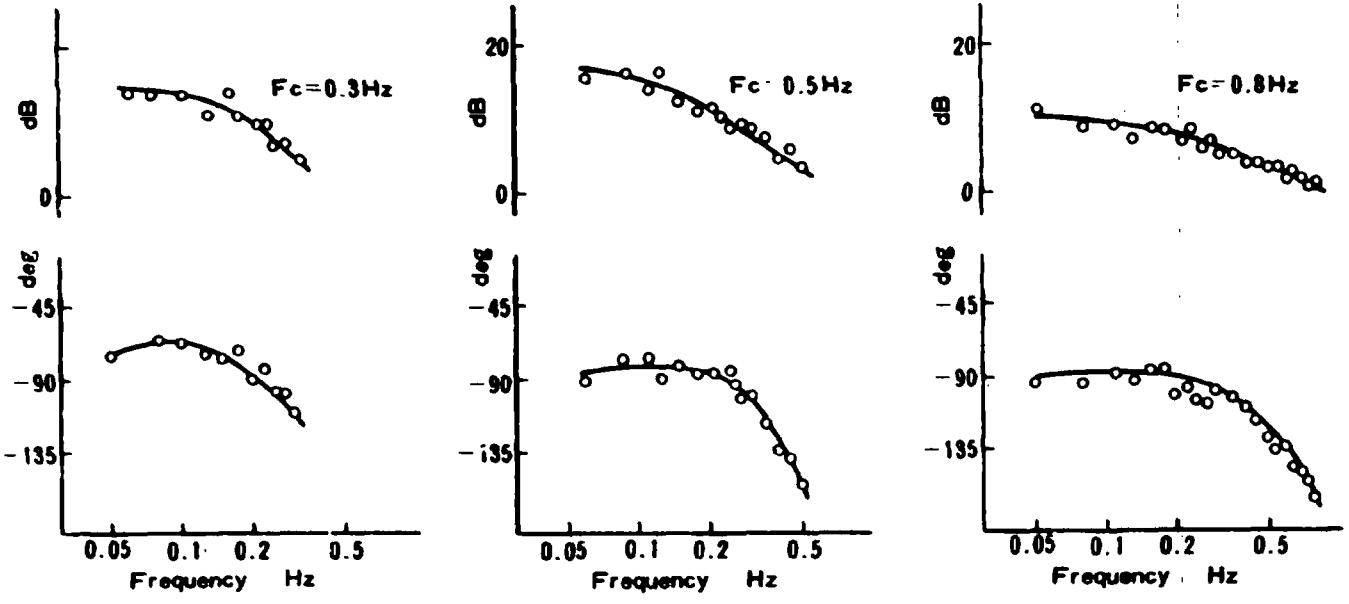


Fig. 5(a) Operator Transfer Functions(Linear Transformation).
(F_C: Cutoff Frequency of Low-Pass Filter Used to Limit the Bandwidth of command signals)

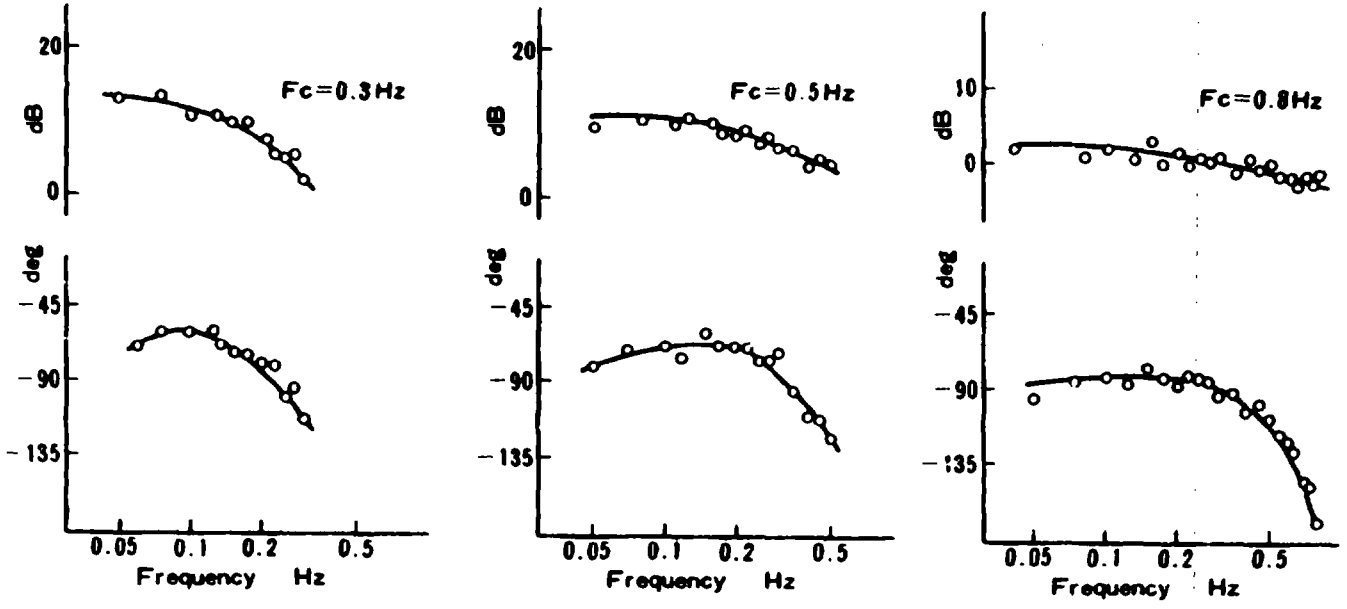


Fig. 5(b) Operator Transfer Functions(Threshold Transformation). (F_C: Cutoff Frequency of Low-Pass Filter Used to Limit the Bandwidth of Command Signals)

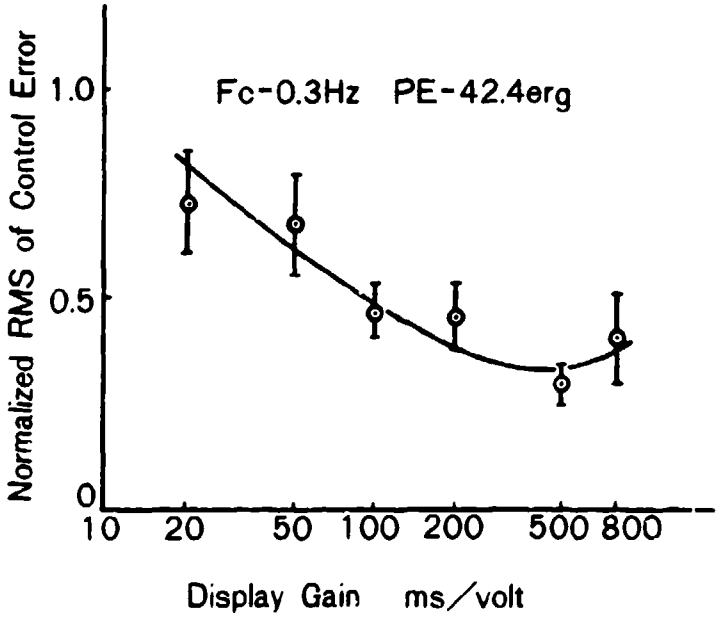
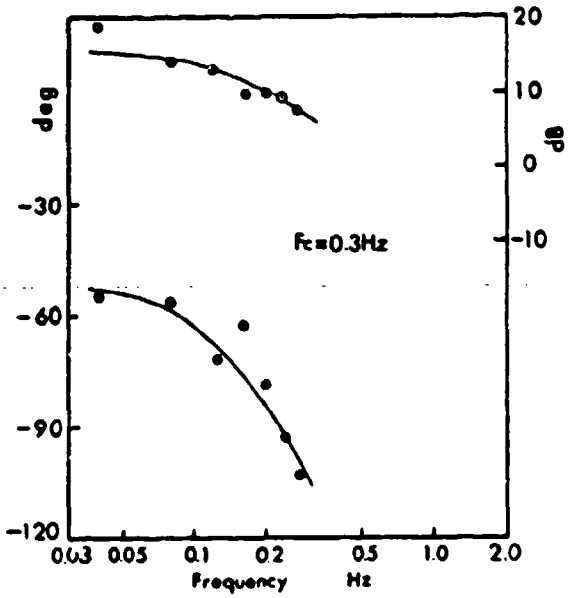
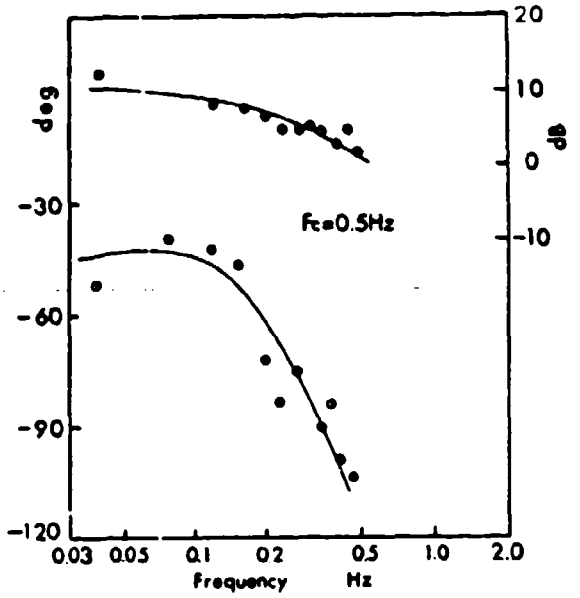


Fig. 6 Relation between Display Gain and Normalized RMS Value of Control Error under the Use of Frequency Transformation(F_C: Cutoff Frequency of Low-Pass Filter Used to Limit the Bandwidth of Command Signals, PE: Stimulus Pulse Energy).





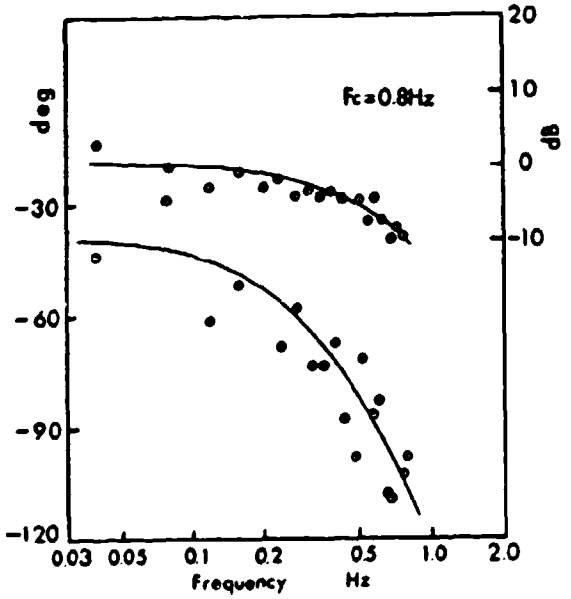


Fig. 7 Operator Transfer Functions (F_C: Cutoff Frequency of Low-Pass Filter Used to Limit the Bandwidth of Command Signals, Stimulus Pulse Energy=42.4 erg).

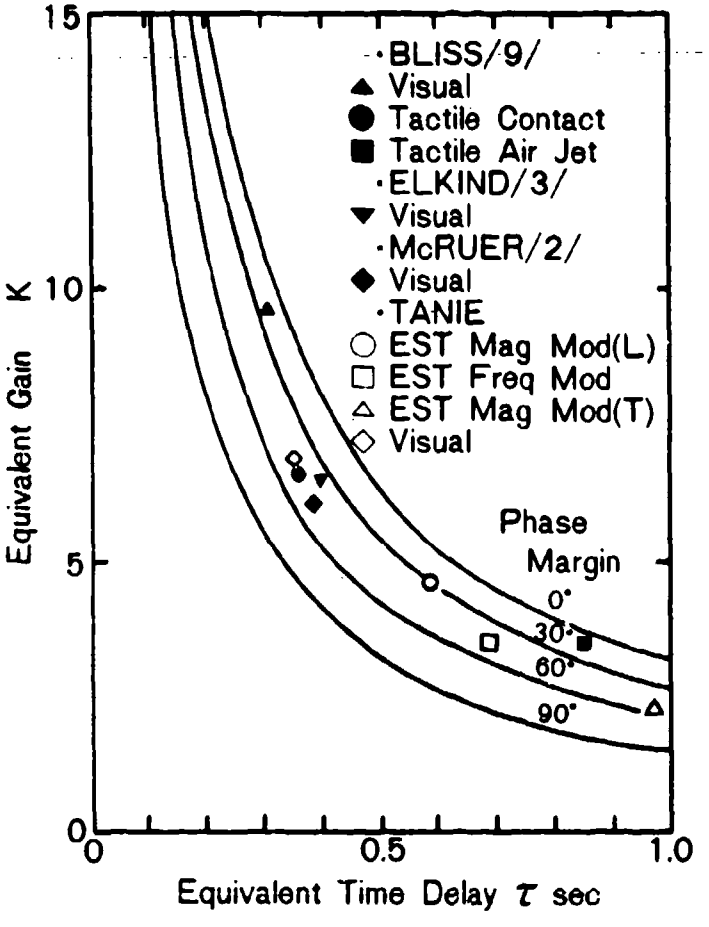


Fig. 8 Comparison of Displays(EST: Electrocutaneous Stimulation, Mag Mod(L): Linear Transformation Using Magnitude Code, Freq Mod: Frequency Transformation, Mag Mod(T): Threshold Transformation Using Magnitude Code, /2/,/3/,/9/: Indicate the Numbers of References).

SESSION 5: HUMAN-MACHINE SYSTEMS AND CONTROLS

Moderator: Ronald A. Hess, NASA Ames Research Center

5

AD P001638

A TRANSFORM APPROACH TO COLLISION FREE PROGRAMMING OF MULTIDEGREE-OF-FREEDOM STRUCTURES

by

J. G. Kreifeldt
S. H. Levine
Department of Engineering Design
Tufts University, Medford, MA.

Paper Presented at the 18th Annual Conference on Manual Control. Dayton, Ohio, June 1982

ABSTRACT

An approach is presented which transforms a dynamic articulated, collision-prone machine, such as a remote manipulator or robot, into a static multidimensional representation of its collision possibilities. Among other advantages of this approach, the basic problem of finding collision free and optimized machine movements from an initial to a final configuration becomes immediately clear and conceptually solvable. The major disadvantages of the approach at this point appear to be related to computational procedures and data storage. However, the clarification of the control and guidance problem gained through the transformational approach as well as its elucidation of many related problems remain prime conceptual tools. This paper presents a number of aspects of this approach as well as several simplified examples.

INTRODUCTION

Partial or fully automated, flexible computer control of multidegree of freedom structures such as robots, remote manipulators, numerically controlled machines, or other articulated structures faces two basic problems.

(1) The automatic determination of an integrated course of movement (trajectory) of each degree of freedom such that the change from machine configuration A to configuration B (e.g. initial to final configuration) can proceed in some optimal fashion without the machine structures or pieces undesirably colliding with each other or with surrounding structures.

(2) Execution of or movement control along the integrated trajectory with a precision sufficient to ensure achievement of the goal while maintaining acceptable safe tolerances from possible collisions.

There are four distinct direct approaches to the solution of these problems.

1. The artificial intelligence approach would try to emulate "intelligent" behavior by equipping the structure with sensors for proximity, vision, etc. and algorithms intended to allow the machine to find its way safely from A to B.

2. The repertoire approach uses preprogramming of all possible movements that may be required and executes these as needed as in automated assembly.

3. Model mimic techniques use a faithful geometric computer model of the machine which can be moved according to an ad hoc trial plan with collision possibilities predetected by checking the position of the model machine's structures at points along the trajectory. A collision free check would imply an acceptable trajectory.

4. The transformational approach maps all of the collision possibilities of the dynamic three dimensional geometric machine with its infinity of configurations into a set of finite zones within a static hyperdimensional configuration space. Any possible machine configuration is a point (vector) in this configuration space and any machine movement no matter how dynamically or spatially complex is a point trajectory in this space. Finding optimal and collision free machine movement corresponds to finding trajectories through only the collision free areas.

Each of the above methods has its own advantage and disadvantages. The artificial intelligence approach puts considerable premium on the invention of sensors and heuristic algorithms but lacks a systematic interpretation and approach to the general problem. Neither does it offer a methodology for designing dynamic articulated structures so as to reduce collision proneness.

The repertoire technique is pragmatic and practical and gets specific jobs done. It is best suited to repetitive, fixed movements although there is some capabilities for synthesizing new movements from existing elements. However, it suffers the same drawbacks as given above for the artificial intelligence approach.

The model mimic technique is a feasible approach whose main drawbacks are that (1) by itself it does not offer any methodology for automatic determination of collision free trajectories and (2) point-by-point prechecking of a trajectory becomes increasingly time consuming as the required probability of not detecting an impending collision diminishes to zero. As with the previous two approaches, this technique does not provide a general understanding of the problem nor offer guidance to the design of less collision prone machines.

The transformational approach has the drawback that it is computationally challanging at present and the result may suffer from lack of visualization. It trades a three dimensional visually recognizable but movement complex dynamic structure for a conceptually simple and static but hyperdimensional construct. It is similar in this respect to linear programming which transforms a lower dimensional problem of inequalities into a higher dimensional one of equalities for the sake of a uniform solution methodology and provides deeper insights into the original problem. The major advantages of the transformation approach include a unified framework for representing and solving the problem of the automated determination of optimal collision free machine movements. It also offers the potential for performing a reverse transformation for designing a machine of reduced collision proneness. Further advantages include the ability to represent the margin of tolerance at the closest approach between machine elements which is useful in conjunction with control algorithms for actually moving the machine according to the

optimally determined trajectory.

It is likely that the ultimate solution to computer automation and control of multidegree of freedom structures will be a combination of the four approaches. However, this paper will pursue development of the transformational approach in order to define its potentials and produce a generally useful theoretical and practical tool for computer automation in this class of problems i.e. articulated structures.

CONFIGURATION SPACE AND COLLISION ZONES

Before describing the transformational approach, its two basic operational problems are stated.

1. Determination of collision zones (colliding machine configurations) in configuration space, and

2. Definition and determination of optimal collision free trajectories in configuration space.

1. Configuration Space

A "Machine" can be defined for our purposes as any set of rigid structural elements whose relative orientations one to another and to a frame of reference can be altered. Except for pedagogical purposes, most machines have a purpose e.g. a remote manipulator, robot etc. At some level a machine can be dissembled into its next level structural elements. The level we chose is that in which the elements are separately identifiable and potentially capable of colliding one with another or surrounding objects. The separate elements are the primitives of the machine. Figure 1 shows a schematic drawing of a radiotherapy machine and Figure 2 shows its reduction to its primitives.

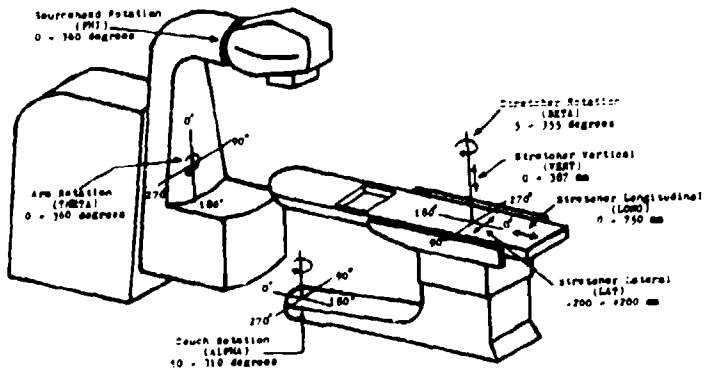


Figure 1. A Radiotherapy Machine

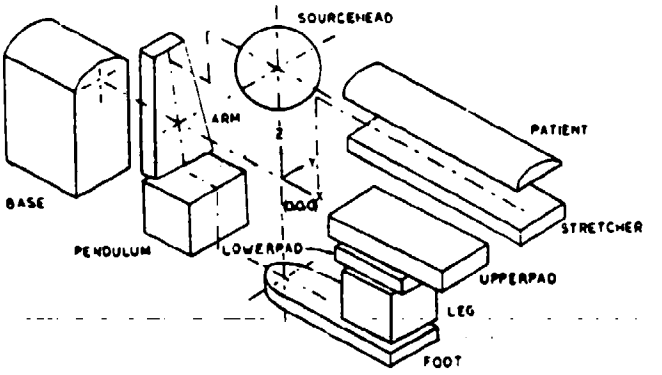


Figure 2. The Radiotherapy Machine Reduced to its Primitive Elements for Computer Modeling

In addition to its separability, each primitive element has one or more degrees of freedom for its movement. In general, any rigid body has 3 translational (x, y, z) and 3 rotational (pitch, roll, yaw) degrees of freedom. Practically, however, most such elements have less than six degrees of freedom. For machine control, it is essential that the translational and rotational orientation of each element be knowable, e.g. by position feedback potentiometers. This information on each degree of freedom of each element may be absolute (to an external frame of reference)or relative (e.g. to its connecting neighbor). Any configuration of the machine is then completely specified by the vector of values of each degree of freedom. The collection of degrees of freedom, absolute or relative, form an orthogonal basis for a hyperdimensional space (CONFIGURATION SPACE) in which any machine configuration is uniquely represented by a vector (point) in this space. It is worth emphasizing that this point in configuration space may have different directional dynamics as well as being the unique representation of the machine configureation. Any path (trajectory) traced by this point in configuration space defines a unique sequence of machine configurations.

2. Collision Zones

A. Concept

By itself, configuration space as described is not a very illuminating concept although for example it can quantify the similarity or proximity of two different configurations by the scalar distance between the corresponding configuration points. The essential concept for machine control is the presence of COLLISION ZONES within Configuration Space. In the simplest case, a collision zone is that compact set of points in configuration space representing all the values of the degrees of freedom of two machine primitives which place them in collision with each other. If the two elements have 5 degrees of freedom between them, and if the two can collide, the collision zone will be essentially five dimensional. Thus a second feature of configuration space is that there will be at most as many collision zones as

there are pairs of collidable machine elements although two or more zones may intersect.

As an illustration, Figure 3 is a simple machine with two elements each with a single degree of freedom. Although it does not seem to have an obvious purpose this will be discussed later.

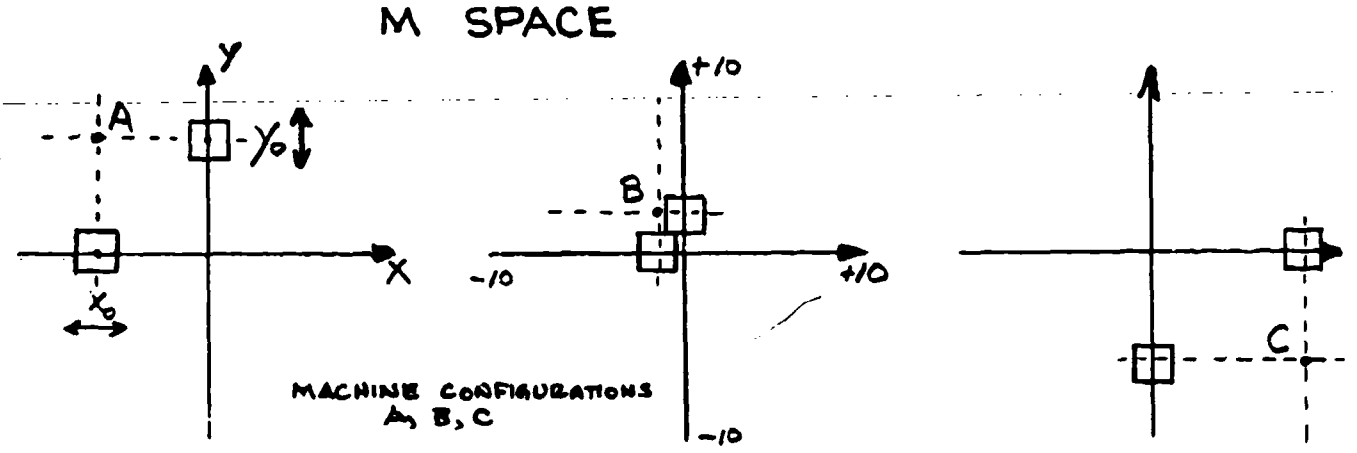


Figure 3. A Simple Two Degree of Freedom Machine

Configuration space for this machine is likewise two dimensional. The machine's configuration is defined by the x coordinate of the center of one element and the y coordinate of the center of the other. This (x, y) coordinate specifies the configuration of the machine and is a point in a configuration space whose orthogonal basis is the (x, y) degrees of freedom. The two elements can collide for certain values of their degrees of freedom and all of these values are represented by the rectangular collision zone in configuration space as in Figure 4.

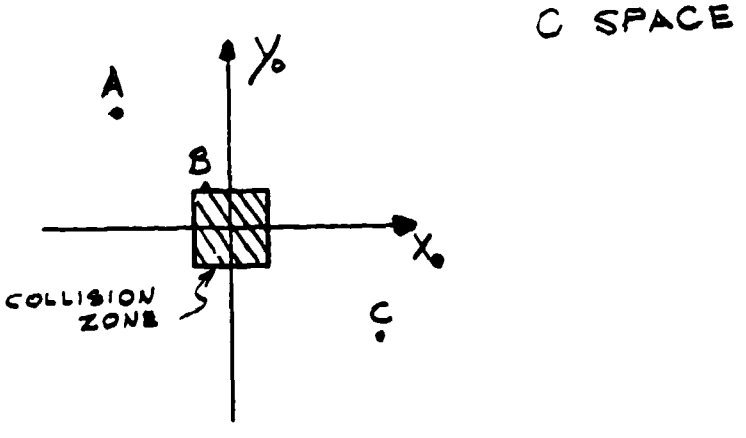


Figure 4. Configuration Space with Collision Zone for a Simple Machine

This Configuration Space with its Collision Zone is a transformation of the machine. The 3 points in configuration space uniquely identify 3 different machine configurations, one of which (8) indicates that the two pieces are colliding. Although it would seem that the interior of a collision zone is unreachable and thus not significant because it would involve penetration of the pieces, in fact there is a utility to the concept of interior points as will be discussed later.

Thus a dynamic realworld machine can be transformed into a static representation in which all of the conceivable colliding configurations of its parts are immediately obvious by the collision zone (s). This simple example can also be used to show that the transformation from M (machine) space to C (configuration) space is not unique. Figure 5 shows two apparently different machines which have an identical transformation.

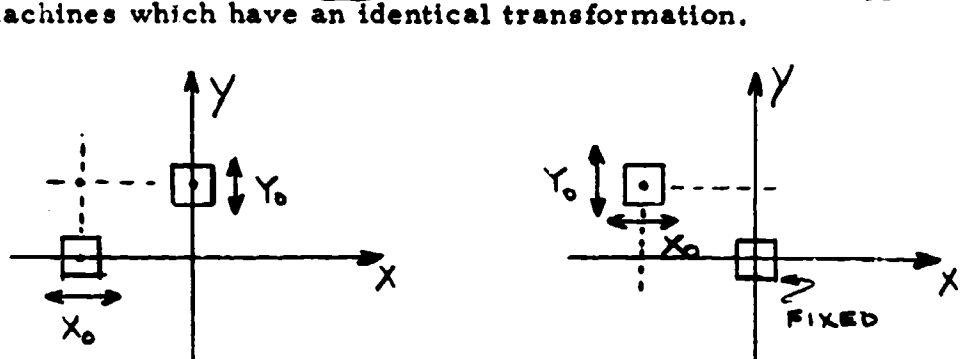


Figure 5. Two 2 Degree-of-Freedom Machines with an Identical Transformation

Thus a many-to-one mapping from M space to C space exists. This poses an interesting question. If different machines have the same transform to C space, what are their basic similarities in M space other than having the same number of degrees of freedom? A possible interpretation comes from study of the purpose of machines A and B in Figure 5. If the purpose is to grind the two blocks by keeping them in contact then either mechanization may be used. In fact, machine A can be considered as machine B with the frame of reference in B taken on the center of the vertical block. From that reference point, it appears that the horizontal block moves around it as in machine A. Perhaps other machines exist with the same C space transformation. Although the mapping from M to C is straightforward, if tedious, no procedure has been developed for performing the inverse C to M transformation. Because the two machines of Figure 5 have the same C space transformation, they may be called isoforms of each other.

Another example of two machines with the same transform to C space will be given later.

B. M Space to C Space Transformation - Collision Zones

The preceding example had a close topographical correspondence

between C space and M space and between machine elements and collision zone. In fact it is known( ) that if two elements are convex polyhedra and move only in translation without rotation, then their collision zone will be a convex polyhedron as well. That is, in general two 3 dimensional convex polyhedral elements, each free to move in 3 translational degrees of freedom, generate a 6 dimensional convex polyhedral hull in C space.

However, if at least one element is non convex and/or at least one element has at least one rotational degree of freedom there is no guarantee that the collision zone in C space for this pair is convex or even polyhedral non convex. The utlility of convexity and polyhedrons will be discussed later.

An example of non convex non polyhedral collision zone transformed from a simple machine with rotary and translational movement elements is shown in Figure 6.

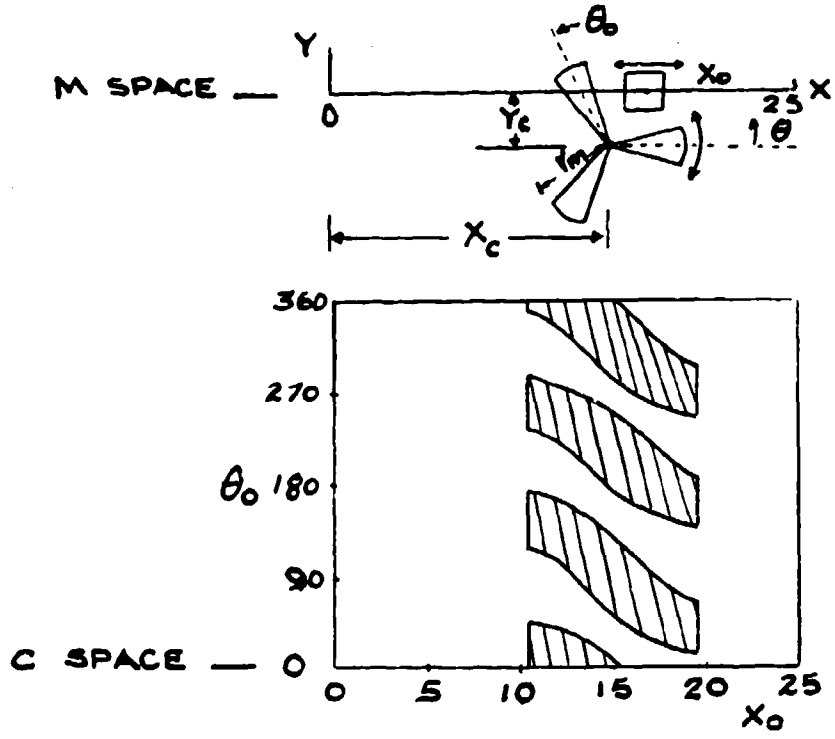


Figure 6. A Simple Rotary and Translational Machine Transforming to a Non-Convex Non-Polyhedral Collision Zone

As a further example, Figure 7 shows a planar machine with two translating and one rotating element with a total of three degrees-of-freedom (x, y, θ) This machine and its resulting transform can be considered a type of superposition of the previous two simple machines of Figures 3 and 6.

These three degrees of freedom form the basis for a 3 dimensional C space. There are three collision zones which intersect because of the location of the machine elements. Figure 8 shows this union of the 3 separate zones.

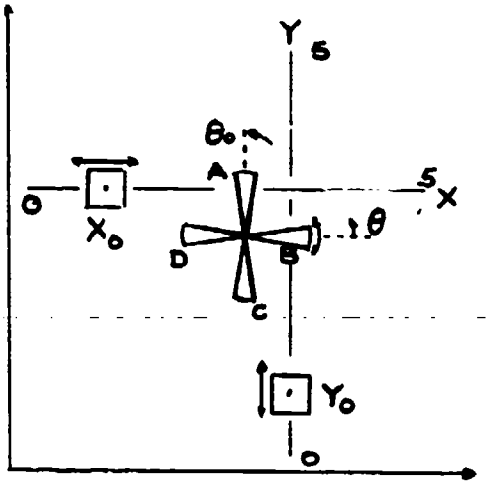


Figure 7. A Simple Three Degree of Freedom Machine

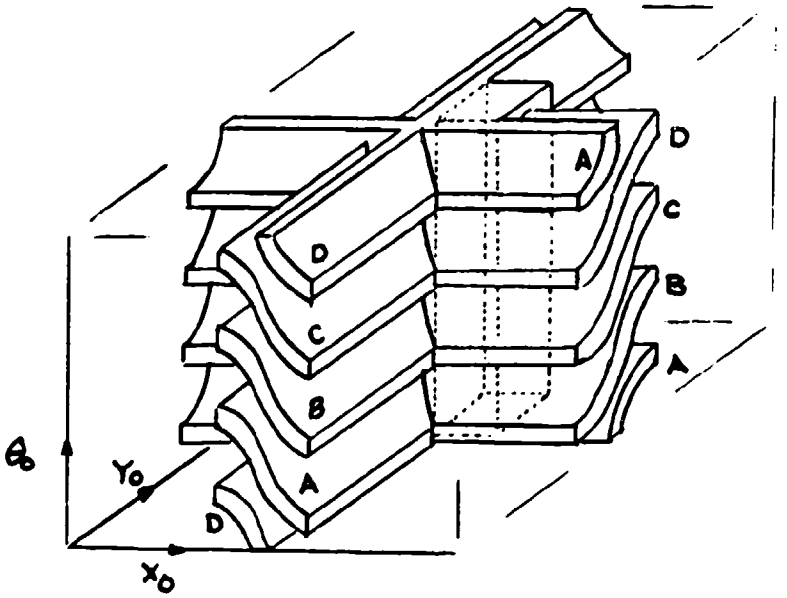


Figure 8. Configuration Space and Composite Collision Zone for the Simple Machine in Figure 7

Even though only one "zone" is apparent it should be remembered that a zone represents machine collision configuration for a pair of elements and the separate collisions by pairs can still be identified by references to each zone. In general, if there are n machine elements there can be at most $\frac{n(n-1)}{2}$ collision zones all or some of which may intersect to form one or more composite zones. Clearly, even if each individual zone is a convex polyhedron, their union need not be.

Because each blade of the fan element can be involved in a collision with

a slider block, there are multiple fan blade-block zones. A second feature of C space that may seem obvious but is worth mentioning is that each dimension of the space can be scaled to plot to the same length which is a useful concept in trajectory related issues.

C. Determination of Collision Zones

There are 3 approaches to obtaining the collision zone for a pair of collidable elements. These are: (1) empirical, (2) Monte Carlo, (3) correlation computations.

1. The empirical approach would utilize the machine itself or a faithful model of it which can be (laboriously) manipulated so as to map out the zone for an element pair. Although tedious, this may be the only tractable way to obtain the zone for complex element shapes. This technique was actually used to determine the 3 dimensional collision zone for the sourcehead-stretcherfoot element pair of an actual machine schematized in Figure 1 in which collisions between these two elements depend on the arm rotation angle, head rotation angle and foot rotation angle. The zone produced is "slightly" concave and not polyhedral.

2. The Monte Carlo approach requires a faithful computer stored geometric model of the machine's primitive elements. Basically, a random point in C space is selected and model given the corresponding orientation. The collision status of all pairwise elements than can possible collide is then checked. Pairs of elements that can not collide should be initially eliminated to save time. The intersection of a pair of elements can be determined by several techniques depending upon their geometries and degrees of freedom. These techniques include: analytic geometry and projective geometry.

i)the analytic geometry approach utilizes some type of interpiece distance calculation and is best suited for very regular convex objects, e.g. two spheres, or a sphere and cylinder with spherical end caps, etc.

ii) in the projective geometry technique, the two structures in question are projected onto suitably oriented orthogonal planes. Convex polyhedral structures are in collision if and only if their convex projections overlap on two orthogonal projections. In the case of a sphere, overlap in all 3 orthogonal projections is necessary and sufficient. In one sense, the projective method is fundamental for determining intersection of any two elements regardless of their geometries. The difficulty of course is in finding a suitable set of orthogonal planes for checking overlapping projections. The projection technique mimics the "moving eye" method of checking for collisions. If a "hit" is signaled for collision between two elements, the configuration point is stored for that pair (collision zone). Optimized search techniques (e.g. Fibonacci, golden mean, or others) may be used to search for the boundary of the zone along a dimension in C space. Obviously the Monte Carlo approach needs considerable sophistication in application. However, collision zones need only be computed once for any machine and this can be done off-line. Considering the time required to design and construct any machine, the time to transform it to C space would seem a reasonable investment if it leads to improved automation control and/or design.

In either the analytic or projective geometry technique, the utility of convex polyhedral or regular conic solids for machine elements is obvious. As the machine elements acquire sculpting such as compound and developed surfaces, any mathematical methods for accurately determining collision zones become increasingly difficult. Attempting to model an actual machine by regular primitives such as is done in CAD/CAM applications is an obvious approach. The application of CAD/CAM in design may in fact be an essential step in the M to C transformation. Hidden line algorithms may be related to the transformation problem.

Computation of collision zones by a correlation technique is outlined below.

3. Computation of the Collision Zone by Correlation Type Methods

i. Some specific examples.

The determination of the collision zone requires a mathematical description of each component and its movement, and the determination of when different components intersect. As a simple example consider the machine of Figure 3. Defining

$$p(x/b) \equiv \begin{cases} 1, & |x| \le b \\ 0, & \text{otherwise} \end{cases}$$

$$q(x/b) \equiv \begin{cases} 1 - \frac{|x|}{b}, & |x| \le b \\ 0, & \text{otherwise} \end{cases}$$

the two squares are then respectively

$$f_x(x-x_0, y) = p((x-x_0)/1)\, p(y/1),$$

$$f_y(x, y-y_0) = p(x/1)\, p((y-y_0)/1).$$

The degree to which they intersect, as a function of x_0 and y_0, is given by correlation type expression

$$z(x_0, y_0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_x(x-x_1, y)\, f_y(x, y-y_1)\, dx\, dy \qquad (1)$$

which reduces to

$$z(x_0, y_0) = \int_{-1}^{1} p((x-x_0)/1)\, dx \int_{-1}^{1} p((y-y_0)/1)\, dy$$

Noting that

$$\int_{-1}^{1} p((x-x_0)/1)\, dx = \begin{cases} x_0+2, & -2 \le x_0 < 0 \\ 2-x_0, & 0 \le x_0 \le 2 \\ 0, & \text{otherwise} \end{cases}$$

which is in turn $2q(x_0/2)$, we can write

$$z(x_0, y_0) = 4q(x_0/2)\, q(y_0/2),$$

a rectangular pyramid whose base is the collision zone of Figure 4 and whose maximum height, at $x_o = 0$, $y_o = 0$, is equal to 4.0. The height of $z(x_o, y_o)$ is a measure of intersection (collision severity) between the two components of the machine.

A similar approach could be used in computing the collision zone for the machine in Figure 6. The propeller is conveniently described in polar coordinates and the resulting mix of polar and rectangular coordinates leads to significantly greater mathematical difficulties.

These two examples represent machines where degrees of freedom are independent. Thus, the location in M space of any component is a function of only one degree of freedom. By contrast, for a multiple joint arm individual movable joints are linked together and more complicated expressions result.

ii. More generalized considerations

In the previous examples of the sliding squares the collision zone was calculated using a spatial correlation type equation. Equation (1) can be generalized as

$$z(x_o, y_o) = \int_Y \int_X f_x(x - x_o, y)\, f_y(x, y - y_o)\, dx\, dy \qquad (2)$$

with f_x and f_y the corresponding shapes (i.e., squares in the example shown). A change in coordinates,

$$u = x - x_o ,$$

leads to

$$z(x_o, y_o) = \int_Y \int_X f_x(u, y)\, f_y(u + x_o, y - y_o)\, du\, dy ;$$

this expression is similar to the two dimensional cross correlation, ϕ_{xy}, of f_x and f_y. Specifically

$$z(x_o, y_o) = \phi_{xy}(-x_o, y_o)$$

where the $-x_o$ results from the $u + x_o$ rather than $u - x_o$ term. Thus $z(x_o, y_o)$ is the cross correlation $\phi_{xy}(x_o, y_o)$ reversed in the x(or u) direction. The outline of $z(x_o, y_o)$ can thus be determined by holding f_x stationary, moving f_y around it such that the two remain tangent, recording the shape outlined by the coordinate center of f_y and then reversing the resultant figure in the x direction. (This recalls the empirical method for collision zone determination.) The previous example, where f_x and f_y were squares, obscures this reversal due to symmetry, but a later example will illustrate this point.

iii Orthogonalizing a Machine

The question arises as to how general the correlation approach is for planar figures. Consider the machine shown in Figure 9a. The two squares no longer move in orthogonal directions; in particular, the expression for the block located by y_o must change since its value as a function of x is now dependent on y_o. As a result the simple correlation relationship of equation(2)

no longer holds.

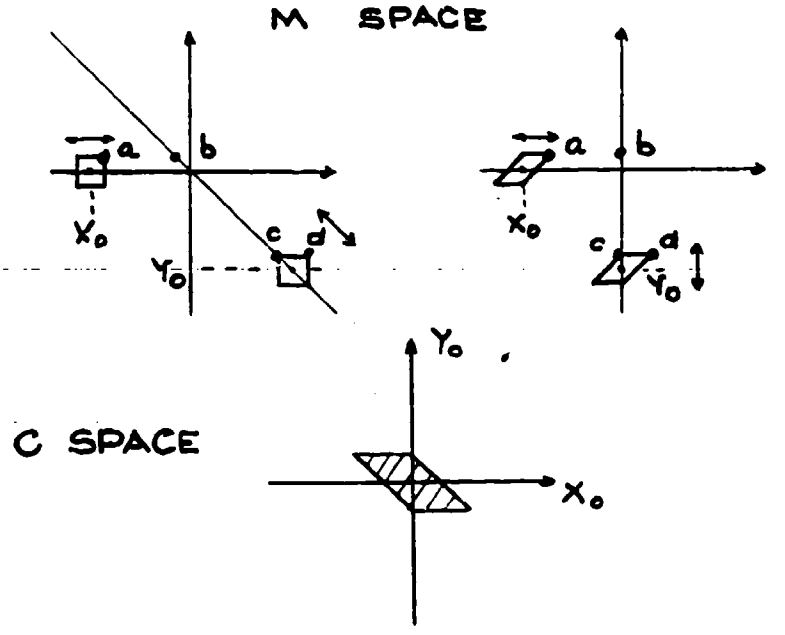


Figure 9. A Non-Orthogonal Machine, Its Collision Zone and Orthogonal Isoform

One approach is to seek an equivalent orthogonal machine where by equivalence we mean generating the same collision zone (an isoform). This collision zone, determined by trial and error, is shown in Figure 9b. Figure 9c indicates an equivalent orthogonal machine, again determined from Figure 9b, by trial and error. Note that given the orthogonal machine the collision zone could then be determined using the correlation approach, that is, by holding the horizontally sliding component and moving the vertically sliding component tangentially around it. The necessary reversal of this resulting shape becomes evident in this case.

Of course the approach taken here does not follow the desired order. What we wish to do is to determine the equivalent orthogonal machine directly and then utilize the correlation approach to determine the collision zone. Inspection of Figures 9a and 9c indicate that a simple transformation suffices. Note that the horizontal distance ab is conserved where b is on the track of the "vertically" moving object. The same holds for all other horizontal distances. Thus the simple coordinate transformation

$$(X, Y) \rightarrow (X + Y \tan\Theta, Y)$$

allows determination of the orthogonal machine. A comparison of the collision zone in Figure 9b to that in Figure 4 indicates the possibilities for directly transforming the collision zone itself.

This orthogonalizing technique has potential application to the computation of collision zones when components move along curved trajectories. Figure 10 indicates such an example. An incremental region is indicated for

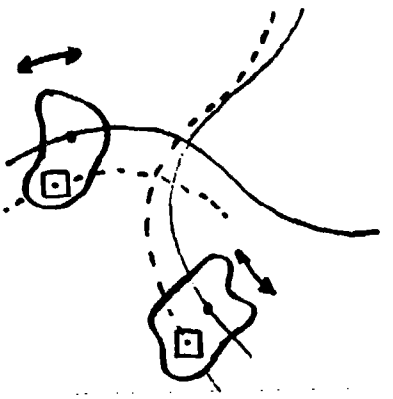


Figure 10 Differential Elements in Collision Along Curved Paths

each component. The trajectories of these incremental regions intersect and contribute to the overall collision zone. These latter trajectories can be approximated near their intersection by two straight lines and the orthogonalization approach can be used.

The examples chosen for illustration resulted in no more than a 3 dimensional C space. Obviously a machine such as in Figure 1 has a nonvisualizable C space. However, the hyperdimensional C space and collision zones do not differ in priniciple from those for a 3-D C space.

The problem of inverse transformation from C space to a generating machine of elements and degrees of freedom has not been approached in any detail yet. The apparent many-to-one M-to-C space transform of Figure 5 would suggest that the inverse transform will have to be carefully considered. A basic question is whether the two machines of Figure 5 are in fact "different" because of their movement structure or identical in that their purpose is to grind one block against another. If a third machine with the same transform could be found whose structure and purpose are both essentially different from the two machines, it would suggest a fundamental problem in inverse transformation (but not in the utility of the M to C transformation). Inverse transformation is useful for design purposes. The forward transformation is useful for control purposes.

Because a collision zone represents restricted machine configurations, the notion may be extended to limitations on the range of movement of a degree of freedom regardless of collision concerns. Such a zone may occur as a "wall" in C space to define the extent of an axis. An unrestricted rotational degree of freedom will generate a periodic structures (as in Figure 8).

C. OPTIMAL TRAJECTORIES IN C SPACE - MACHINE CONTROL

I. INTRODUCTION

As mentioned previously, the M to C operation transforms a dynamic,

geometric machine with many moving parts to a static representation in which the instantaneous machine configuration is a point in C space. The control problem is to change the machine from configuration A to B without producing any undesired collisions and usually in some "optimal" fashion. This corresponds to finding a C space trajectory connecting A and B, that does not pass through any collision zone and that has the optimal properties. There can be several optimal criteria which include

. time optimal - e.g. minimum time from A to B
. tolerance optimal - no closer than E to any collision zone
. minimum reversals of motions
. minimum information storage.

Before discussing such trajectories several basic properties of a trajectory in C space are given.

Figure 11 presents the machine of Figure 3 with its C space transform and several trajectories between A and B.

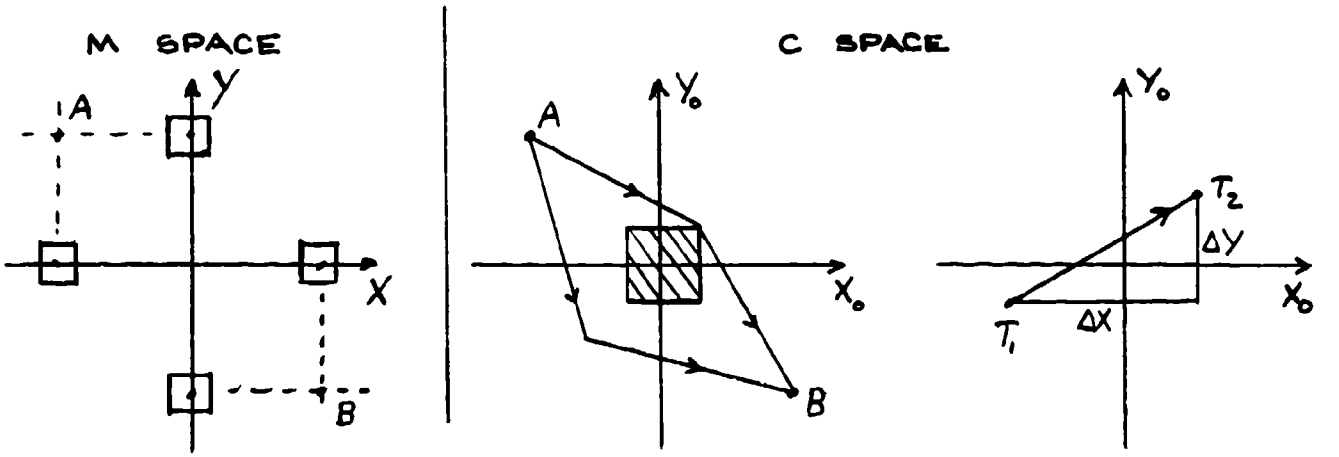


Figure 11. Trajectories in C Space

An essential property of a trajectory is that any point on it represents not only a unique machine configuration but also the instantaneous relative velocities of the degrees of freedom. Time is not implicity nor explicitly evident on a trajectory. Thus a straight segment of a trajectory simply means that at any point on the line the two blocks maintain the same instantaneous relative velocities that is:

$$m = \frac{\Delta X_o}{\Delta Y_o} = \frac{\Delta X_o}{\Delta T} \cdot \frac{\Delta T}{\Delta Y_o} = \frac{V_x}{V_Y}$$

$$V_x = m V_y$$

Thus the blocks may accelerate, decelerate, stop, move at constant velocity and as long as the instantaneous velocities retain the same ratio, the trajectory will describe a straight line in C space. Or inversely, the local slope in the trajectory is determined by the instantaneous ratio of velocities of the degrees of freedom. A straight line trajectory can be a minimal time trajectory but need not be the only one. The shortest distance in C space is not

necessarily the only time optimal policy. The meaning of the path slope and time is complicated by the dynamics of the machine elements.

A simple example of the above statments is shown in Figure 12 which shows two dimensional C space reticulated into a grid in which movement can only be from point to point as shown with the objective of reaching point O in the shortest time from any other point shown.

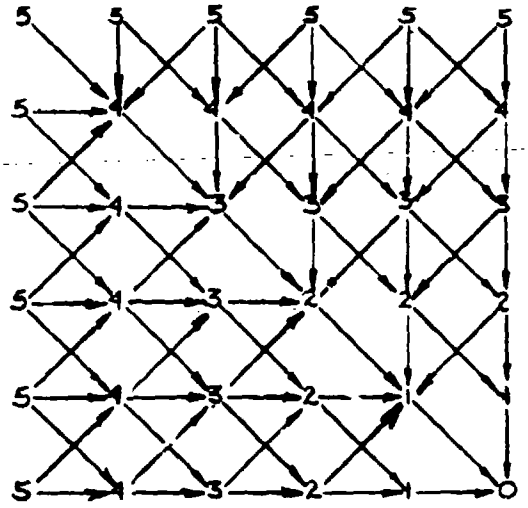


Figure 12. Minimum Time Trajectories in a Reticulated 2 Dimensional C Space

This figure assumes that each degree of freedom can move at the same maximum velocity in C space (by suitable scaling). The points labeled with the same numbers then indicate that they are the same time distance away from O. The figure shows that a 45° line connecting two points is the "fastest" trajectory. For example, once ona 45° line to point O, no movement off it is allowable while retaining minimal time. However, at any point off the diagonal to point O, there can be a number of time optimal trajectories. This reticulated example may be considered to exist on as small a scale in C space as desired.

A more rigorous analysis of trajectories and obstacles is given below.

II. AN EXAMPLE IN SOME DETAIL; TRANSLATIONAL MOTION ONLY

The configuration space, with collision zone, of the machine in Figure 3 is again repeated in Figure 13. Three possible trajectories from an initial configuration, $x_o = -p$, $y_o = -p$, to a final configuration, $x_o = p$, $y_o = p$, are shown.

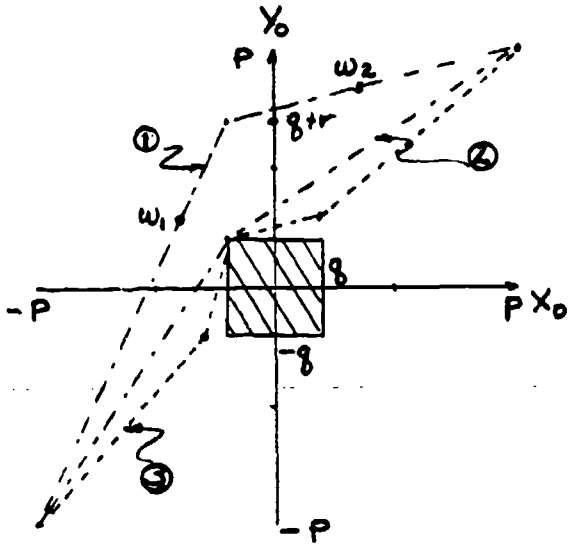


Figure 13. Trajectories in a Two-Dimensional Configuration Space

The slope, m, of a trajectory indicates the relative velocity of the two squares. Thus if $m > 1$ the "y- square" is moving more rapidly, if $m < 1$ the reverse is true. In all three slopes shown the y- square first moves more rapidly, then the x-square does.

In considering these trajectories the following two simplifying assumptions will be made.

1. One or the other of the two squares is always moving at the highest allowable speed, s_{max}.

2. Instant acceleration is allowed (thus thecorners in the trajectories).

Given assumption 1 it follows that for

$$m>1 \;,\; V_y = S_{max} \;,\; V_x = S_{max}/m$$

$$m=1 \;,\; V_y = V_x = S_{max}$$

$$m<1 \;,\; V_y = m S_{max} \;,\; V_x = S_{max} \;.$$

Finally, note that trajectories 2 and 3 are "just miss" trajectories while trajectory 1 included a larger safety factor.

A. Minimum Time Trajectories

Since one or both of the squares are always moving at top speed, and the value of m determines which one, the following rules are used to compute the time required for a trajectory:

1. determine those sections of the trajectory for which $m > 1$, $m = 1$ and $m < 1$.

2. Determine the total length of the trajectory's projection on the y_o

axis where $m > 1$, on the x_o axis where $m < 1$ and on one or the other where $m = 1$.

3. Divide this total projection length by s_{max} to obtain the trajectory time.

Consider the trajectories of Figure 13 using the rules just listed, the following times can be computed.

a. Trajectory 1.

b. Trajectory 2.

c. Trajectory 3.

We note that the two "just miss" trajectories take equal time, though in configuration space they are not the same length. We now state a general criterion for the shortest time trajectory between two points in configuration space that can be connected by a straight line not intersecting a collision zone (i.e. a pair of "unobstructed" points).

Criterion -

Given two "unobstructed" points, $u_1 = (x_1, y_1)$ and $u_2 = (x_2, y_2)$, and assuming that

$$\left|\frac{y_2 - y_1}{x_2 - x_1}\right| > 1 ,$$

then <u>any</u> trajectory from u_1 to u_2 such that $|m| \geq 1$ over the complete trajectory is a minimum time trajectory between u_1 and u_2. For $\left|\frac{y_2 - y_1}{x_2 - x_1}\right| < 1$ the condition is $|m| \leq 1$. If this criterion is not met the trajectory is not minimum time. (In physical terms this means that one specific component goes at top speed the entire trajectory.)

We now introduce an optimality principle - a minimum time trajectory consists entirely of minimum time subtrajectories between unobstructed points. This is a necessary but not sufficient condition as we shall see.

Next, we investigate the trajectories of Figure 13. Trajectories 2 and 3 meet the criterion over their complete lengths, and are in fact minimum time trajectories. Thus minimum time trajectories are not unique; an infinite number may exist. The underlined word <u>any</u> in the criterion ensures this. Trajectory 1, as determined, takes longer. Note points w_1 and w_2. The trajectory between these points fails the criterion and ensures that trajectory 1 is not minimal time.

Before preceding we should note that if the maximum speed of the two squares differed, either the criterion would have to be adjusted or the figure scaled appropriately. Scaling according to maximum speeds seems to offer the better remedy in terms of ease of interpretation. Moreover, considering Figure 6 where one movement is rotational, the relative scaling is arbitrary to begin with. Appropriate scaling means that an equal distance on the figure, either vertically or horizontally, should require an equal time at the maximum speeds. The assumption of Figure 6 is that the propeller can turn 90^o in the same time the square moves five units, both going at top speed.

B. Minimum Information Storage Trajectories

We now consider the amount of information it takes to specify a trajectory. If the machine control is to be computer driven then this determines the number of instructions required. If information storage or the time required to input this storage are important constraints on operation it may become desirable to minimize (to some degree) the required instructions.

If a trajectory consists of straight line segments then the breakpoints must be given and the required slopes can then be determined. The initial and final configuration points are required, but are common to all trajectories. Comparing the trajectories of Figure 12 we see that trajectories 1 and 2 require less information than trajectory 3, with two line segments to the latter's four. Could we do better perhaps with a smooth curve from the initial configuration point to the final configuration point. The minimum we we would require to avoid the collision zone is a quadratic curve, specified by three points, the initial and final points being two. This is equivalent, in information, to the two line segment trajectory and is therefore also a minimum information trajectory. We might note that the quadratic passing through the corner point, $x_o = -q$, $y_o = q$, is also a minimum time trajectory as well.

III. An Example with Rotational and Translational Motion

Figure 14 indicates an initial and final configuration for the machine of Figure 6, and indicates three different trajectories. The scales are measured in units of d, according to the scaling concept previously defined. Assuming again some scaled maximum speed, s_{max}, we find that for

Trajectory 1 $T_1 = ((3.7-2.4)d + (3.9-2.1)d + (3.7-2.0)d)/s_{max} = 3.6d/s_{max}$

Trajectory 2 $T_2 = ((3.9-1.6)d + (0.7-0.0)d + (4.0-3.8)d)/s_{max} = 3.2d/s_{max}$

Trajector $T_3 = ((2.4-1.3)d + (4.4-2.1)d)/s_{max} = 3.4d/s_{max}$

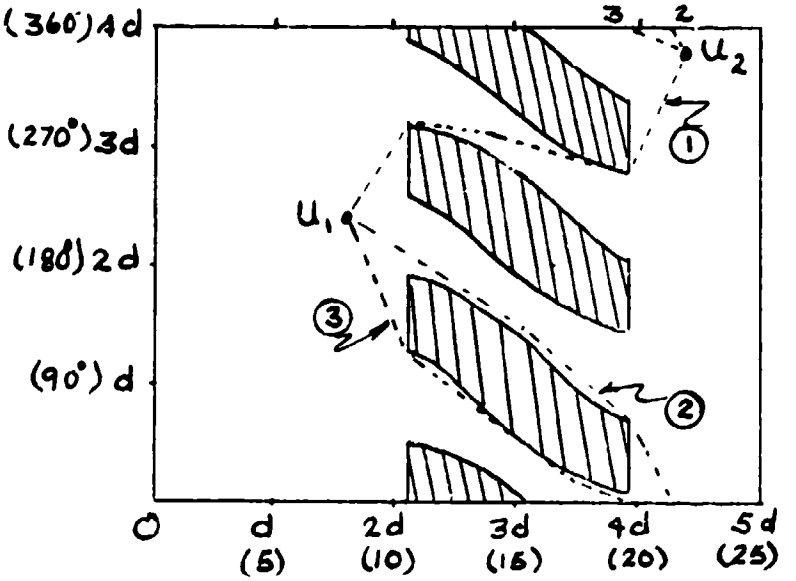


Figure 14. Trajectories in a Two Dimensional Configuration Space: Machine with Rotational Motion

Trajectory 2 is the shortest, in spite of the fact that trajectory 1 requires less overall movement of the propeller. The long segment of trajectories 2 and 3 at a -45° angle correspond to both moving parts moving at near top speed simultaneously, and make these faster than trajectory 1.

All three trajectories meet the optimality condition noted earlier. Trajectories 1 and 3 correspond, evidently, to local minimums, whereas trajectory 2 is a global minimum, though again, not unique.

Figure 15 illustrates another point. The trajectory shown requires the elapsed time of

$$T = ((3.8 - 1.6)d + (4.4 - 3.8)d)/s_{max} = 2.8d/s_{max}$$

and is a global minimum. u_1 and u_2 are not an unobstructed pair of points, yet the trajectory touches no point or edge of the collision zone. While some of the globally optimal trajectories do have such "touch" points clearly such a safety margin is desirable when it can be achieved.

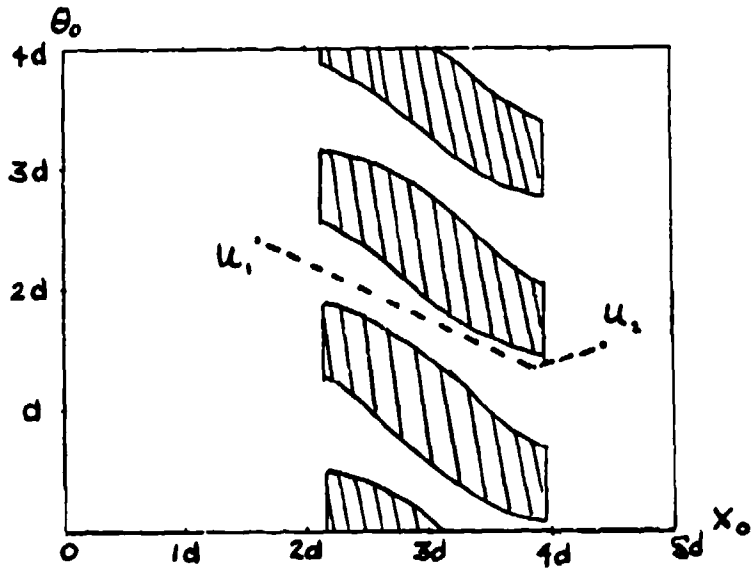


Figure 15. An Optimal Time Trajectory Not Exhibiting "Just Miss" Characteristics

IV. Trajectories in More Than Two Dimensions

If we consider minimum time trajectories in more than two dimensions we once again can assume that at all times at least one component is moving at maximum speed. An optimality principle similar to the two dimensional case would seem to exist. Once again, the trajectory between two unobstructed points should include one component going at top speed the entire way, with the other components traveling at appropriately reduced velocities. Multiple dimension problems are not direct extensions of the two dimensional case, however, due to the pairwise nature of collisions. Thus while two squares on orthogonal tracks generate, as shown, a square collision zone, three cubes on orthogonal tracks do not generate a cubic collision zone, but rather a more complicated figure.

Each machine degree of freedom, as read from a position feedback potentiometer for instance, can form an orthogonal basis for C space. In the event that one degree of freedom is not independent of another over part of their ranges (as, for example, if the two slider blocks in Figure 3 were connected by a two bar linkage) then the restriction is not on the C space itself but on the allowable trajectories in C space. Again such restrictions may be viewed as "collision" zones in which trajectories lying on (not in) the surface of the zone are acceptable.

TOLERANCING AND Ad Hoc TRAJECTORY CHECKING

Although C-space with its collision zones provides the conceptual framework for planning optimized collision free trajectories for the transformed machine, actual determination of such a path in hyperdimensional space is not always evident. The methods described by Perez[6] of slicing a three dimensional polyhedral zone into a series of parallel 2 dimensional polyhedral zones and then proceeding from vertex-to-vertex through the stack is a useful approximation method of procedure.

Another approach that has been tried[2,a] is to plan a type of "straight-line" trajectory from known initial to final configuration and then with computer model mimic technique to move the model into successive positions along the trajectory, checking for model collision at each point. Should no impending collisions be signaled, the trajectory is declared "safe" and movement may be implemented. The 3 basic questions with this method are:

1. Are configurations not checked also safe?
2. How close is this trajectory to a collision zone?
3. If the checked trajectory is signaled as unsafe, how is another trajectory chosen for checking?

The first two questions will be elucidated by examples from actual experience with computer control of a radiotheratron machine.

The third question will not be treated here since it is a fundamental problem with this approach and usually employs machine specific ad hoc strategies.

Basically, prechecking every point (machine configuration) on a selected trajectory is obviously impossible. However, even checking a finite but large number of configurations may be practically impossible if time is critical. For any trajectory checked at a finite number of points and signaled "safe" (i.e. none of the checked configurations are colliding ones), there is a definite probability that a segment of the trajectory passes through or touches a collision zone. Furthermore this probability increases as the number of checked points decreases, for example, in an effort to save computer time. An example of this relationship is shown in Figure 16 which plots the number of prechecked points against the probability of a trajectory signaled "safe" actually being totally outside of a collision zone. The collision possibility being checked is that between the large sourcehead of the machine and a half cylinder representing a patient on the treatment stretcher. In addition to the detection probability, a "severity" of penetration of the sourcehead inside the cylinder (patient) is also plotted.

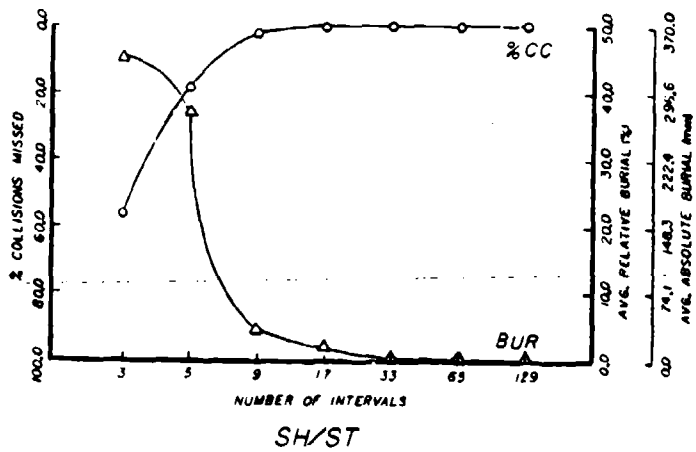


Figure 16. Predetection Probability of a Collision and the Collision Severity as Depending on the Number of Points Checked

As intuitively obvious, as the number of equally spaced points checked on a trajectory increase, the chance of falsely signaling it as safe and the severity of the undetected collision decrease. The main point, however, is that for any finite number of points prechecked, there is always a probability of an undetected collision and the most that can be hoped is that the undetected collision is a "grazing" one in which the trajectory touches but does not penetrate a collision zone. The significance of a grazing collision depends of course on the application.

The second problem of "how safe is safe" refers to the distance between a trajectory and the nearest collision zone. During movement control, the machine may stray from the trajectory for many reasons. If the trajectory is not sufficiently far from a zone, the new trajectory may enter it.

A commonly suggested attempt to solve both problems involves putting a tolerance around each collision zone in C space or, equivalently, artificially enlarging one or both machine elements in a collision pair. This solution has two drawbacks. The first is that while increasing the probability of signaling a true colliding trajectory such enlargement results in the probability of false alarms in which a safe trajectory is rejected. The second drawback is that zone enlargement reduces the free space in which trajectories must lie increasing the difficulty of their selection.

The relationship between original and inflated zones and the probability of missing a true collision or falsely signaling a safe trajectory as unsafe is shown in Figure 17.

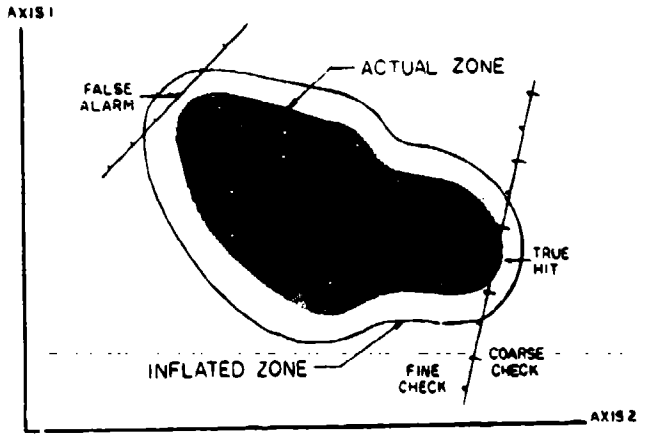


Figure 17. Zone Tolerancing and False Alarm Introductions

Figure 18 presents the Hit and False Alarm detection curves for collisions between the radiotheratron sourcehead and half cylinder (patient) for various percentages of sourcehead enlargements. The number of points prechecked on the straight line trajectory are plotted parametrically.

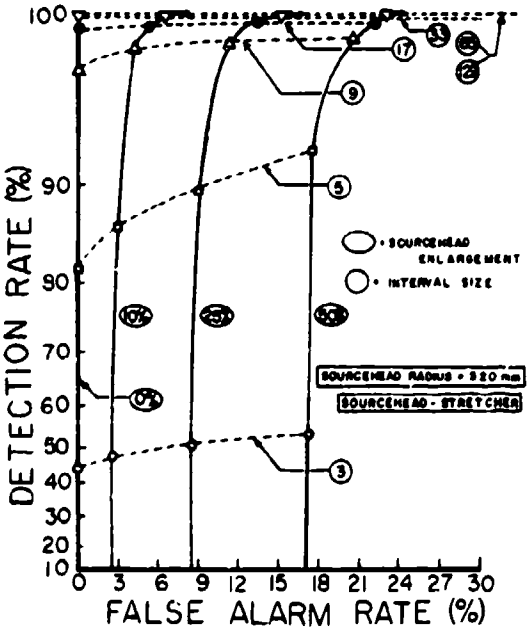


Figure 18. Hit and False Alarm Detections of Unsafe Trajectories for Various Amounts of Machine Part Enlargement

Again, the particulars in this figure may be expected to vary depending upon the machine specifics. However, the general nature will remain.

The significance of false alarms is that an otherwise safe trajectory will be rejected and a new one must be postulated and checked which unnecessarily adds to the time involved.

Thus although tolerancing is a reasonable suggestion, the penalties attached can be significant.

DISCUSSION AND CONCLUSIONS

A general transform approach to collision free computer control of a multidegree of freedom machine has been briefly presented together with several of its implications and problems. Our experience with this approach has resulted in progressive understanding of the problem as well as uncovery of related issues. This approach appears to lead to continual insights and connections with other seemingly unrelated areas. At present we are investigating dynamical aspects of machine movement and their trajectory implications.

The major drawbacks with the transform approach are the difficulty of visualization on which much insight depends and certain computational and storage problems related to collision zone determination and storage and non-convex and/or non-polyhedral zones.

Future work is directed to rapid visual display of zones, trajectories, machines, etc. on a CAD/CAM system, polyhedral approximations to machine elements and/or collision zones and optimal trajectory synthesis and study of the reverse transformation problem as a possible aid to rational design of machines with reduced collision proneness.

The utility of CAD/CAM design of machines may also lie in the fact that such computer aided design also implies that all relevant machine geometry thereby exists in computer storage. Thus computation of collision zones directly from this information would seem highly feasible.

This paper attempted to present some of the basic concepts and considerations of configuration space, collision zones and trajectories for machine control - primarily as they were developed for computer control of a radiotheratron machine. Further discussion of this approach can be found in the references.

REFERENCES

1. Kreifeldt, J., Neurath, P., Seymour, R., Hsieh, D., "Collision Avoidance Programming for a Radiotherapy Machine". AAPM, Chicago, Ill., Dec. 1972.

2. Kreifeldt, J., Hsieh, D., "Collision Avoidance Programming for Computer Delivery of Radiotherapy Treatment", Proc. 2nd Annual New England Bioengineering Conference, Worcester, MA., March 29-30, 1974.

3. Kreifeldt, J., Kiel, R., Richichi, F., "Fundamental Aspects of Computerized Collision Detection for Radiotherapy Machines in the Treatment of Cancer". Proc. 3rd Annual New England Bioengineering Conference. Tufts University, Medford, MA. May 9-10, 1975.

4. Kreifeldt, J., Neurath, P. W., "Improved Delivery of Radiotherapy Treatment Through Computer Automation". IERE Conference Proceedings No. 34. IERE 1976

5. Hsieh, D., Collision Detection and Avoidance Programs for a Theratron 80 Cobalt 60 Radiotherapy Machine. Unpublished MS Thesis. Dept. of Engineering Design, Tufts University, Medford, MA. Oct. 1973.

6. Lozano-Perez, T. "Automatic Planning of Manipulator Transfer Movements." IEEE Transactions on Systems, Man, and Cybernetics. Vol. SMC-11, No. 10, Oct. 1981.

7. Kreifeldt, J. Computer Control of Radiation Therapy Cancer Treatment, Final Report to NIH-NCI. Grant CA 14960. March, 1978.

AD P001639

CROSS MODALITY SCALING: MASS AND MASS MOMENT OF INERTIA

Margaret M. Clarke
Oak Ridge Associated Universities*
Oak Ridge, Tennessee 37830

Stephen Handel
University of Tennessee
Knoxville, Tennessee 39716

John G. Kreifeldt
Tufts University
Medford, Massachusetts 02155

Previous research has shown that we use perceptual inputs from sensations of both mass and mass moment of inertia in handling and manually controlling objects in our environment (1). Mass, the resistance to linear acceleration, and mass moment of inertia, the resistance to angular acceleration, are fundamental properties of objects in both a physical and perceptual sense. The relative importance of mass moment of inertia as a perceptual input has been largely overlooked. Experiments are being conducted ~~in three stages~~ to study the perceptual equation of these two properties.

The first stage involves the perceptual scaling of mass moment of inertia, covering the range typical of simple manipulative tasks (0.15 to 1.15 inch pounds sec sq). pounds sec^2). Preliminary experiments have been completed and are discussed below. The second stage will involve perceptual scaling of mass or weight, covering the range most often found in tools used for simple manual tasks (4 oz. to 10 lbs.). The third stage will involve asking subjects to equate the perceived mass moment of inertia of one set of stimuli with the perceived mass (weight) of a second set.

*Oak Ridge Associated Universities performs complementary work for the U.S. Department of Energy under contract No. DE-AC05-960R00033.

APPARATUS

Figure 1 illustrates the apparatus used. A 6" plexiglass T handle was fixed to a plexiglass shaft 13" long and 1" in diameter. The shaft rested on two ball bearings mounted 6" apart. A 24" length of plexiglass was glued at the end of the shaft, with 12 holes drilled symmetrically at different

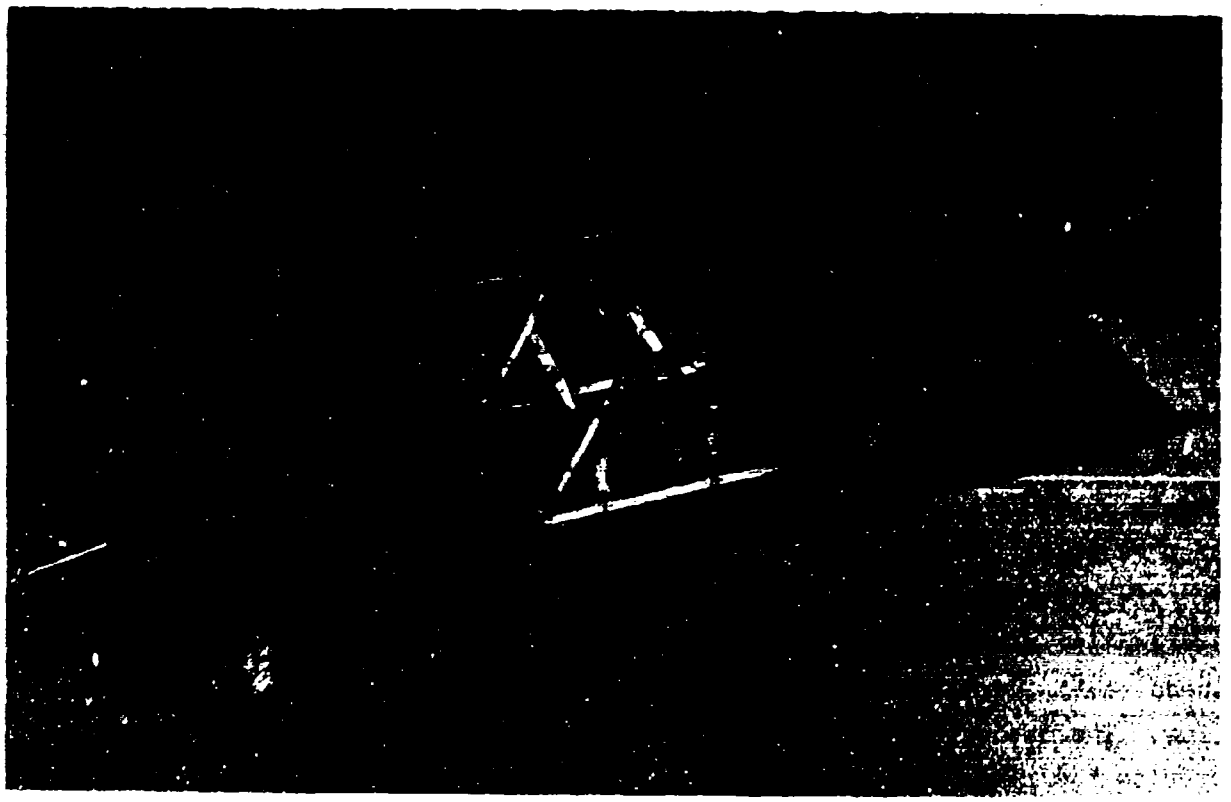


Figure 1. Mass Moment of Inertia Apparatus

positions along its length. The hole positions were selected so that the mass moment of inertia increased by equal ratios when a pair of 2 lb. weights were moved from 2 holes equidistant from the center point of the 24" length to the adjacent holes. The mass moment of inertias were thus linearly spaced on a logrithmic scale. Six different mass moments of inertia, one for each symmetrical pair of holes, could be generated.

SUBJECTS

Subjects were 17 volunteers, employees or family members of employees, of Oak Ridge Associated Universities.

EXPERIMENTAL DESIGN

Each of the 6 mass moments of inertia was presented to the subjects once in each of 3 trials, varying the order of presentation from 1 to 6 randomly within each trial. Subjects were divided into 3 groups, each group being given different instructions.

Group 1 (3 Subjects)

The weights were positioned at the third hole from the center. Subjects were instructed to "move the T bar with your preferred hand from side to side and concentrate on the feeling of 'top heaviness' or 'wanting to keep on going' when you reverse direction. Concentrate on this feeling and give it a number, any number you wish." When the subject had assigned a number, instructions continued: "Now, I'm going to change that feeling of top heaviness and I want you to concentrate on the new feeling and give it a number. If the feeling is twice as large double the number", etc.

Group 2 (4 Subjects)

These subjects were not asked to give an initial reference point. The subjects were merely asked to assign numbers to reflect the magnitude of the sensation.

Group 3 (10 Subjects)

These subjects were given the same instructions as Group 1 except they were told to assign the number "100" to the first stimulus position (weights at position 3).

DATA ANALYSIS

Starting at the center holes, the mass moment of inertia for each weight position was calculated in inch lbs sec^2 as follows: 0.15, 0.20, 0.29, 0.43, 0.70, 1.15. The geometric mean of each subjects' three judgements of "top heaviness" was then calculated for each of the 6 mass moments of inertia. Other research on perception (2) indicates that the sensation magnitude grows as a power function of the stimulus magnitude. On this basis the best fit power function relating perceived "top heaviness" to mass moment of inertia was calculated. The criterion was minimal sums of squared deviations.

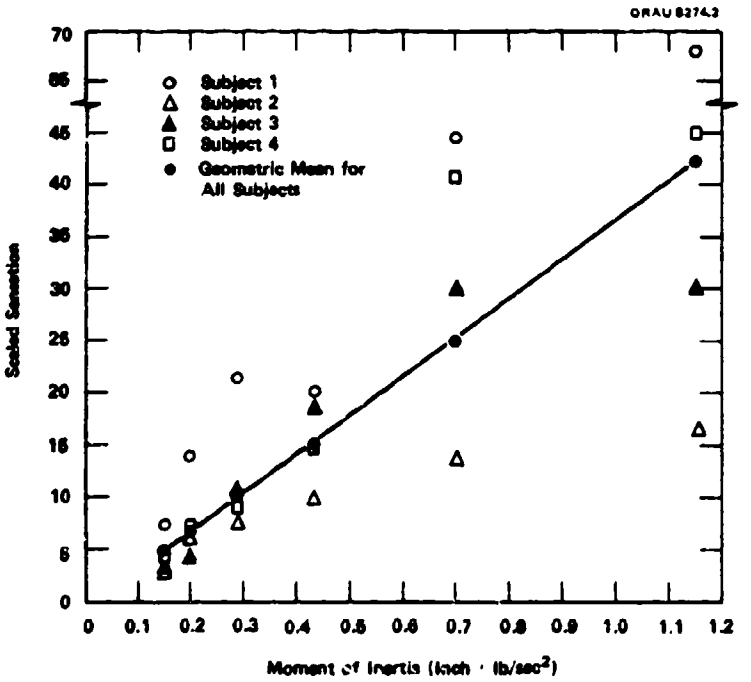


Figure 2. Scaled Sensation and Corresponding Mass Moment of Inertia for four Subjects

Data for Group 2 have been extensively analyzed and are shown in Figure 2 together with their best fit line. The equation for this line averaged across subjects is

$$S = 36.31\ I^{1.06}$$

where S = perceived "top heaviness"
I = mass moment of inertia

There was surprisingly little variation among subjects. The exponent for each subject was 1.00, 0.76, 1.2, 1.22.

DISCUSSION

Under these experimental conditions the relationship between mass moment of inertia and sensation was approximately linear. The power function 1.06 fits within the range of exponents found in scaling experiments for other senses: length 1.0, pain 3.45, loudness 0.67, heaviness 1.45, brightness 0.33 to 1.0, angular accelertion 1.4, cold 1.0, and warmth 1.6 (2). The sensation of mass moment of inertia may be perceived to increase less rapidly than for heaviness, tentatively suggesting that increasing the heaviness or mass of a tool associated with a manually controlled operation may be more noticeable than increasing the mass moment of inertia.

These results will be useful in designing handling systems where trade-off decisions might have to be made concerning the relative effects on perception and task performance of these two properties and in designing further studies in which these two properties may be studied simultaneously.

Future plans include the following:

1. Complete analysis of all present data.
2. Perceptual scaling of mass and of mass moment of inertia in larger reach envelopes.
3. Cross-modality scaling where subjects will be asked to equate the perceived mass moment of inertia of one set of stimuli with the perceived mass of a second set.

REFERENCES

1. Kreifeldt, John G. and Chuang, Ming-Chuen, "Moment of Inertia: Psychophysical Study of an Overlooked Sensation," Science, 1979, vol. 206, pp. 588-590.

2. Stevens, S.S., "Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects," John Wiley, New York, 1975.

PROMAN: Program for Management and Analysis of Manual Control Experimental Data

Ramal Muralidharan and William H. Levison
Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, MA 02238

ABSTRACT

In this paper we describe some software tools developed recently for the management and analysis of data collected while conducting manual control experiments. As implemented, PROMAN consists of four major stand-alone modules which perform the following functions:

EXPINT: Interface with data collected from the experiment

DBASE: Data Base Organization and Management

STAT2: Perform Statistical Analysis

MODINT: Interface with model matching/analysis software

Of these, the DBASE and STAT2 are general purpose modules. The EXPINT interface will be tailored to the format in which experimental data are collected. The MODINT interface will be tailored to the particular model matching and analysis exercises desired on the experimental data.

The use of PROMAN will be illustrated with reference to a specific manual control experiment. It appears that PROMAN can serve as an extremely useful tool for manual control experiment data base management and analysis.

AD P001640

DEXTEROUS MANIPULATOR LABORATORY PROGRAM

Roy E. Olsen
Grumman Aerospace Corp.
Bethpage, N.Y. 11714

ABSTRACT

This paper describes a program currently underway to develop telepresence systems for the support of space missions such as satellite maintenance and repair. A force-reflecting, dexterous manipulator system is emphasized as the basis for the development of integrated telepresence systems. Alternate controllers, control implementations, modified end effectors, and candidate vision systems will be evaluated to resolve human-machine design issues.

INTRODUCTION

A major consideration being addressed as we enter the Space Shuttle era concerns the extent of extra-vehicular activity (EVA) to perform tasks in space versus remotely operated and automated systems. Mission analysis studies have identified many complex tasks which will be required in the areas of satellite services and space construction (Ref. 1). There are numerous reasons for strongly favoring remotely operated and automated systems for certain applications. These reasons include removal of the astronaut from the vicinity of hazardous operations (such as transferring propellant) and reduction of the non-productive operational time associated with EVA.

Remote human presence, i.e., "telepresence", is envisioned as the needed capability for many tasks because of their low frequency of occurrence and/or high level of difficulty. Telepresence systems provide the capability to perform difficult operations with human versatility by providing "natural" control interfaces, vision, and mechanical arms to the operator.

The evolution of space telepresence systems from the present conceptual stage to operational flight status will require development along the lines indicated in Fig. 1, starting with ground-based laboratory tests and simulation and progressing through flight demonstration programs leading to operational systems. The latter will include orbiter-attached configurations, space station support versions, and both manned and unmanned free-flying applications.

Initial studies make it clear that current technology in the fields of nuclear energy, underwater exploration, and industrial robotics are not sufficient to meet the needs of space telepresence. Technology and application issues were identified which require resolution before such systems can be fully developed. Issues such as the type of controller needed, control mode, force reflection applicability, working volume, time delay effects, number of arms, and degrees of freedom must be addressed before operational systems can be defined.

The key to developing space telepresence technology is to start with manned dexterous manipulators and evolve the control implementations, vision

systems, and end effectors before embarking on automated systems. Toward this end, a Manipulator Technology Laboratory was established to investigate and resolve telepresence issues. An electromechanical master/slave dexterous manipulator system was built based on design requirements derived from analysis of future space missions. This system has six active degrees of freedom with a manual seventh, bilateral force reflection, and automatic counterbalancing. The 2-meter-long slave arm is more than three times larger than the master, and indexing capability permits operation of the slave arm throughout its 1100 ft^3 working volume while keeping the master handle within a 10-inch cube. The manipulator slave arm is the basic development hardware around which telepresence systems are being built and tested.

LABORATORY TESTING APPROACH

Initial evaluations have emphasized two types of manipulator system testing:

- Type 1 - Quantitative measurements of man-machine performance to establish a baseline with which to compare future improved systems
- Type 2 - Preliminary applications tasks which represent the kind of operations which would be needed for on-orbit missions.

Type 1 testing utilizes a task board (Fig. 2) which consists of elementary operations, such as pushing buttons, for which performance time and accuracy can be measured (Ref. 2). Type 2 testing uses standard electrical and mechanical components in a realistic setting to evaluate performance ability in a more qualitative sense than does Type 1 testing.

TESTS STATUS

An initial Type 1 test series using 12 test subjects has been completed. The subjects were first screened for distance acuity, depth perception, and color discrimination. Only test subjects who had no experience with manipulators were used. After a short briefing on system operation and the tests to be performed, the subjects were directed through a series of 20 trials with a 5-minute break after each five-trial segment. Tests were conducted with different position gains, with and without force feedback, and for both direct and black and white closed-circuit TV viewing.

Typical results for one test subject are presented in Fig. 3 for one test condition. A learning process is observed for each five-trial segment with average task times about equal for each segment but with average error reduced by almost a factor of five from the first to last segment. In both cases, the standard deviation is significantly lower in the later trial segments.

Figure 4 presents a summary of the same test for four different test subjects for the same test conditions and also using TV. The mean times and standard deviations are only for the last five trials, which represent performance after an initial learning period. In three out of four cases, performance degrades (as expected) when using TV. Task time standard

deviation appears to increase for TV operation, but the standard deviation of accuracy generally decreases with TV use. The latter observation is probably due to increased operator proficiency rather than the means of observing the task being performed.

FUTURE PLANS

Future testing is expected to include the following:

- Digital control manipulator operations
- Improved end effectors
- Multiple-camera vision systems including stereo
- Alternate hand controllers
- Improved display capability
- Arm-mounted miniature TV cameras
- Mission task mockup testing
- Supervisory control operation.

Future plans also include integrating the telepresence system with a large-amplitude motion base to simulate various carrier systems and operational interfaces.

REFERENCES

1. Satellite Services System Analysis Study, NASA Contract NAS 9-16120, Grumman Aerospace Corporation, August 1981.

2. Study of Modeling and Evaluation of Remote Manipulator Tasks with Force Feedback, John W. Hill, SRI International, March 1979.

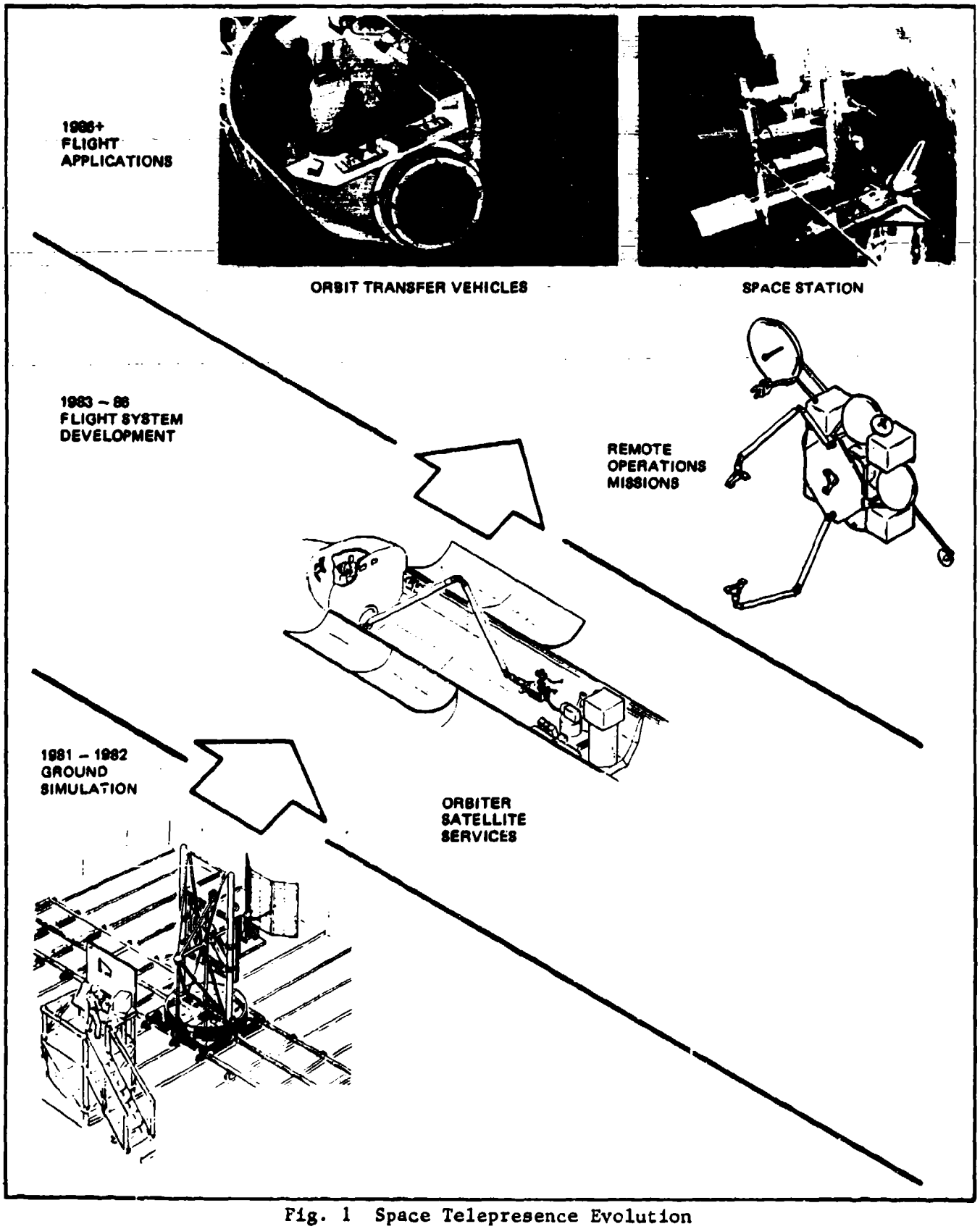


Fig. 1 Space Telepresence Evolution

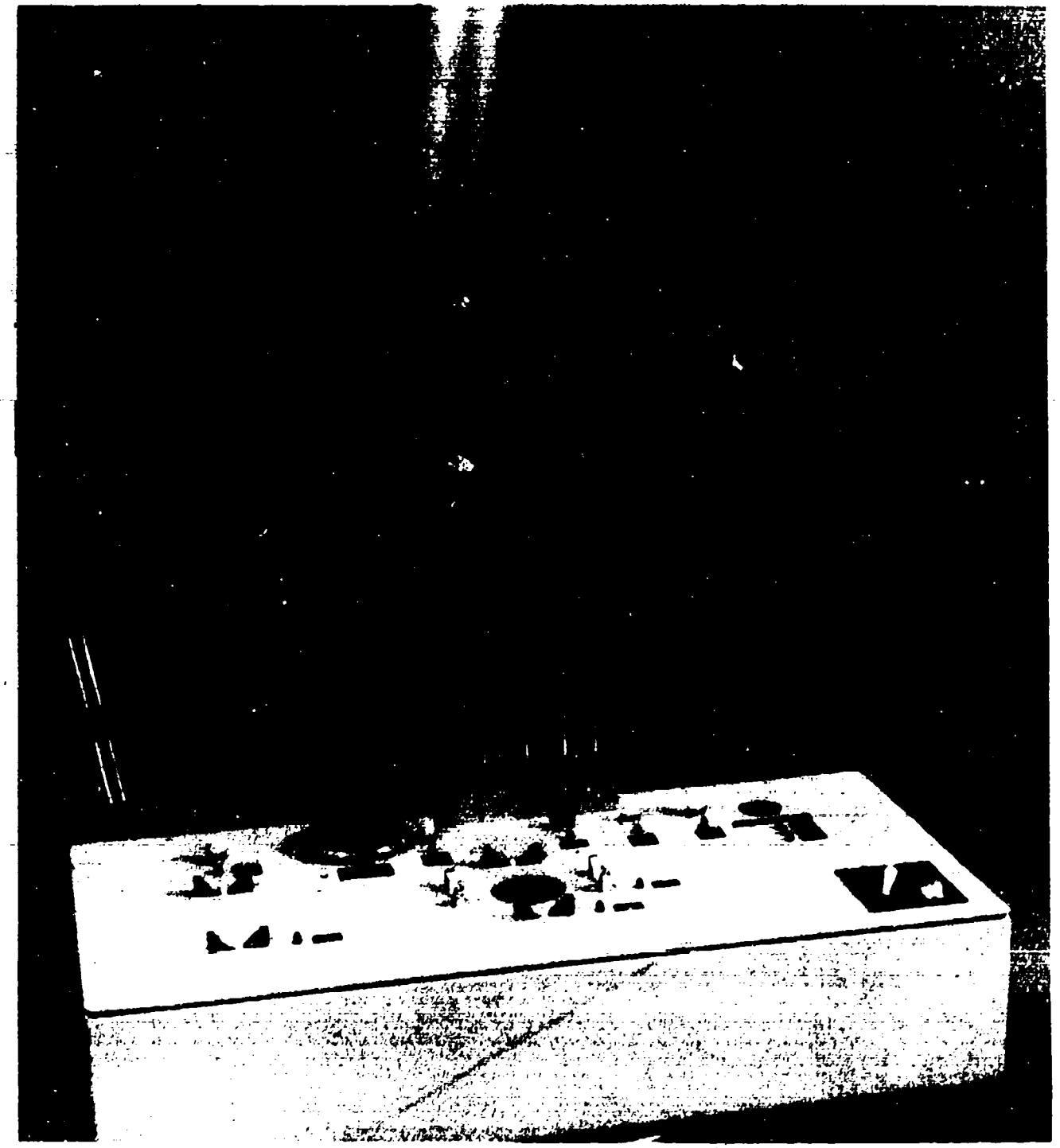


Fig. 2 Manipulator Task Board

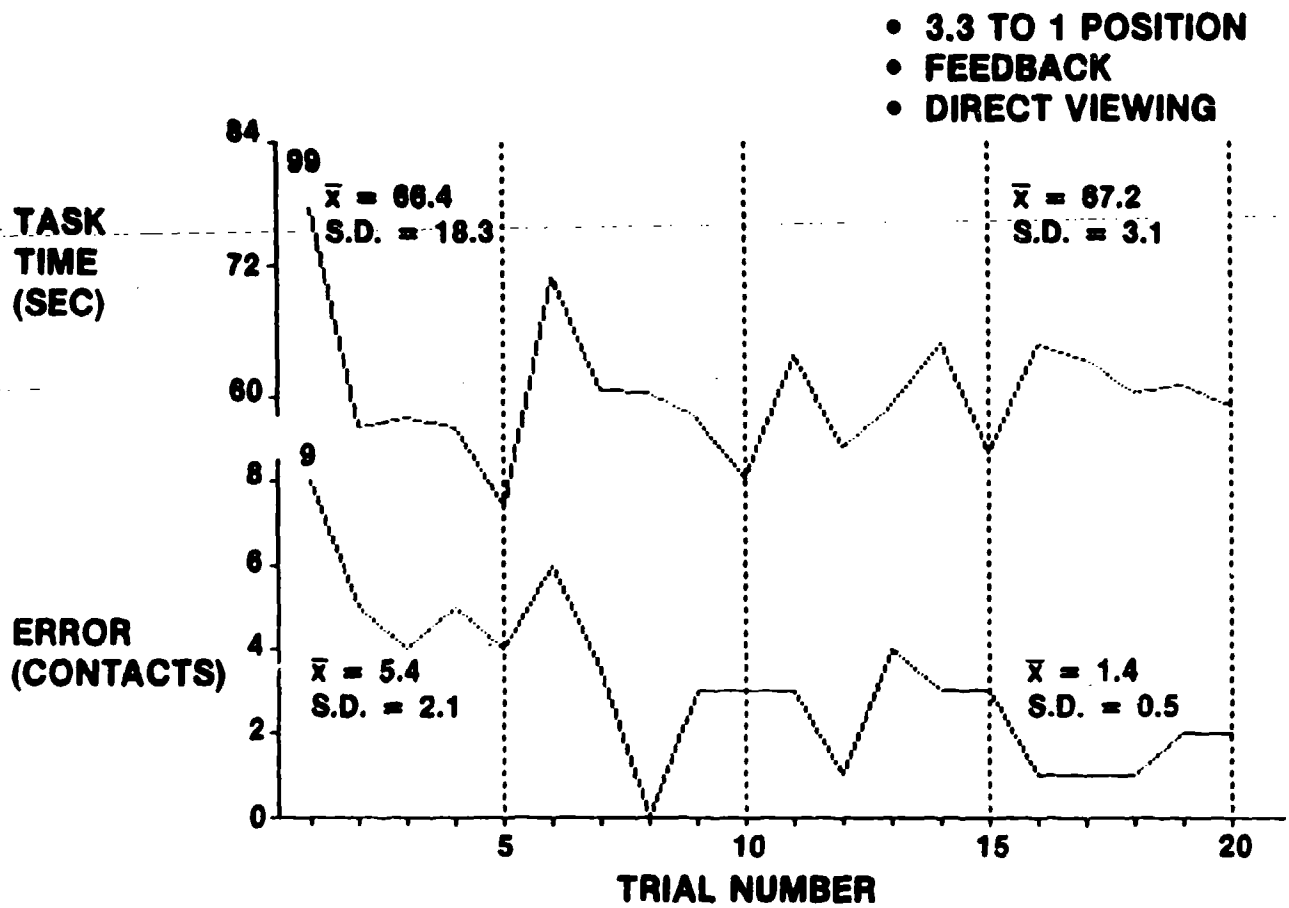


Fig. 3 Typical Test Results

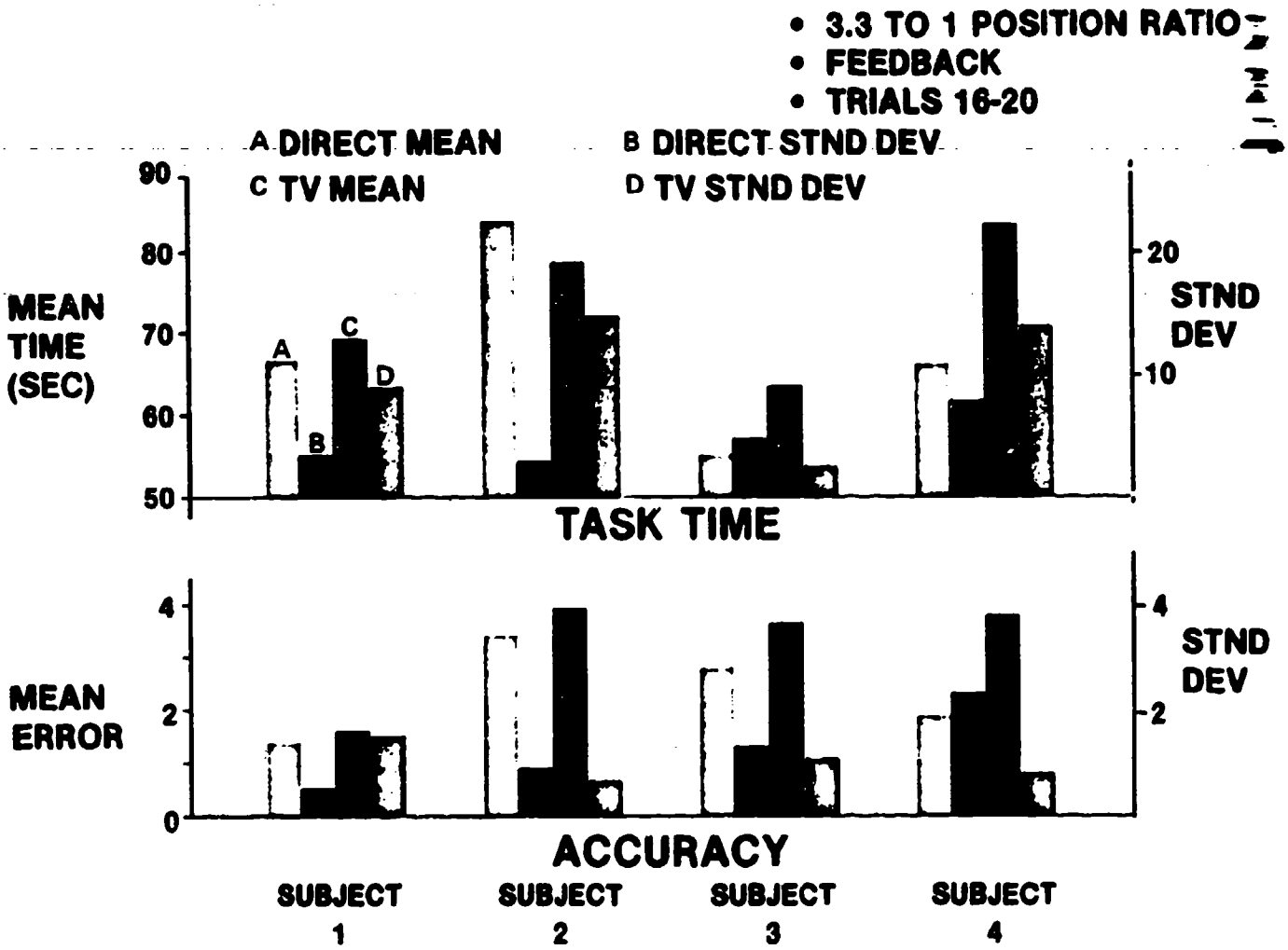


Fig. 4 Task Performance Comparisons

AD P001641

Stick Controller Design For Lateral Acceleration Environments

D. W. Repperger*, W. H. Levison**, V. Skowronski*, B. O'Lear***,
J. W. Frazier*, K. E. Hudson****

* Air Force Aerospace Medical Research Laboratory, Wright Patterson Air Force Base, Ohio 45433

** Bolt, Beranek, and Newman, Cambridge, Massachusetts 02138

*** Raytheon Service Company, Dayton, Ohio

**** Systems Research Laboratory, Dayton, Ohio

SUMMARY

A study is being conducted at the Air Force Aerospace Medical Research Laboratory, Wright Patterson Air Force Base, Ohio to investigate different types of stick controllers that can be used in a complex G environment. The purpose of this study is to develop impedance models to show the biomechanical effects of stick feedthrough as the human body is subjected to low frequency G stress fields. These impedance models will extend previous work developed in the vibration area (reference [1]) to the low frequency vibration fields that commonly occur in modern fighter aircraft (e.g. the AFTI/F-16 with lateral acceleration fields). Another purpose of this study is to investigate how human subjects use displacement and force feedbacks in side arm controllers.

INTRODUCTION

The design of the hand controller in a manual tracking situation may substantially influence the performance characteristics of the closed loop man-machine system. In a previous study in the vibration environment [1], it was found (figure (1)) that the dynamics of the stick controller significantly influenced the performance characteristics of the closed loop system under a variety of vibration disturbances. For example, (figure (1), top plot), the difference between closed loop e_{RMS} tracking scores when averaged over the seven subjects in the experiment was approximately a 25% change in tracking performance. This effect (i.e. the controller's dynamic characteristics) gives rise to a performance decrement which is an important effect to be considered in the experimental design. A further study by W. H. Levison investigated these impedance models developed in the vibration area by computer simulation of the Optimal Control Model over a variety of different stick parameters (reference [2]). An interesting result from this study was the existence of an optimal electrical gain (top plot in figure (2)) which minimized closed loop tracking error variance.

As it has been demonstrated, performance can vary over a wide range for different choices of stick parameters (figure (2)). The next question that is asked is how much further is it possible to design stick parameters to optimize closed loop tracking error or reduce feedthrough effects? From observation of figure (2), it appears that a reduction in stick feedthrough by a factor of about 2 might be realized from the adjustment of the electrical gain characteristics. To study some of these effects, a special type of biomechanical model was developed (reference [3]) for lateral acceleration environments

which occur in the AFTI/F-16. In figure (3), a description of the biomechanical model is given. The forearm of the subject is illustrated and moves side to side in the Y direction. The wrist of the subject is at the point where the forces $F_1(t)$ and $F_2(t)$ act on the stick. The stick is modeled by the mass M_s, the spring K_s, and the damper B_s as indicated in the dotted box. The objective of the study reported in reference [3] was to make the best choices of M_s, B_s, and K_s so as to minimize stick feedthrough. In this manner it is possible to choose stick parameters to develop a minimum feedthrough stick. The minimum feedthrough stick may not be the optimal stick in the sense of best performance (minimum tracking error). A trade off may exist betweeen the extreme of a minimum feedthrough design versus the other extreme design requiring excessive control forces by the pilot. Figure (4) illustrates the results of the computer simulation of this biomechanical model for different choices of the independent variables M_s, K_s, and B_s.

In figure (4), stick feedthrough H (units of $Degrees^2$ - displacement) is a measure of the stick feedthrough obtained via computer simulation of the biomechanical model illustrated in figure (3). On the axis B_s/M_s = constant = 0, it is observed that a minimum feedthrough stick appears as a hybrid combination between a force stick and a displacement stick. Quite interestingly enough, the stick feedthrough is reduced to a level of 10^{-2} at the minimum but achieves levels one or two orders of magnitude higher for both the force stick region and the displacement stick region. This substantial reduction of feedthrough indicates the existence of some minimum feedthrough designs that may have distinct advantages over conventional designs used today. Along the K_s/M_s = 0 axis, one also observes that feedthrough is not sensitive to the parameter B_s/M_s. This, quite interestingly, concurs with results obtained in reference [2] (cf. plot 2 on figure (2)) in which error variance is very insensitive (the plot is very flat) to stick damping over a range of two orders of magnitude. From the early work of W.H. Levison (figure (2)) and the results of the biomechanical simulation (figure (4)), it was hoped to design a study involving human subjects to examine the effect of different stick parameters on closed loop tracking error.

METHOD

Design of the Forcing Function - A zero mean, constant variance, sum of sines, disturbance input tracking task was developed consisting of component frequencies ranging in value from .192 radians/second to 64.005 radians/second. The task was specially designed to investigate various aspects of the man-machine system based on frequency spectrum characteristics of the controller that would effect the human's responses. The tracking task was of duration 165 seconds and the subjects had a 1 minute rest between tracking sessions. All of the subjects tracked nine tasks each day.

Apparatus - An F-4 aircraft seat was situated in a semi-isolated room and a television monitor was located approximately 56 inches from the subject's eyes. The subject was displayed a 1 1/4 inch wide circular pipper and a target (cross) of equal width. The subject would move a side arm controller in the lateral direction and control the pipper on the inside out display. The control signals developed in the loop were sent to a SEL computer for storage on magnetic tapes. At the SEL computer, the system being controlled (1/s dynamics) was simulated on an analog computer diagram.

Subjects - Nine subjects participated (8 male, 1 female) in part I of the stick study. These subjects were active duty Air Force members of the Acceleration Panel and received hazardous duty pay for participation in centrifuge experiments. Their ages ranged from 24 years to 39 years and they had a variety of experience to both G exposures and manual tracking tasks.

Experimental Design Conditions

Three spring gains (denoted as WDS (weak displacement stick), SDS (strong displacement stick), and FS (force stick)) were used to generate the independent variable K_s. The units of spring gain are pounds force/displacement in degrees (figure (5)). Three electrical gains (or display gains) were used to vary the force characteristics versus speed of the target on the screen of the TV monitor. The units of the electrical gain are volts/pound (figure (5)). Each subject was required to track for 3 days of data collection for each experimental condition. On each data day, the subject would have one stick configuration and track 9 runs or trials. The nine trials consisted of the 3 electrical gains presented randomly 3 times each. The high electrical gain was designed to be over sensitive, the mid value of electrical gain was chosen to be optimal, and the low electrical gain was designed to require excessive control inputs. The presentation of the spring gains (K_s) also occurred randomly each data day.

Training Orientation - The subjects tracked until they appeared to show asymptotic performance levels. This level was determined as less than a 5% change in RMS error scores for 3 successive data days. Calculations were also made of the ratio $\hat{\sigma}/\hat{\mu}$ (standard deviation of the 3 replications for one data day divided by the mean e_{RMS} score). If this ratio was of the order of 0.1, it was assumed that training was asymptote.

Results

The results presented herein are somewhat preliminary in that only part I of this experiment was completed at this point in time . Figures 6,7,8 illustrate the results of these data for the nine subjects, the three spring conditions, and the three electrical gain conditions. For each spring condition, (force stick, strong displacement stick, and weak displacement stick), the three electrical gain conditions are illustrated (G=1, G=10, and G=90). The three plots are kept on separate axes because a G=1 plot for the force stick gives rise to a different electrical gain (K_e) than G=1 for the strong displacement stick. What is interesting to note is that the electrical gain G=10 appears to give optimal (lowest e_{RMS} error scores) for both the force stick and the strong displacement stick. This result, however, does not appear to hold for the weak displacement stick for data averaged over the nine subjects.

Since the greatest variation in this experiment is due to biodynamical differences in the subjects (and their strength characteristics), it is necessary to separate out the intra-subject variability (one subject all conditions) from the inter-subject variability (all subjects, one condition). To illustrate the effects of the spring gains of the stick controller on one subject, figure (9) is representative of the performance scores obtained. The learning data in figure (9) represents the three sessions prior to the data collection. An ANOVA was conducted on the performance scores using the electrical gains and different sticks as factors. The results of the ANOVA and a Newman-Keuls Multiple Comparison Test Procedure (alpha=0.1) for all 9 subjects are illustrated in Table I:

Table I - Newman-Keuls Multiple Comparison Procedure (alpha=0.1)

Electrical Gain	G=10	G=90	G=1	
mean	.4535	.4747	.5795	Force Stick
s.d.	.0785	.1094	.1519	
Electrical Gain	**G=10**	**G=1**	**G=90**	
mean	.5668	.5871	.8205	Strong Displacement
s.d.	.1366	.1217	.3315	Stick
Electrical Gain	**G=1**	**G=10**	**G=90**	
mean	.6314	.65223	.9752	Weak Displacement
s.d.	.2895	.2810	.4205	Stick

* Parameters underlined are not significantly different.

The results of the Newman-Keuls Multiple Comparison Test as illustrated in table I indicate for the force stick that there were no significant differences in mean scores due to the change in the electrical gain. For the strong displacement stick and weak displacement stick, however, the mean score for the gain conditon G=90 was shown to be significantly different from the other two gains.

Summary and Conclusions

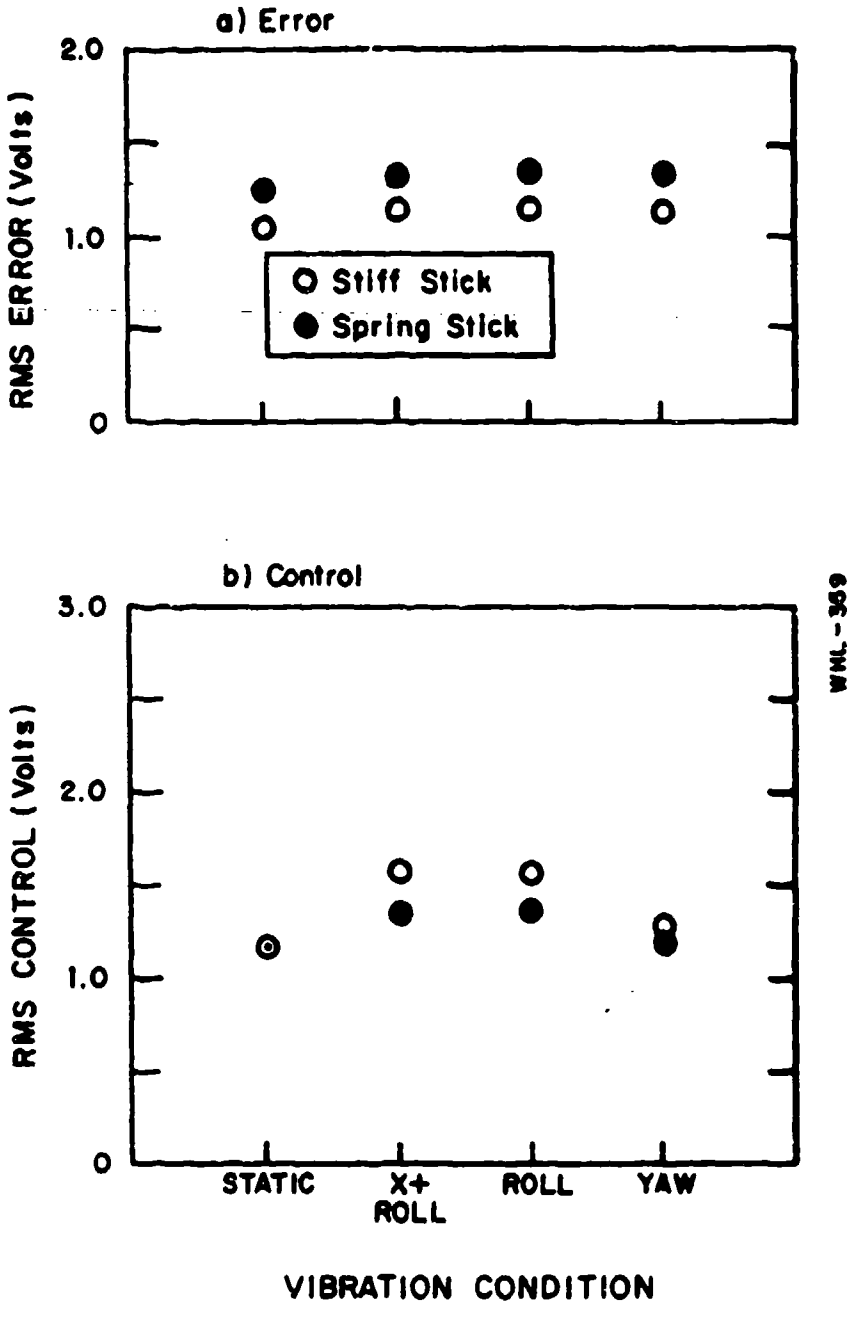
A study was conducted on different types of stick controllers to be used in G environments that occur in modern side slip vehicles. A choice of 3 electrical and 3 spring gains were used with test subjects from the centrifuge panel. From this investigation, which is exclusively static and contains no G disturbances, the stick parameters influenced tracking performance across subjects for two of the three sticks considered. These results show high variation when averaged over all 9 panel members; however, the results appear to show differences when averaged across replications for one subject. This high inter-subject variability reported here may be due to wide va iation in the biodynamic response characteristics of the pool of subjects.

References

[1] Levison, W. H., and P. D. Houck, "Guide For The Design of Control Sticks in Vibration Environment", Feb, 75, AMRL-TR-74-127

[2] Levison, W.H., "Biomechanical Response and Manual Tracking Performance in Sinusoidal, Sum-of-Sines, and Random Vibration Environments", AMRL-TR-75-94.

[3] Repperger, D.W. and W.N. Bianco, "Optimization of Stick Design Under ± Gy Stress Using Biomechanical Modeling and Limitations of Resonance Effects", Proceedings of The 1980 Aerospace Medical Association Meeting, pp. 92-93, May, 1980.



Effect of Vibration on Rms Tracking Performance Scores, Roll-Task

Figure (1) - Force Stick (Stiff Stick) versus Displacement Stick (Spring Stick) in the vibration environment. From Reference [2].

COMPARISON TO VIBRATION RESEARCH - TR-AD -A008 533- "GUIDE FOR THE DESIGN OF CONTROL STICKS IN VIBRATION ENVIRONMENTS"

- William H. Levison, ET. AL. Bolt Beranek and Newman

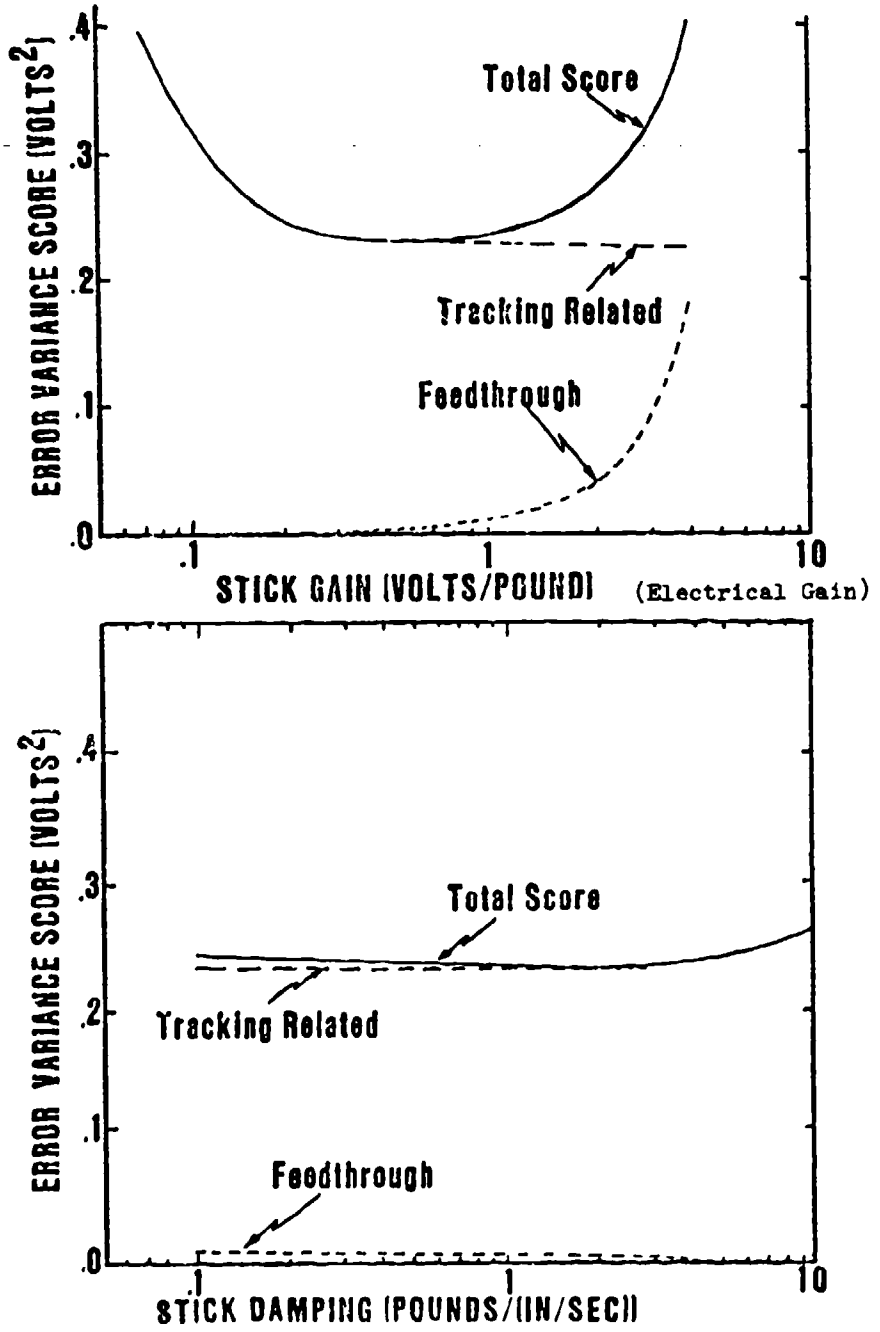


Figure (2) - Simulation of the Optimal Control Model for different electrical gains and spring damping characteristics for impedance models developed under vibration. From reference [1].

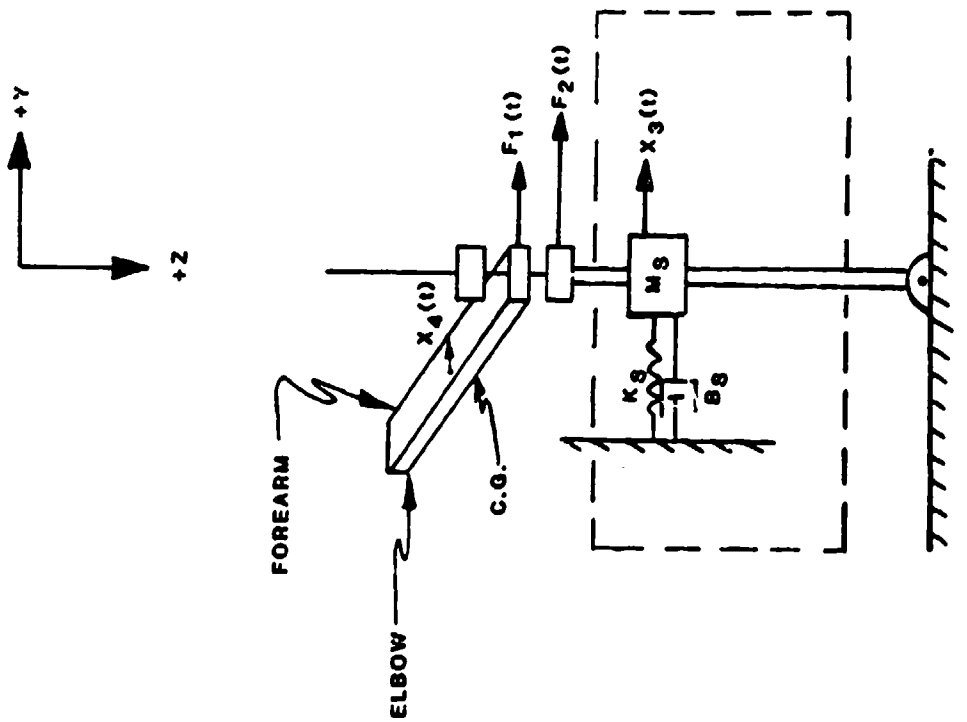


Figure (3)

The biomechanical model developed for lateral acceleration environments

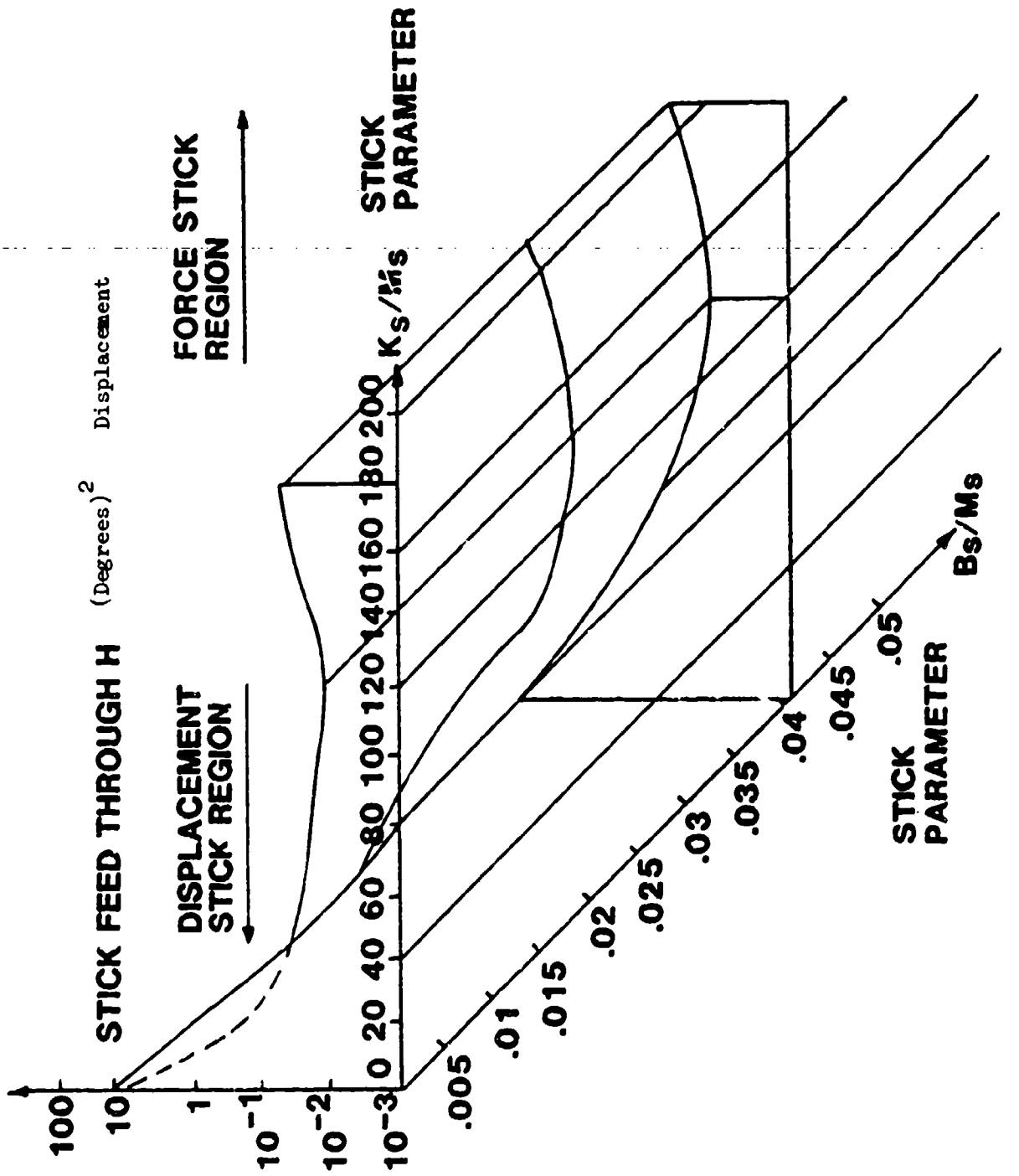
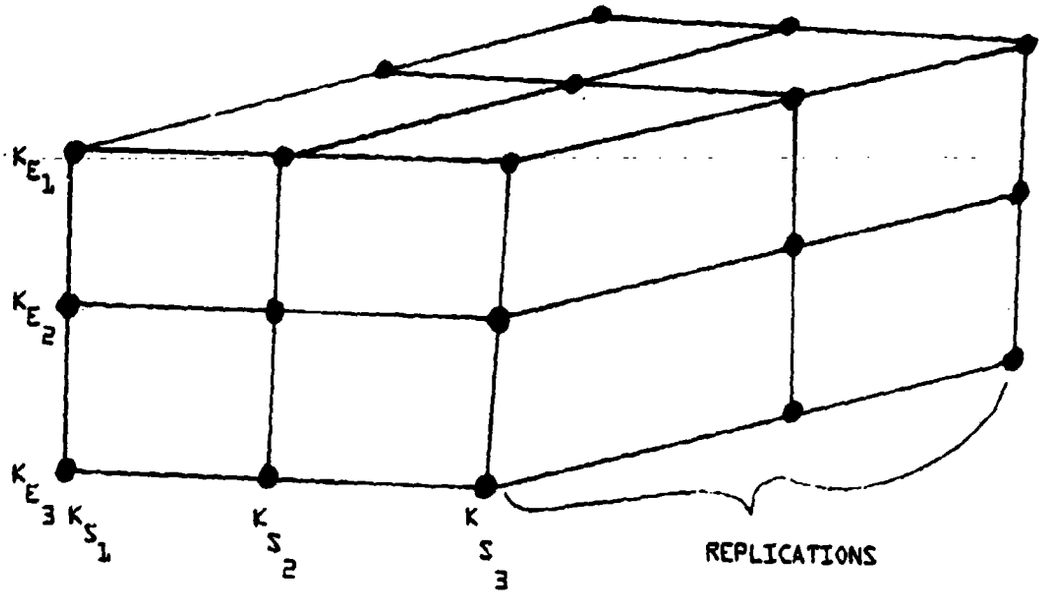


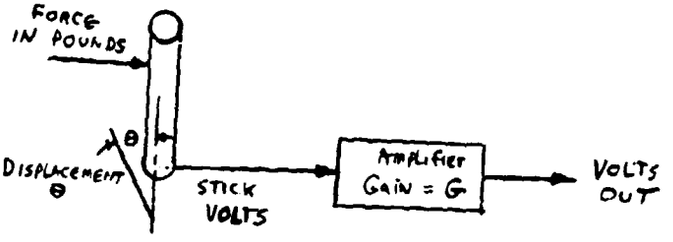
Figure (4) - Stick feedthrough as a function of B_s, M_s, and K_s for a specified G field. From reference [3] .

EXPERIMENTAL DESIGN CONDITIONS



ELECTRICAL GAIN K_E =

Force Stick	Displacement Stick
Volts / Pound	Volts / Displacement



SPRING GAIN

$$K_S = \frac{\text{POUNDS}}{\text{DISPLACEMENT (DEGREES)}}$$

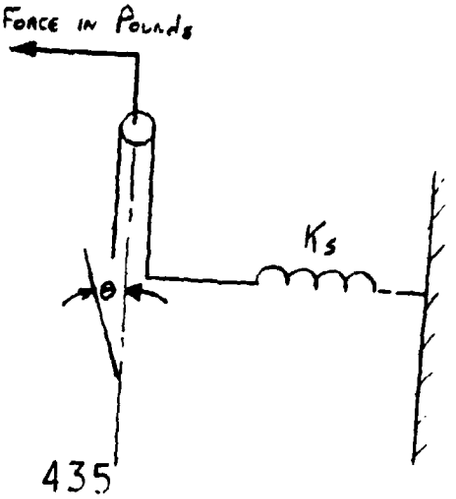


Figure (5)

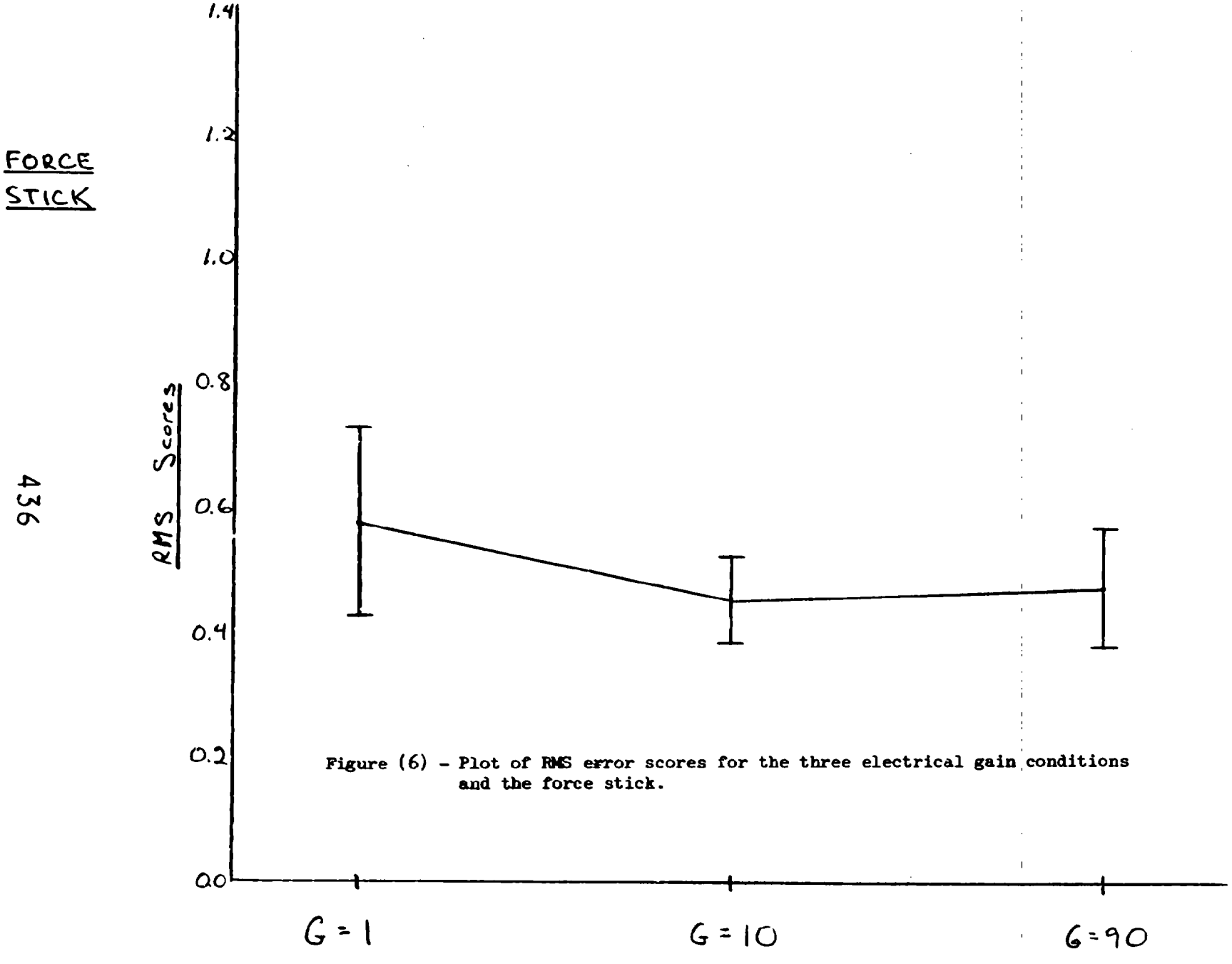


Figure (6) - Plot of RMS error scores for the three electrical gain conditions and the force stick.

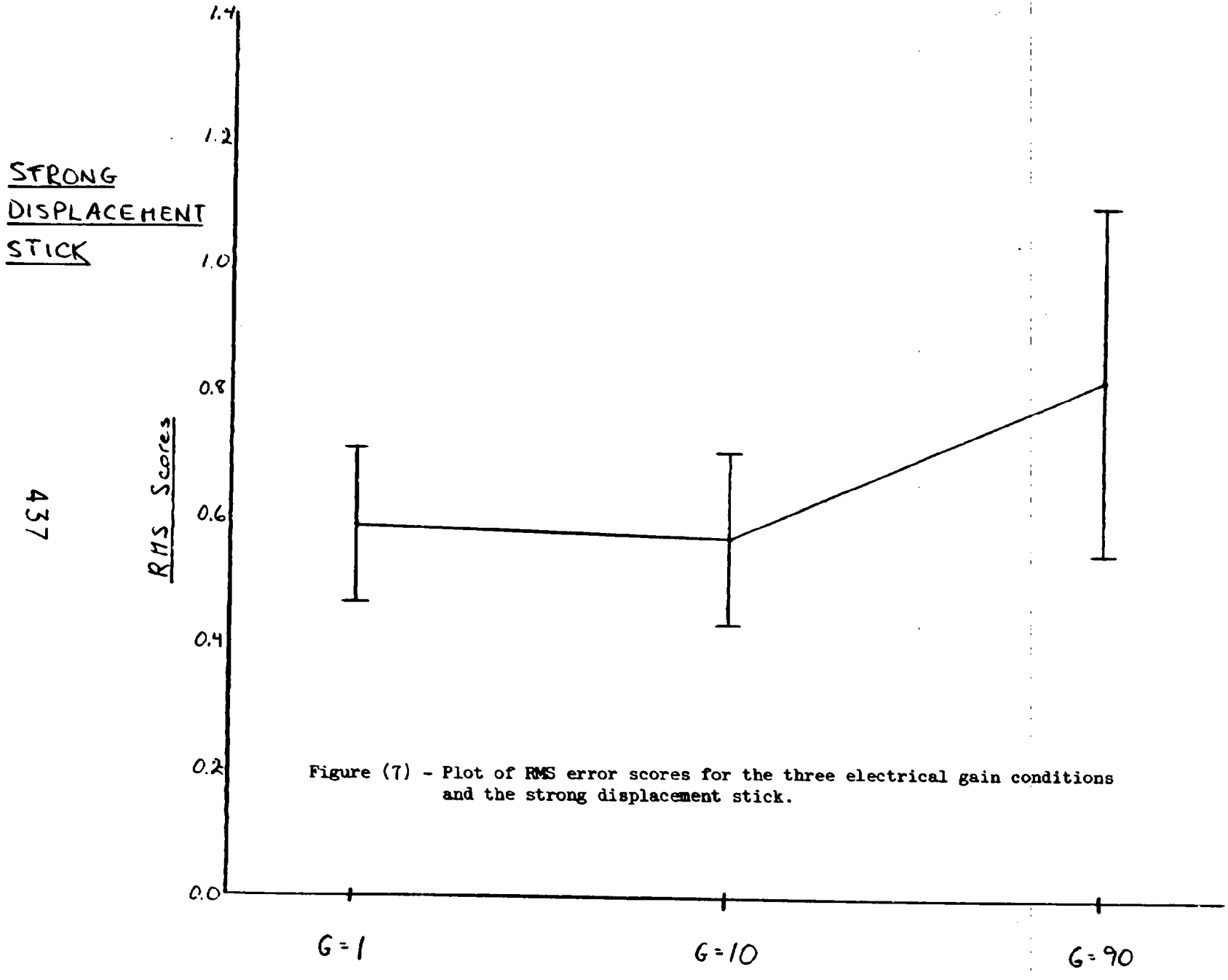


Figure (7) - Plot of RMS error scores for the three electrical gain conditions and the strong displacement stick.

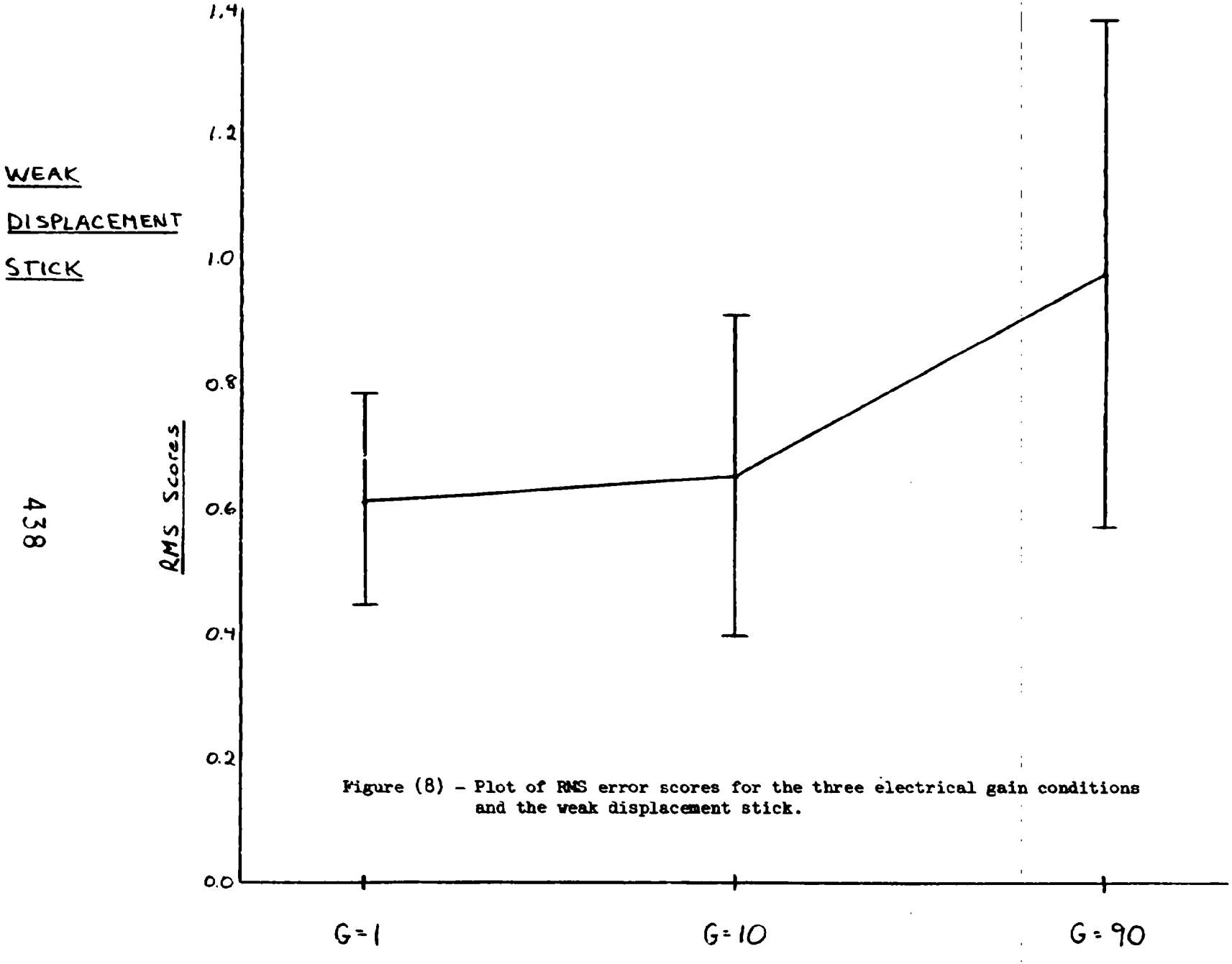


Figure (8) - Plot of RMS error scores for the three electrical gain conditions and the weak displacement stick.

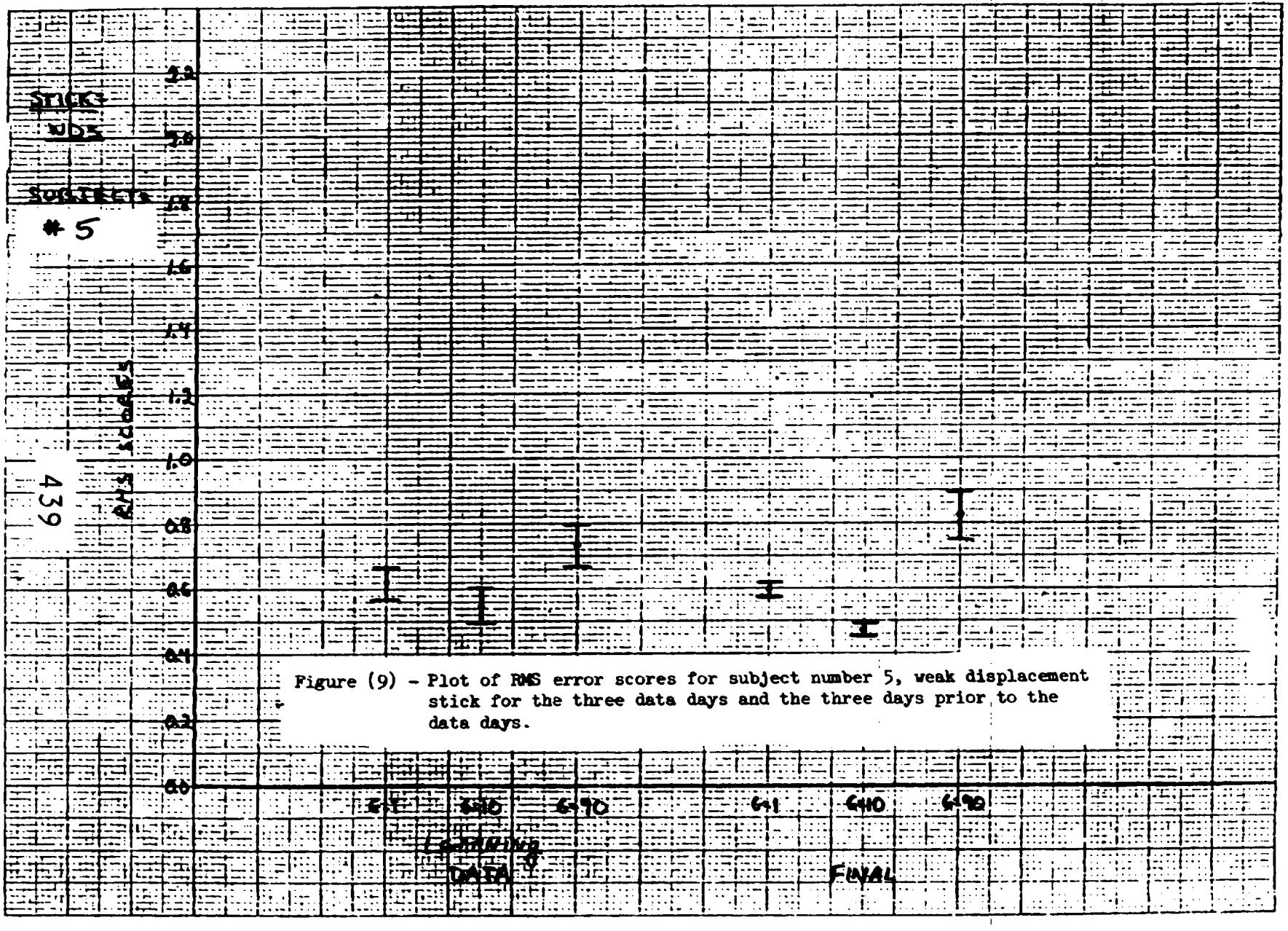


Figure (9) - Plot of RMS error scores for subject number 5, weak displacement stick for the three data days and the three days prior to the data days.

AD P001642

FORCE-TORQUE CONTROL EXPERIMENTS WITH THE SIMULATED SPACE SHUTTLE MANIPULATOR IN MANUAL CONTROL MODE

A. K. Bejczy and R. S. Dotson
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

J. W. Brown and J. L. Lewis
Spacecraft Design Division
Johnson Space Center
Houston, Texas 77058

SUMMARY

An experimental six degree-of-freedom force-torque sensor system has been developed at the Jet Propulsion Laboratory (JPL) and integrated with the simulated full-scale Space Shuttle Remote Manipulator System (RMS) at the Johnson Space Center (JSC). The sensor system provides data on forces and torques acting at the end effector. The sensor data are shown to the operator on a graphics display. The operator's response to the sensor data is through resolved rate manual control of the RMS. The opening/closing of the end effector is also in rate control mode. Two sets of control experiments have been conducted using force-torque sensor information. The first set involved the use of a task board equipped with "tools" and "modules". The operator's task was to remove the "modules" from their holes in the task board or insert them back to their holes. The removal or insertion of one of the modules required the use of tools which also were placed in holes in the task board. The second set of experiments involved the berthing of a payload into a latching mechanism through four V-shaped guides. Altogether more than fifty training runs and one hundred ten recorded test runs have been performed. The preliminary control experiments were successful and established the utility of the force-torque sensor system for geometrically and dynamically constrained operations. The sensory feedback significantly reduced the maximum force and torque loads during the constrained operations and permitted operational techniques that were perceived as "risky" previously. This paper briefly describes the experimental "sensor-end-effector-display" system and the control experiments, and presents a brief summary of the experimental results.

INTRODUCTION

Motivated by future space application requirements, an experimental system was developed at JPL for sensing and displaying the three orthogonal forces and the three orthogonal torques acting at the base of the end effector (or "hand") of an approximately 16 meter (50 feet) long robot arm at JSC. The arm at JSC simulates the functions of the Space Shuttle RMS. The experimental system contains the following main components and capabilities:

a) Two force-torque sensors; one is operating in the 0 to 100 lb (0 to 445 N) range, the other in the 0 to 200 lb (0 to 890 N) range.

b) A servo-controlled end effector drive system using a brushless DC torque motor in position or rate control mode; the rate control can be

proportional or preselected fixed rate control.

c) An interchangeable three-claw and four-claw end effector, interfaceable to both force-torque sensors, and controllable by the same end effector drive system.

d) A computer graphics terminal for displaying the acting forces and torques to the operator in real time; the graphics display is programmable for alternative scales and formats, the selection of which can be controlled manually or by a computer-recognized voice command system in real time.

e) A network of dedicated microcomputers supporting the sensor data handling, the control of end effector drive system, the graphics display and the voice command system. Capability is provided for real-time data transfer between this microcomputer network and the SEL 32/35 minicomputer in the JSC Manipulator Development Facility (MDF).

f) Control input peripherals for position, fixed rate and variable rate control of the end effector.

g) An eight-channel analog chart recorder, interfaced to the microcomputer network, for recording sensor data and end effector status for performance evaluation.

More on the technical details of the "sensor-claw-display" system can be found in Ref. 1. The "sensor-claw-display" system and the overall experimental system configuration are shown in Fig. 1.

The experimental "sensor-claw-display" system has been integrated with the simulated full-scale Space Shuttle RMS at JSC and control experiments were conducted for geometrically and dynamically constrained manipulation tasks. The testing and evaluation experiments were aimed at studying the enhancement of the Space Shuttle RMS performance in operations which involve manipulation tasks with narrow geometric and dynamic tolerances. Such tasks are involved in operations like satellite retrieval, maintenance and repair, in-orbit servicing of reusable vehicles, handling of large payloads which may require the docking of large masses, assembly of structures in space, etc.

CONTROL TEST SCENE

Two sets of control experiments were performed using force-torque sensor information. The first set of tests involved the use of a task board equipped with "tools" and "modules". The task board was designed and built at JSC. The second set of tests involved the berthing of a payload into a latching mechanism through four guides. The payload simulated the Plasma Diagnostic Package (PDP) under zero gravity conditions.

Task Board Tests

The "tool" and "module" handling task board is shown in Fig. 2. It was placed in the bay of the Shuttle mock-up, about 8 meters (25 feet) from the

Shuttle cockpit. The task board contained (a) a box, (b) a keyed cylinder, (c) a screwdriver, and (d) a square-base wrench. To accommodate the triangular and quadrangular grasp geometries of the three- and four-claw end effectors, two sets of tools and modules were used: one set with triangular handles and another set with quadrangular handles. The surface of the tool handles were coated with a special paint providing a rough surface texture in order to prevent the claws from slipping. Each end of the tool handles were also equipped with small metal cleats helping to keep the claws on the handles.

The tool and module handling tests had two phases of manipulation activities. The objective of the first phase was to remove the keyed cylinder and box from the task board and place them aside. Figure 3 illustrates the first phase manipulation. The cylinder removal did not require the use of tools, but the box removal did as illustrated in Figure 4. Before box removal, the box had to be unlatched. This required the turning of two retaining latches, each with a different tool. First, the screwdriver was removed from its retaining hole and inserted into the head of the upper latch. Then the latch was turned ninety degrees to the right or to the left, and the screwdriver was removed from the head of the latch and returned back to its retaining hole. Then the wrench was removed from its retaining hole and inserted into the head of the lower latch. The latch was then turned ninety degrees to the right or to the left, the wrench was removed from the head of the latch and returned to its retaining hole. Thereafter the box was removed and set aside. This completed the first phase manipulation. The second phase manipulation was just the reverse of this. The objective of the second phase was to insert the keyed cylinder and box back to their retaining holes in the task board. First the box was picked up and inserted into its retaining hole. The two latches were then turned back using the screwdriver and wrench to lock the box into its place. Thereafter the keyed cylinder was inserted into its retaining hole in the task board. This completed the second phase manipulation. All insertion tolerances on the task board were 6 mm (0.25 inches).

Payload Berthing Tests

The objective of the payload berthing test was to maneuver the simulated PDP payload into a retention or latching mechanism shown in Fig. 5. The latch assembly was placed in the bay of the Shuttle mock-up about 10 meters (30 feet) from the Shuttle cockpit as shown in Fig. 6. The berthing tests were performed so that the weight of the mock-up PDP payload (about 250 lb) was counterbalanced through a pulley attached to an overhead crane as also seen in Fig. 6. In this way the only forces and torques generated at the force-torque sensor were those caused by the payload contact with the latch assembly. The counterbalance arrangement allowed all small translational and rotational movements of the manipulator necessary for the tests. The tests started with lowering the guide pins of the PDP payload to the point that they were almost touching the V-shaped guides of the latching mechanism.

The latching mechanism consists of four V-shaped guides as shown in Fig. 5. Two are on the forward end of the mechanism, and two are on the port side. Three microswitches (labeled A, B and C in Figure 5) are closed whenever the payload is level and touching the bottom of the guides. Three indicators inside the flight deck area of the cockpit indicate the on-off state of the

three microswitches. To latch safely requires that all three microswitches are on. This in turn requires a simultaneous contact at points A, B and C. After simultaneous contact the payload is latched by a motor which pulls it aft along the horizontal notch in the guides and locks it in place. It is noted that the location (A, B and C) of the three microswitches is invisible to the operator once the payload guide pins are inside the V-shaped guides.

The initial placement of the payload during the test runs contained some small translational and/or rotational misalignments relative to the V-shaped guides and relative to the contact plane of the latch assembly as shown in picture (a) of Figure 7. Consequently, during a "down" motion of the MDF arm, some of the payload guide pins contacted the inside edge of the V-shaped guides generating undesired contact forces and torques between the payload and the latch assembly. This can lead to (a) an unever contact between the payload and the bottom plane of the latch assembly preventing the simultaneous closing of all three (A, B and C) microswitches in the contact plane, (b) overloading the latch assembly, or (c) jamming the payload guide pins into the V-shaped guides. Ideally, only a small "down" force should be acting between the payload and the latch assembly at the terminal contact, and all lateral forces and all torques should be zero or near zero. Pictures (a), (b) and (c) of Fig. 7 show a full berthing sequence.

INFORMATION AND CONTROL CONDITIONS

The operator had three basic sources of information for guiding and controlling the RMS during the tests: (a) direct visual access to the scene through the cockpit windows, (b) two closed circuit TV monitors showing the scene selectable from several TV cameras located on the wrist and elbow of the RMS and around the cargo bay area of the Shuttle mock-up, and (c) a graphics display of the force-torque sensor information.

The basic RMS control was manual using two three-dimensional hand controllers for RMS control in resolved rate control mode: one hand controller (operated by the left hand, see Fig. 8) controls the three translational components of RMS end effector motion, the second hand controller (operated by the right hand, see Fig. 8) controls the three rotational components of RMS end effector motion. The on-off switch, which controls the opening and closing of the RMS end effector, was replaced with a linear potentiometer arrangement providing proportional rate control capability for opening and closing the claws.

The direct visual and TV information sources and the basic RMS control are Shuttle baseline arrangements. The graphics display and the proportional claw control were specifically developed for the force-torque control experiments.

Graphic Display

The forces and torques measured by the sensor at the base of the claws were displayed to the operator on a 9-inch B/W monitor in graphics format. This monitor was mounted to the right of the TV monitors as shown in Fig. 8. The graphics display generator used in the present experimental system has a resolution of 512 by 512 pixels and is capable of displaying up to eight colors.

(The color capability was not used since the baseline flight deck monitors are B/W.)

The initial format chosen for displaying forces and torques is a very simple "bar chart" display, and a rotating two dimensional vector. The bars are arranged horizontally in the upper half of the screen (see the top of Fig. 1). Each bar corresponds to one of the six forces and torques. If there is a positive force or torque, the bar elongates to the right. A negative force/torque causes the bar to extend out to the left. There are vertical grid markings on the screen to enable quick estimation of the applied forces. Underneath each vertical grid line is a number indicating the corresponding force or torque. This scale may be changed either from the users console, or by voice command. In addition, by a keyboard or by voice the user may cause each force and torque to be labeled indicating whether it is an X, Y, or Z force or torque and what its exact numeric value is. Currently the forces and torques are shown only in the reference frame of the sensor, although the axis may be labeled as desired by the user via commands to the display computer. The sensor reference frame is shown in Fig. 9 in relation to the four-claw end effector geometry.

Beneath the force/torque bar chart display appears the last word recognized by the voice recognition system. The word blinks if the voice system is active.

To the right of this word display is a small two dimensional vector and coordinate system that shows the direction and magnitude of the resultant of the two forces, orthogonal to the end effector roll axis as it would appear if viewed through the TV camera mounted on the end effector. The position of the tip of the vector is indicated by a small cross to enhance visibility. Adjacent to this is a vertical line indicating the direction and magnitude of the force along the end effector roll axis. If it extends upward, the wrist force is a push. If it extends downward, the wrist exerts a pulling force. This line and the 2D vector is biased and scaled with the same commands that apply to the bar chart.

At the bottom of the screen are horizontal bars indicating the position of the claw. As the claw is closed the bars extend toward the center of the screen. When the claw is fully closed it appears as a solid horizontal bar on the display.

Existing forces and torques measured by the sensor at the base of the claws can be biased out from the readings shown on the display (setting them to zero) at any time and for whatever position/orientation the end effector is in at that particular time. Using this display biasing capability just before an object held by the end effector comes in contact with the environment enables the operator to see the actual contact forces and torques on the display. This display biasing capability in effect simulates zero gravity conditions for payload handling tasks at the moment when contact between payload and environment is established. The display can be unbiased at any time. The unbias command will cause the true forces and torques acting on the end effector (including gravity) to be displayed. The "bias" and "unbias" display commands can be activated manually (by a push-button switch or a TTY key) or by voice.

Claw Control

Three methods were implemented for controlling the closing/opening of the end effector claws: position control, fixed rate and proportional rate control. Position control was achieved with a three-turn potentiometer mounted in a small box and placed next to the flight deck window. (This box also housed the push-button switch for controlling display "bias".) The operator could simply dial in the amount of closure he wanted the claw to assume. When the claws reached the desired position, the brake on the end effector drive motor was automatically set, and the motor turned off automatically.

Fixed rate control was implemented by using a thumb switch for selecting the desired rate and accepting the close/open commands from the rocker switch mounted on the standard RMS flight model orientation control joystick. The thumb switch was mounted to the same box which housed the three-turn potentiometer for position control. Fixed-rate control was not used during the tests.

The predominantly used control for claw opening/closing was the proportional rate control mode. This was implemented by modifying a copy of the standard RMS flight model orientation control joystick. The rocker switch was replaced with two linear potentiometers with about 2 cm (0.75 inches) travel. One was mounted above the other. The upper trigger was for opening the claws, the lower one was for closing. The opening trigger was equipped with a safety cover to preclude accidental opening. Motor velocity was proportional to the distance the trigger was pulled in. The triggers were spring loaded so that they would return to zero position, and the motor would stop moving if the both triggers were released. An exception to this was in the "grasp" mode where a constant grasp force would be applied to the tool after trigger release if the operator had commanded a maximum-travel closing. This grasp force would be applied to the tool handle until the operator issued an open command.

TEST RUNS AND DATA

More than 50 training runs and 110 recorded test runs have been performed by four test operators. The force-torque sensor and claw open/close position data were recorded in both analog and digital formats. An eight-channel strip chart recorder was used to record either the three force and three torque components plus the claw position, or the eight raw (unconverted) strain gage readings.

The force-torque sensor and claw open/close position data in digital form were sent to the JSC MDF SEL 32/35 minicomputer over a serial line as a set of signed decimal ASCII numbers. Each set of readings is separated by carriage-return. Data is buffered at a rate of approximately 5 to 6 Hz in the memory of the M6809 graphics control microcomputer until it could be sent to the SEL 32/35 and to the hardcopy printer. Data buffering was necessary since, in the worst case, the real-time data can be changing faster than the serial transmission line can accept the data. The time of day to the nearest second was appended to the data since in some cases 2 to 3 minutes of data had been buffered up before the SEL 32/35 could accept the data through the serial transmission line.

The time-dating of the F-T sensor data allows correlating the F-T sensor data with the other RMS data recorded by the SEL 32/35 for performance evaluation. The digital data records (either force-torque information or raw strain gage readings) can also be displayed in real time on a CRT console connected to the M6809 graphics control microcomputer.

Familiarization Phase

Four test operators performed familiarization runs for about a week. These runs, which were distributed over a three week time period, were necessary since the use of the force-torque sensor-claw-display system in manual control of a large and flexing robot arm represented a very first and unique experience to everyone involved in the tests.

The operators had to learn about several aspects of the new sensor-claw-display system: (1) The nature and functional meaning of the multidimensional force-torque sensor information; three orthogonal force and three orthogonal torque components are simultaneously displayed to the operators. (2) Correlating the force and torque components displayed in alternative reference frames (sensor/claw or orbiter coordinates -- see Fig. 9) with the use of the two three-dimensional joystick controls. (3) Correlating the displayed force-torque information and the joystick control inputs with the task control at hand; this involved the mental integration of visual and graphically displayed force-torque sensor information and the calibration of joystick control inputs with the dynamic response of the system as seen on the force-torque sensor graphics display. (4) The use of the new position and rate controls of the claw; as a consequence of the training runs, only proportional rate control was used for controlling claw opening/closing during the tests because of its flexibility. (5) The difference in the visual geometric reference for using force-torque sensor information with the three-claw and four-claw end effector; again, as a consequence of the training runs, only the four-claw end effector was used during the tests because of its inherent quadrangular symmetry.

As a result of the familiarization phase, the original test run matrix was limited to the small payload berthing and task board test scenarios.

Payload Berthing Test Data

The test scenario and the operator's control tasks related to payload berthing were described in a previous section. For control reference, the Orbiter Body Reference Frame (see Fig. 9) was used. The berthing tests were performed under three diferent information feedback conditions.

A) Sensor display only: the operator had only access to the graphic/numeric display of the force-torque sensor information; the cockpit window was blocked, and the TV monitors were turned off.

B) Visual access only: the sensor display was turned off, and the operator had to rely on direct vision and/or TV information only.

C) Sensor display and visual access simultaneously: the operator had visual (direct vision and/or TV) access to the task scene and could also see the graphic/numeric display of force-torque sensor information.

The operators had to perform twelve payload berthing test runs in one setting, four runs for each of the three (A, B and C) information feedback conditions. The information feedback condition sequence was predetermined for a twelve-run setting, but the operators did not know under what information condition the next test run will be performed. Each operator had to repeat a twelve-run setting three times, preferably on three different days, or one in the a.m. and another one in the p.m. of the same day. Because of time and subject availability constraints, only two of the four test operators were able to complete the full thirty-six-run test plan for payload berthing. Their average time performance data are shown in Table 1. Typical force-torque time histories recorded on the analog strip chart recorder during the payload berthing test runs are shown in Figures 10 and 11.

The mean time data shown in Table 1 should only be interpreted as indicative rather than conclusive since (a) the data base upon which the mean times were computed is statistically narrow, and (b) the time performance is dependent on the overall hardware configuration employed in the tests. Nevertheless, Table 1 and Figures 10 and 11 show a few significant points:

(1) The most interesting result is that all operators consistently could perform the payload berthing without any visual feedback, relying only on graphics display of force-torque sensor information during the terminal phase of berthing when the payload guide pins were inside the V-shaped guides of the latch assembly. However, operator comments indicated the desirability of having some visual access to the RMS and the task scene.

(2) Using graphics display of force-torque sensor information for guidance, the operators could successfully control the excess contact forces and torques during the terminal phase of the payload berthing task.

(3) Without graphics display of force-torque sensor information, using only visual feedback, the operators had no idea about the magnitude and location of contact forces and torques generated during payload berthing (see the lower part of Fig. 11), though the latching was successfully accomplished.

(4) The time data indicate that the force-torque sensor information may contain more relevant guidance data than the visual information during the terminal/contact phase of the payload berthing task, since the average time under condition A is shorter than under condition B (see Table 1).

(5) The time data also indicate that the use of more sensory information (that is, the simultaneous use of visual and graphics display of

force-torque sensor information) may lead to longer performance time unless the information is properly coordinated in order to ease the operator's perceptive workload. Note that the average time under condition C is longer than under conditions A and B (see Table 1).

It is worth noting two facts: (a) The time data shown in Table 1 are related to 10 to 12 cm (4 to 5 inches) total travel of the payload. This indicates that the terminal/contact phase of the payload berthing task is a time consuming operation. (b) There is a considerable spread of performance times around the mean time (see max. and min. times in Table 1). This indicates that minor variances in the initial conditions (that is, minor variances in the appraoch conditions to the V-shaped guides of the latch assembly) and minor control performance variations have a considerable impact on performance time.

Task Board Test Data

The test scenario and the operator's control tasks related to "module" and "tool" handling on a task board were described in a previous section. For control reference, the Sensor/Claw Reference Frame (see Fig. 9) was used. The nature of the control tasks required that, for guidance/control information, the operators have access to all three information sources all the time: direct vision, TV cameras/monitors and graphics display of force-torque sensor information.

As described previously, a full "tool" and "module" handling test had two phases: (A) removing the modules from the task board and (B) reinserting the modules back to their retainers in the task board. Each phase contained twenty-five subtasks. For Phase B (reinsertion of modules) the sequence of subtasks is listed in Table 2. Test data were only taken for Phase B manipulation activities.

Altogether sixteer Phase B test runs have been performed by four operators with recorded performance data. The mean times for each of the twenty-five subtasks, as listed in Table 2, are computed from ten test runs only since the sequence of subtasks in the other six test runs was somewhat different; the sequence of using the "red" and "blue" tools was there interchanged. (The "red" tool is the "screwdriver", the "blue" tool is the square-head "wrench".) Typical force-torque time histories recorded on the analog strip chart recorder during the "module" and "tool" handling test runs are shown in Figures 12 and 13.

The data shown in Table 2 are not comparative in nature since all test runs were performed under identical information feedback conditions. The same tests were planned with visual feedback only, but the available time did not allow it. The data in Table 2 should therefore be interpreted as indicative regarding the distribution of performance times among the subtasks. Note also the spread of performance times (max. and min. time in Table 2) for a subtask.

Note that the most time consuming subtasks are the insertions of the box and cylinder "modules". These are the riskiest subtasks where jamming can easily occur. Jamming occurs when the force applied in the direction of insertion no longer causes the insertion to proceed. When jamming occurs, it is possible to cause deformation of parts by the applied force. In general, jamming is caused by moving the direction of the applied force outside certain bounds (Ref. 2). The graphics display of force-torque sensor information was most useful for preventing jamming during box and cylinder insertion, illustrated in Fig. 12. The large amplitude variations in the F_Z force shown in the upper and lower part of Fig. 12 indicate situations where jamming could have occurred. The time history of the F_Z force variations shows that the operator prevented the jamming and successfully completed the box and cylinder insertions.

CONCLUSIONS AND PLANS

The preliminary tests established the utility of the graphically displayed force-torque sensor information for geometrically and dynamically constrained operations with a large and flexing robot arm simulating the functions of the Shuttle RMS since:

(1) The sensor system feedback helped in significantly reducing maximum force and torque loads generated during the simulated RMS operations.
(2) The sensor system feedback permitted operational techniques with the simulated RMS that were perceived previously as marginally acceptable or completely unacceptable.

Thus, the force-torque sensor system can be a significant aid in providing force-torque information for sensitive RMS operations.

The future plans include: (a) Integration of visual and force-torque sensor graphics display information. This can take several forms; for instance, superimposing the graphics display on the TV pictures and/or having a small plasma display at the flight deck windows. (b) Development of task-related, more concise and "intelligent" display formats to ease the operator's perceptive and decision burden. (c) Development of an experimental control system for using the force-torque sensor information in an interactive manual/automatic control mode. (d) Conducting further performance evaluation experiments.

Acknowledgment

The force-torque sensor was designed by V. Scheinman of Stanford University, and the sensor was strain-gaged by R. Talmo of Microgage, Inc., El Monte, CA. The end effector drive system, the analog and digital electronics was developed by H. C. Primus. The sensor calibration was performed by T. L. Turner. J. High modified the hand controller for proportional rate input. The precision machining of the sensors and other mechanical components was performed by S. Mihalko, D. Conrad and W. Kunberger. Software was developed by R. S. Dotson, overall system design by A. K. Bejczy.

The research described in this article was partly carried out at the Jet Propulsion Laboratory, California Institute of Technology, under NASA contract NAS7-100.

References

1. A. K. Bejczy and R. S.Dotson, A Force-Torque Sensing and Display System for Large Robot Arms, Proceedings of the IEEE Southeastcon '82, Destin, Fla., April 4-7, 1982.

2. S. Simunovic, Force Information in Assembly Process, Proceedings of the 5th International Symposium on Industrial Robots, Chicago, Illinois, September 22-24, 1975.

Table 1. Time Performance Data of Payload Berthing Tests

Operators / Information Condition	Operator No. 1 max.time	Operator No. 1 min.time	Operator No. 1 mean time	Operator No. 2 max.time	Operator No. 2 min.time	Operator No. 2 mean time	Overall Average Time
A	3:58	0:39	1:40	4:33	1:27	2:49	2:14
B	4:10	0:48	2:11	5:27	1:46	3:17	2:44
C	3:48	0:44	2:13	7:27	2:39	4:16	3:14

time in [min:sec]

A: only force-torque sensor display

B: only visual (direct and/or TV) feedback

C: both visual and sensor display feedback

Note: each "mean time" is computed from twelve test runs

Table 2. Time Performance Data of Task Board "Module" and "Tool" Handling Tests, Phase B

	mean Δt [min:sec]	max. Δt [min:sec]	min. Δt [min:sec]
START RUN			
BOX GRAPPLED	1:19	2:12	:26
BOX MANEUVERED	3:00	5:23	1:43
BOX INSERTED	4:26	21:20	:21
RED TOOL GRAPPLED	2:03	5:03	:01
RED TOOL EXTRACTED	0:30	2:18	:07
RED TOOL MANEUVERED	2:25	4:43	:20
RED TOOL INSERTED	1:19	4:23	:01
RED LATCH CLOSED	:38	1:49	:01
RED TOOL REMOVED	:11	:59	:01
RED TOOL MANEUVERED	2:45	4:28	:17
RED TOOL INSERTED	1:14	4:26	:40
RED TOOL RELEASED	:04	:12	:01
BLUE TOOL GRAPPLED	1:38	3:34	:15
BLUE TOOL EXTRACTED	:13	:21	:02
BLUE TOOL MANEUVERED	1:27	2:59	:34
BLUE TOOL INSERTED	1:45	4:52	:10
BLUE LATCH CLOSED	:34	1:20	:10
BLUE TOOL REMOVED	:13	:21	:02
BLUE TOOL MANEUVERED	1:25	2:53	:22
BLUE TOOL INSERTED	:59	2:36	:08
BLUE TOOL RELEASED	:11	:55	:01
CAN GRAPPLED	2:41	5:16	1:19
CAN MANEUVERED	2:25	5:00	1:13
CAN INSERTED	4:20	13:48	:28
CAN RELEASED	:07	:16	:01
TOTAL TIME AVERAGE FOR PHASE B TASK	37:53		

The mean time for subtasks is computed from ten test runs.

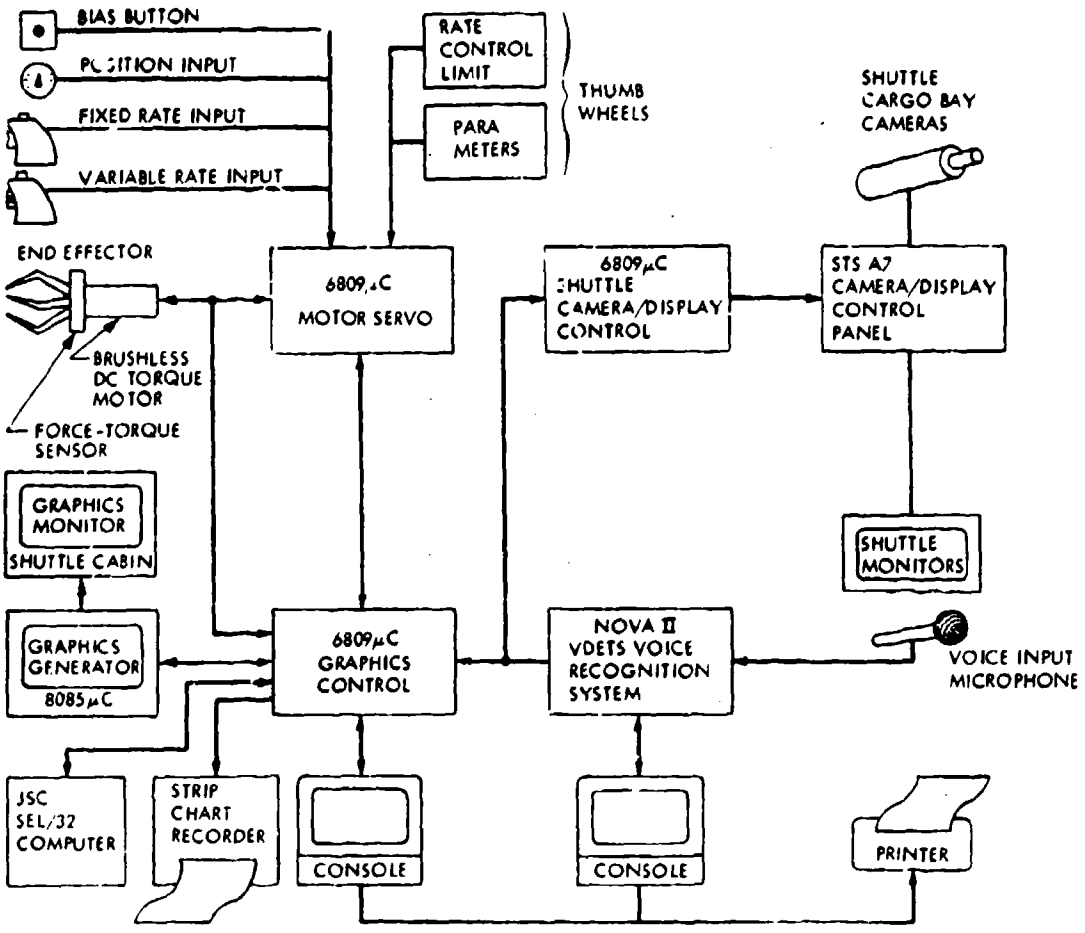


Figure 1. Sensor-Claw-Display System and Overall Experimental System Configuration

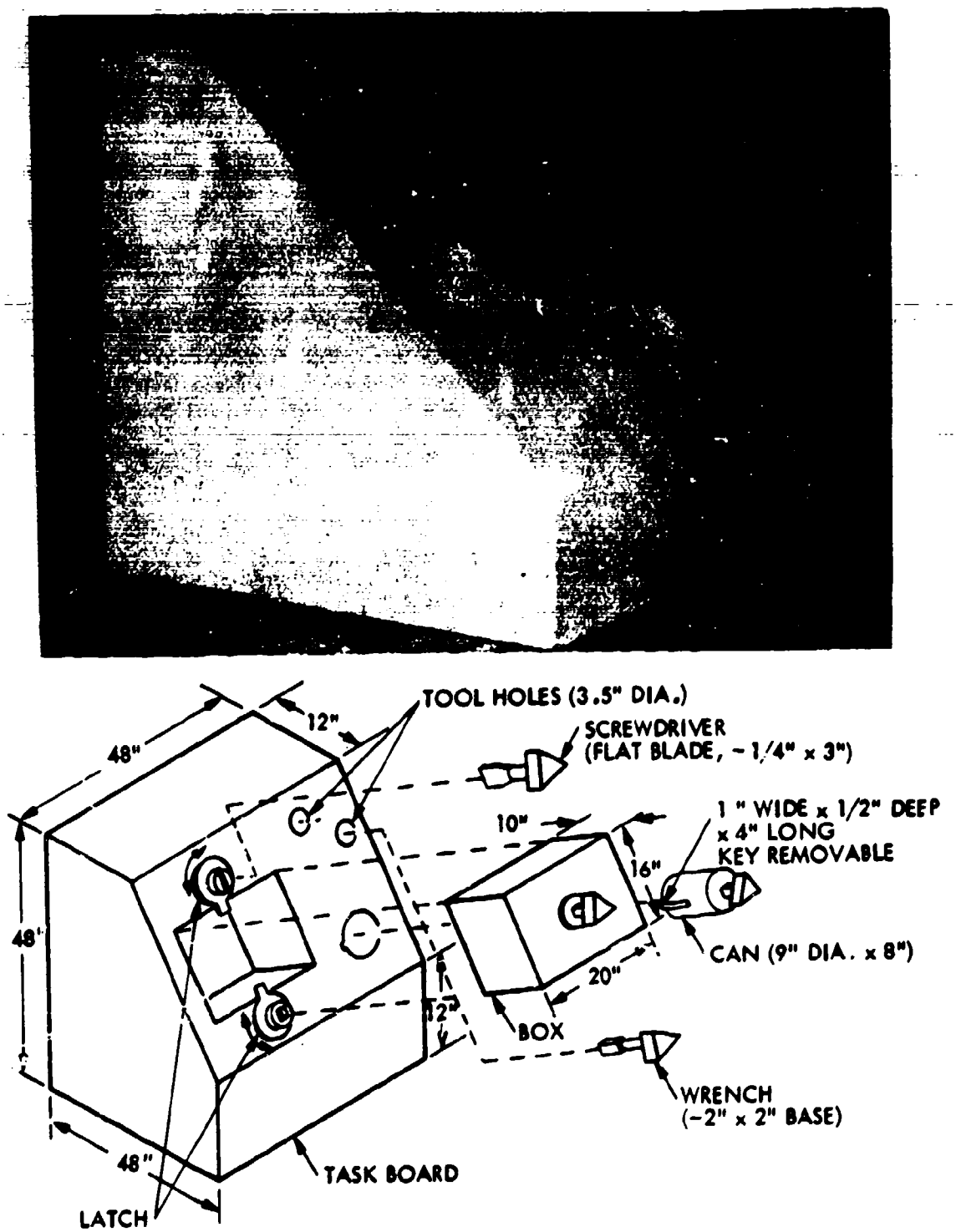


Figure 2. Tool and Module Handling Task Board, Located in the Bay of the Shuttle Mock-Up

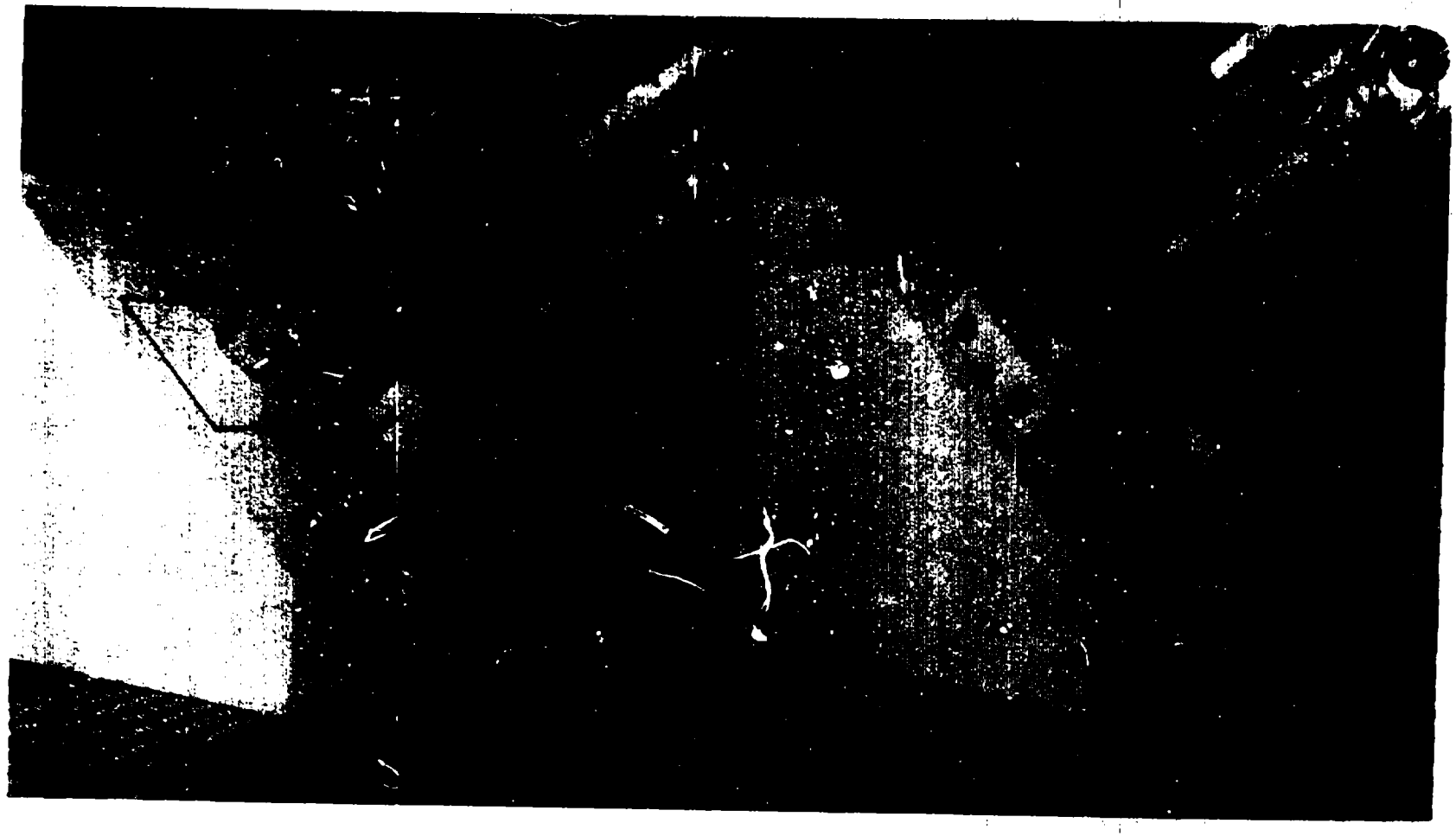


Figure 3. Cylinder and Box Module Insert/Removal Tasks

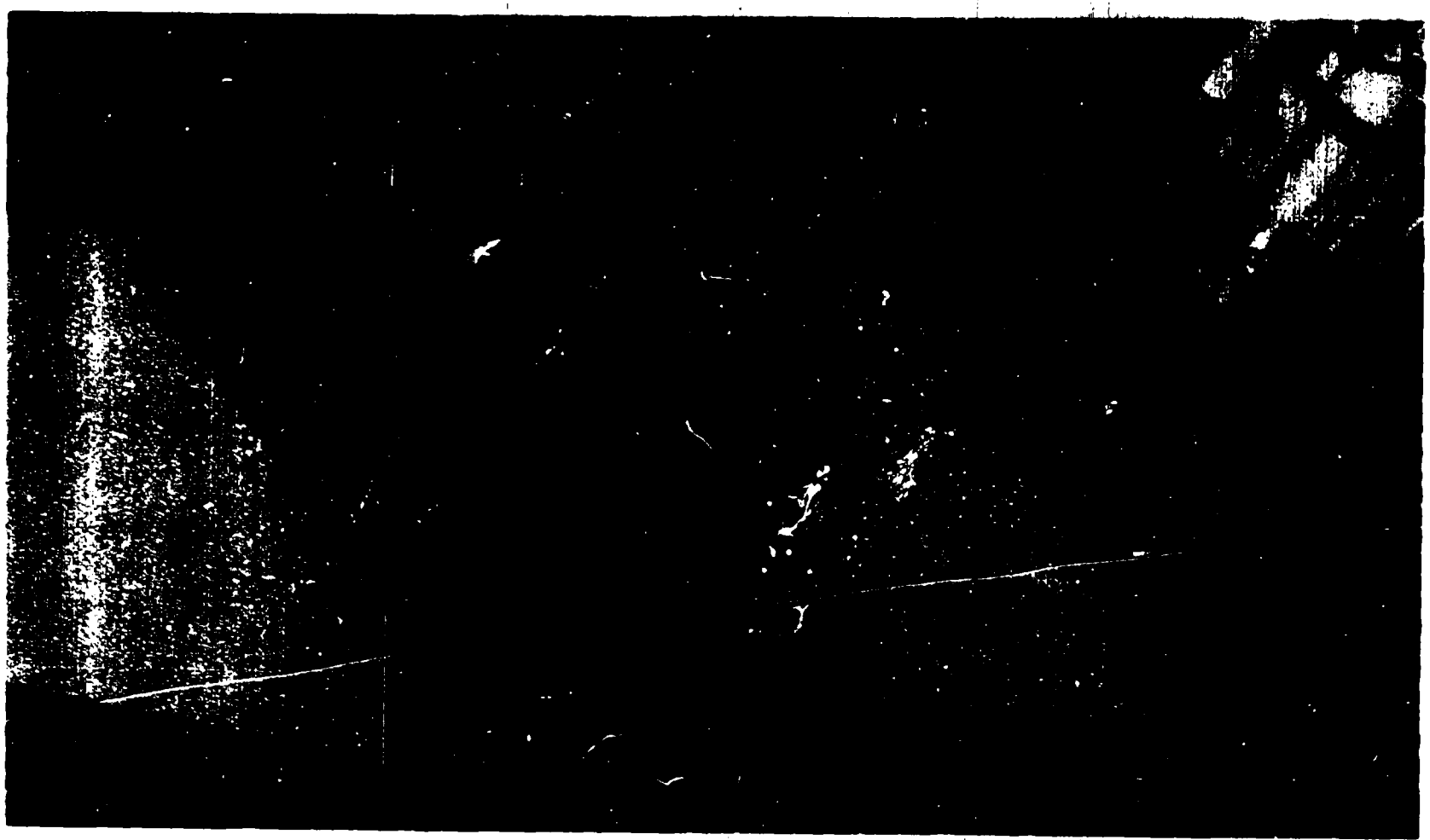


Figure 4. Latch Turning Tasks Using Flat or Square Head Tools (Screwdriver and Wrench)

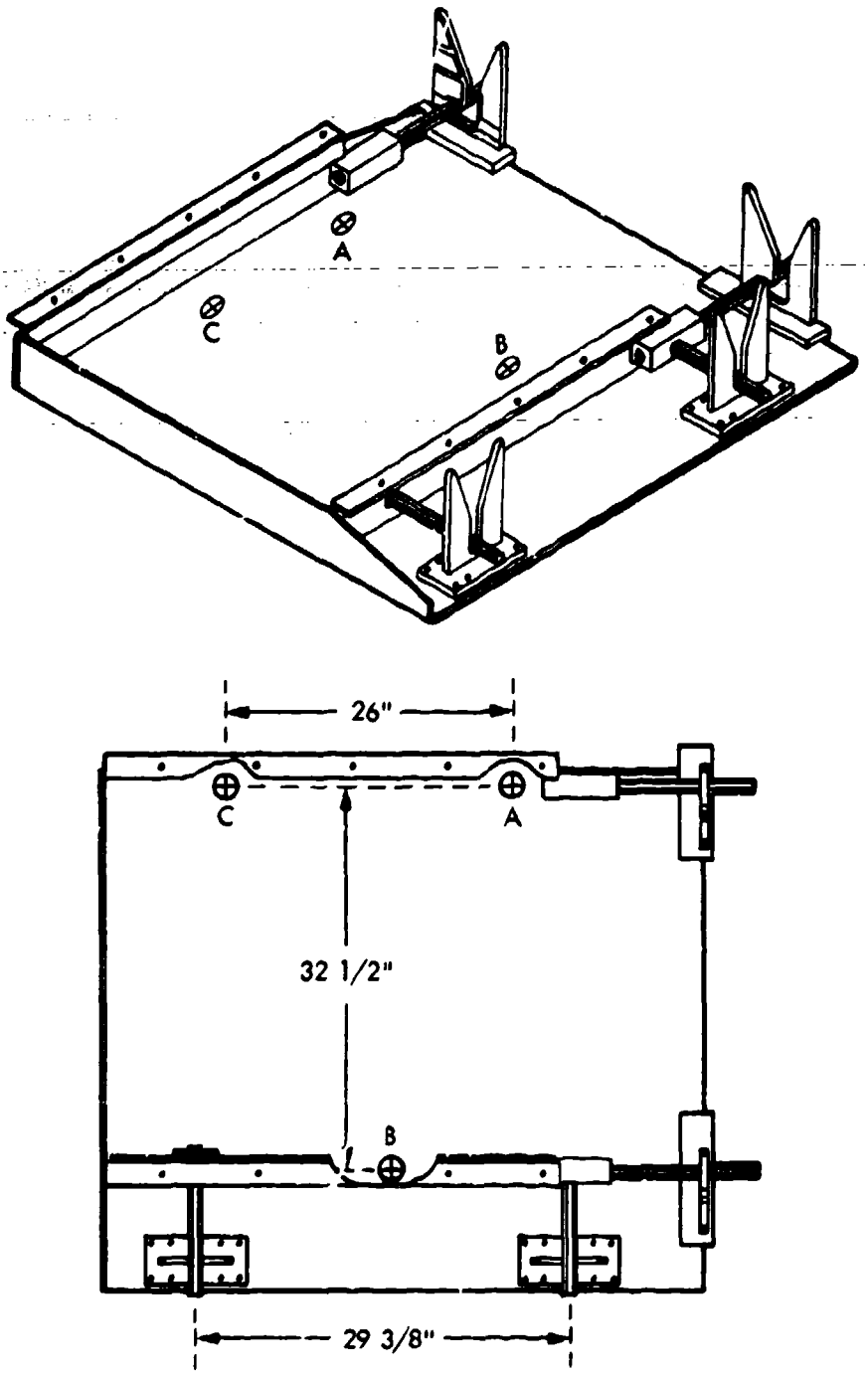


Figure 5. Latching Mechanism for PDP Payload Berthing in the Shuttle Bay

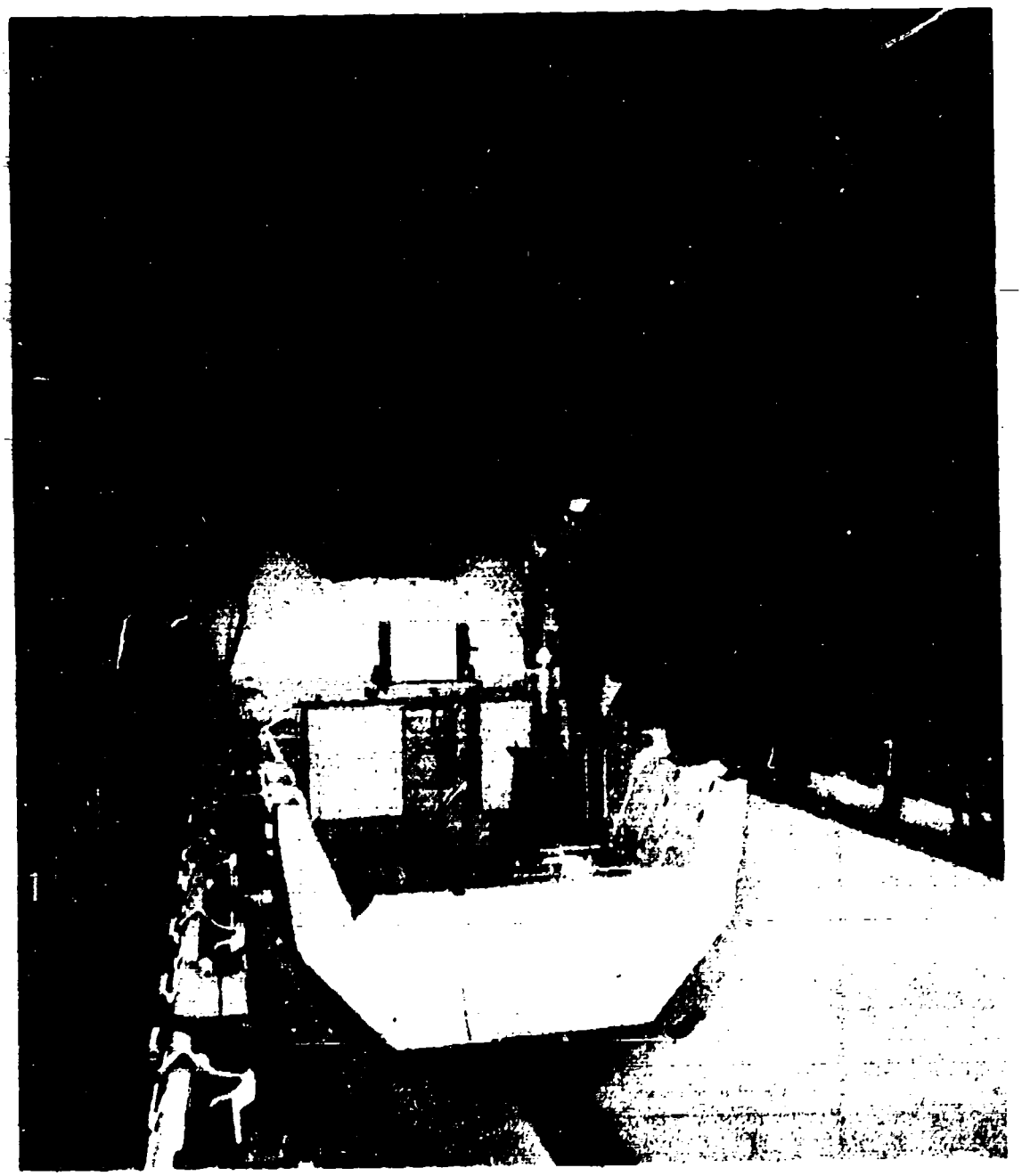


Figure 6. Overall Task Set-Up for PDP Payload Berthing, Including PDP Counterbalance to Simulate Zero-G Condition on the Payload

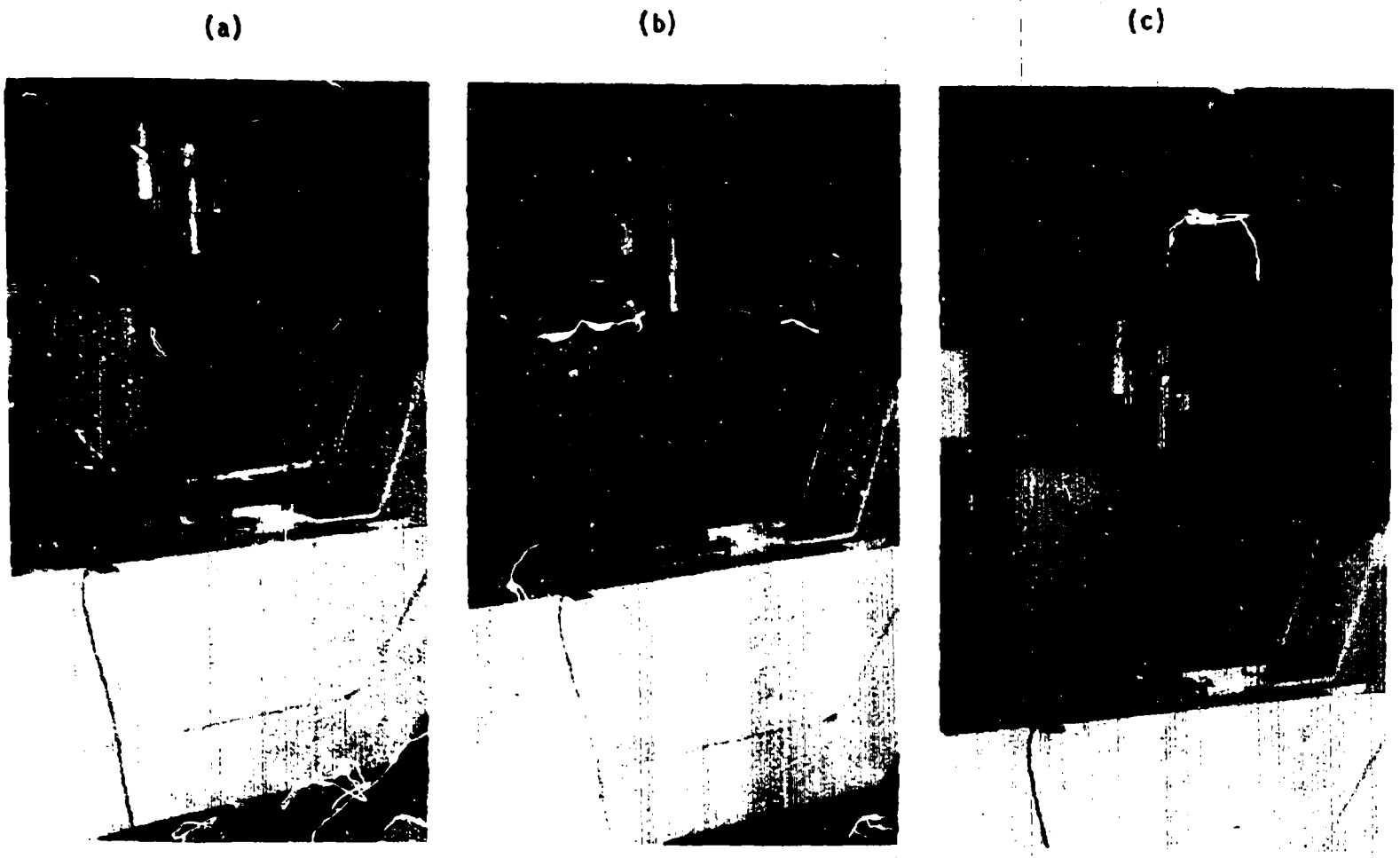


Figure 7. Berthing Sequence Through Mechanical Guides of Latch Assembly (from left to right)

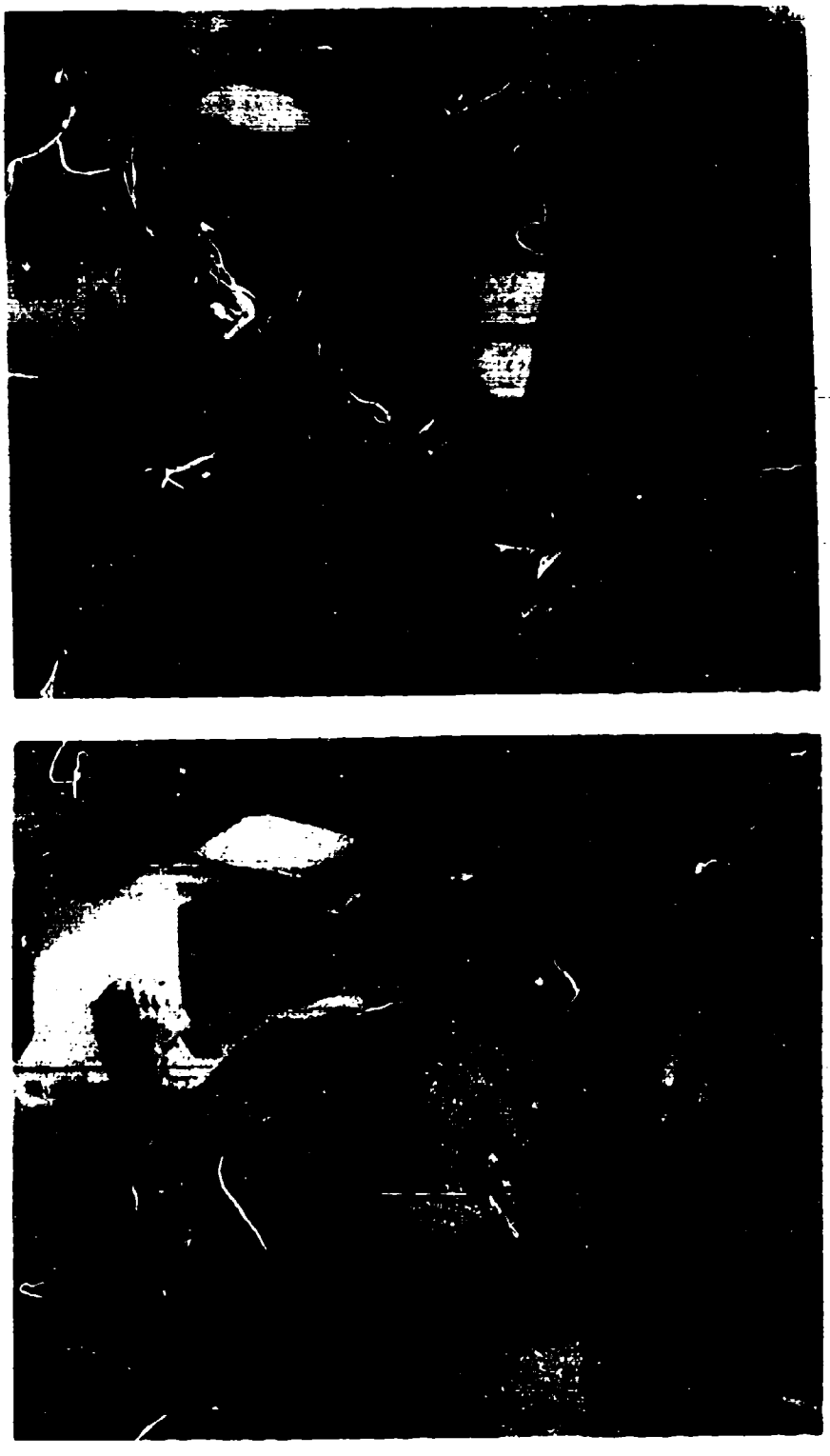
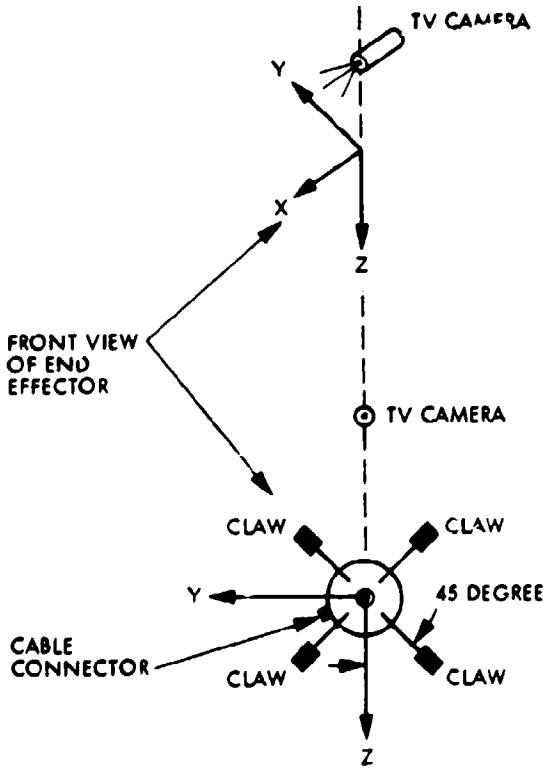
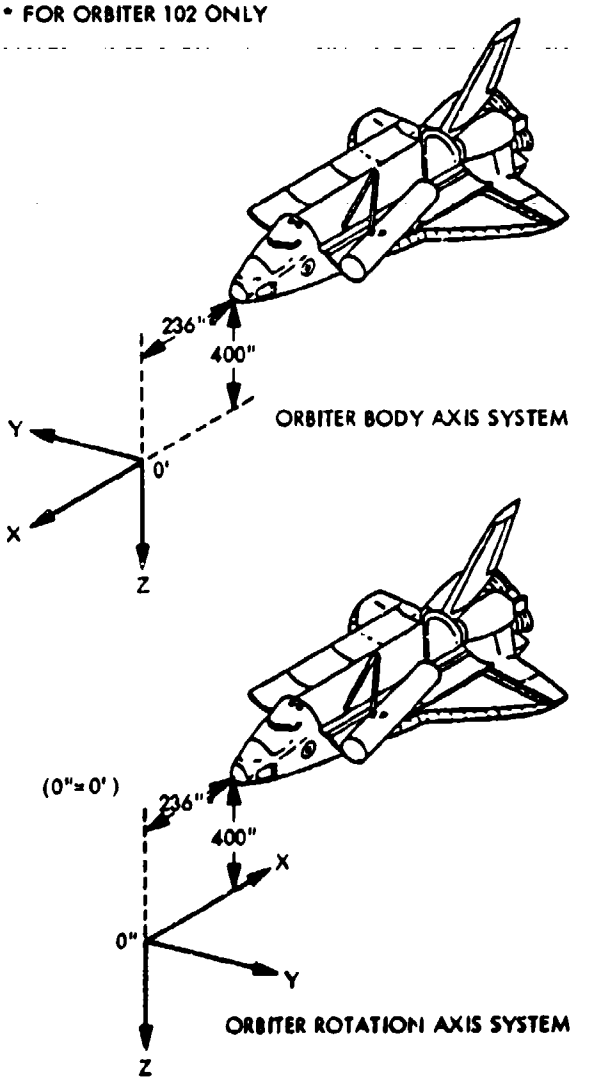


Figure 8. Test Operators in the Shuttle Cabin Use Graphics Display of Force-Torque Sensor Information (Graphics Display is to the Right of the TV Monitors)



F-T SENSOR REFERENCE FRAME IN FRONT VIEW IN RELATION TO FOUR-CLAW END EFFECTOR GEOMETRY

* FOR ORBITER 102 ONLY



ORBITER BODY REFERENCE FRAME

Figure 9. Sensor/Claw and Orbiter Reference Frames

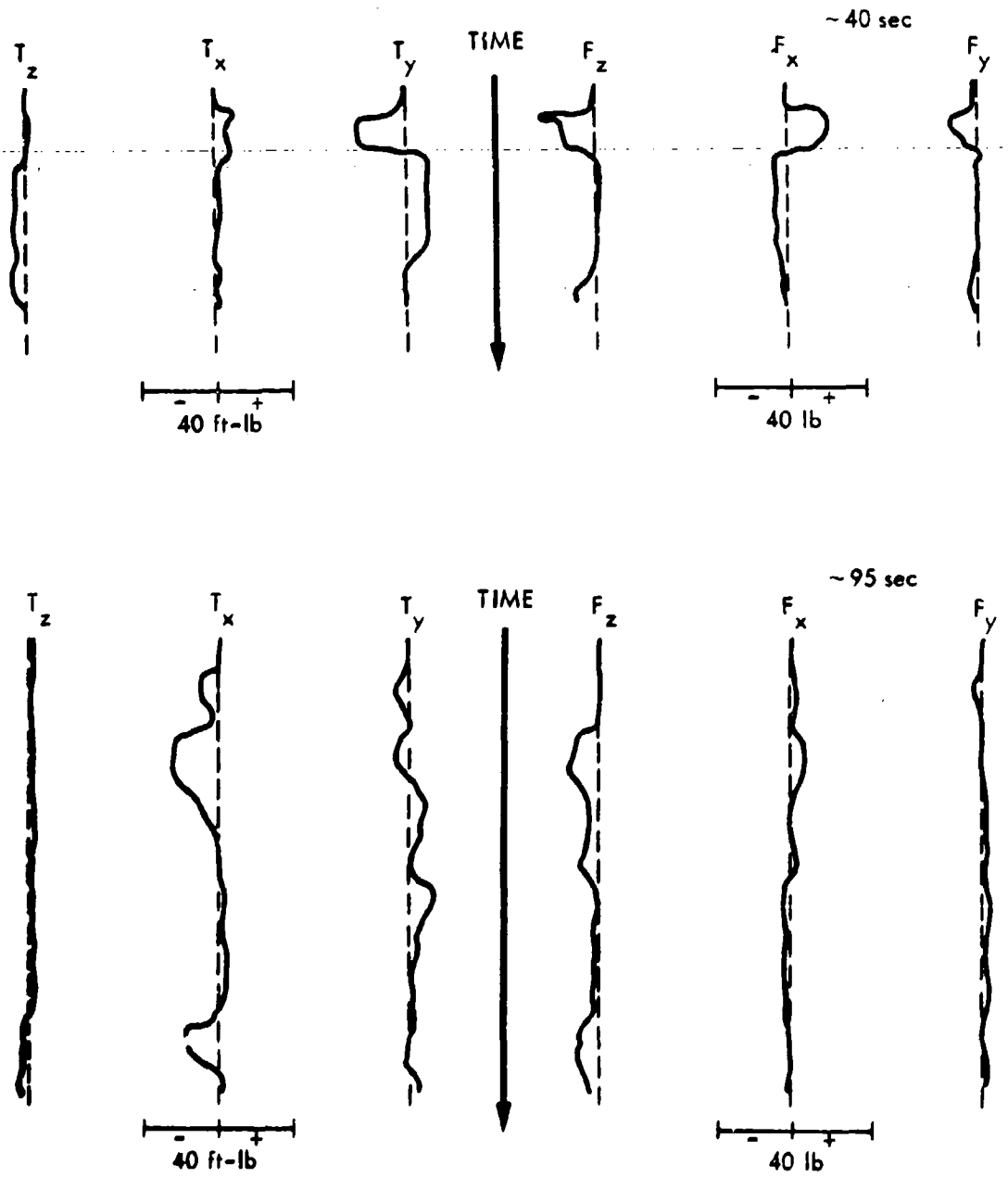


Figure 10. Typical Performance Data for Cases When Only Graphics Display of Force-Torque Sensor Data Were Used for PDP Berthing. (No Window and TV Monitor Viewing.)

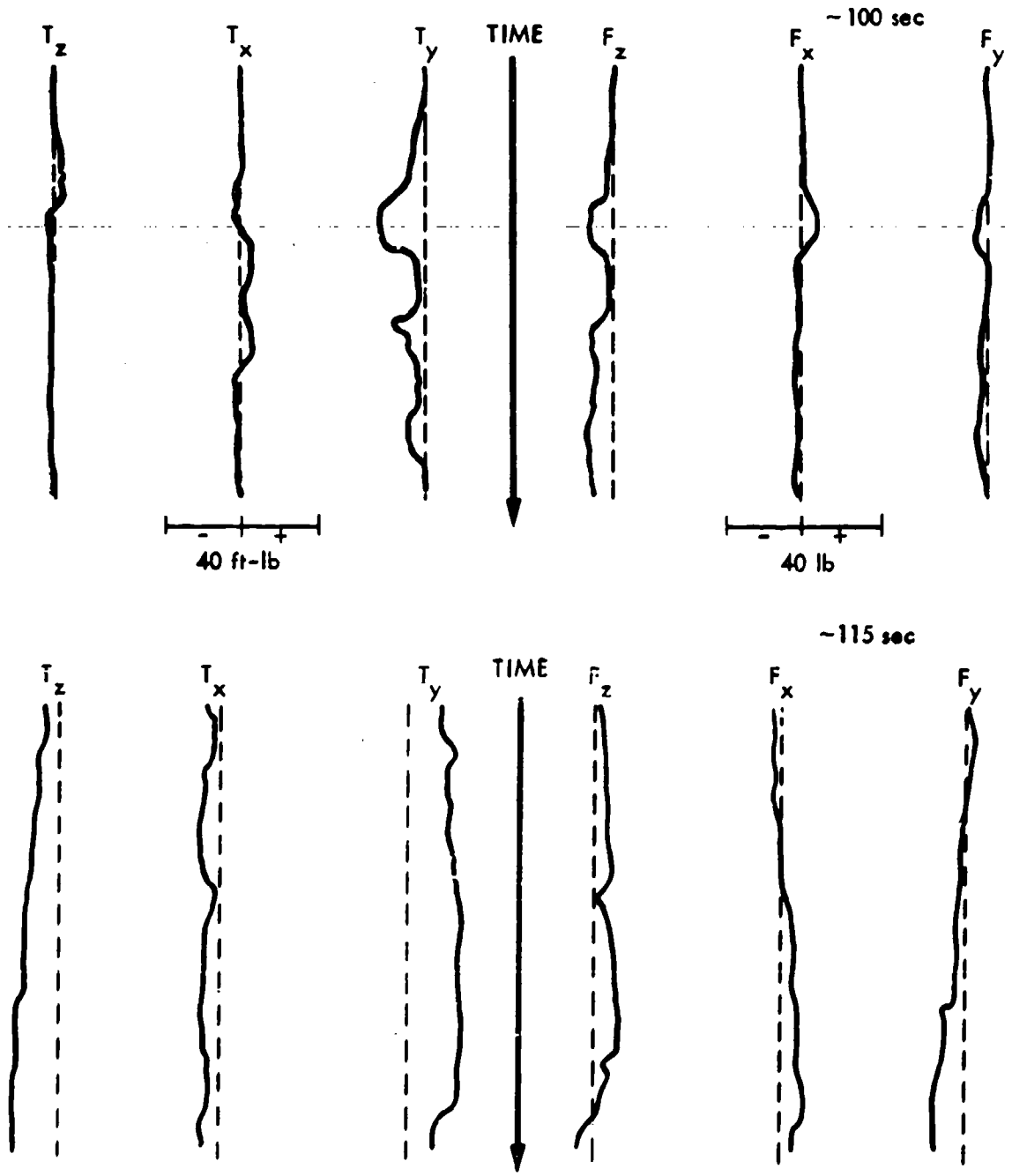
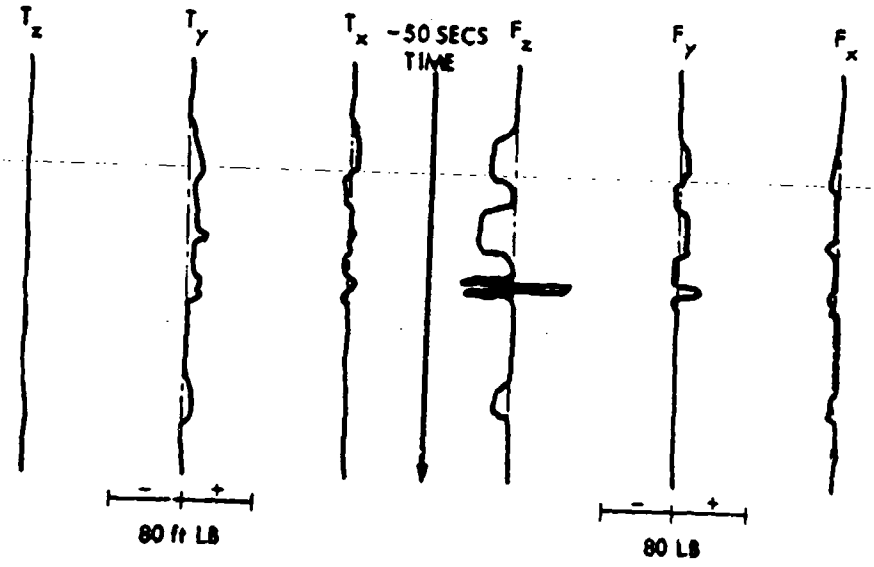


Figure 11. Typical Performance Data for Cases When a) Both Visual and Graphics Display of Force-Torque Sensor Data Were Used (Upper Figure), and b) Only Visual Information was used (Lower Figure)



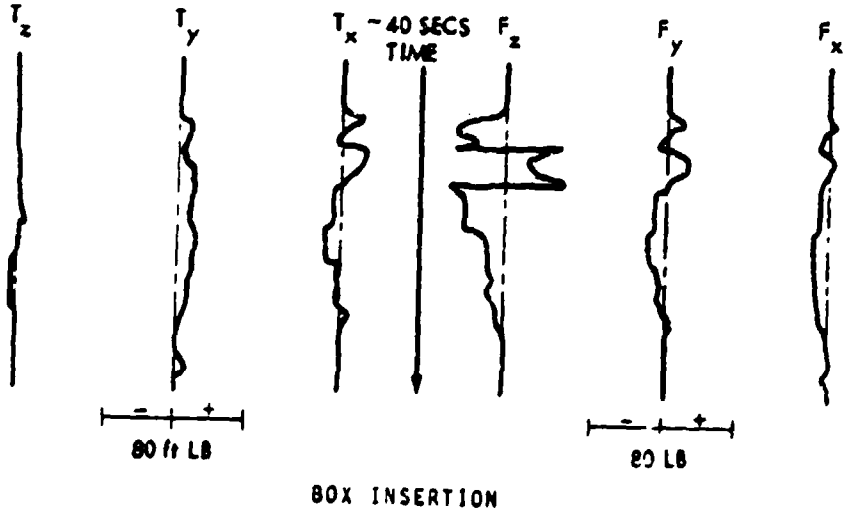
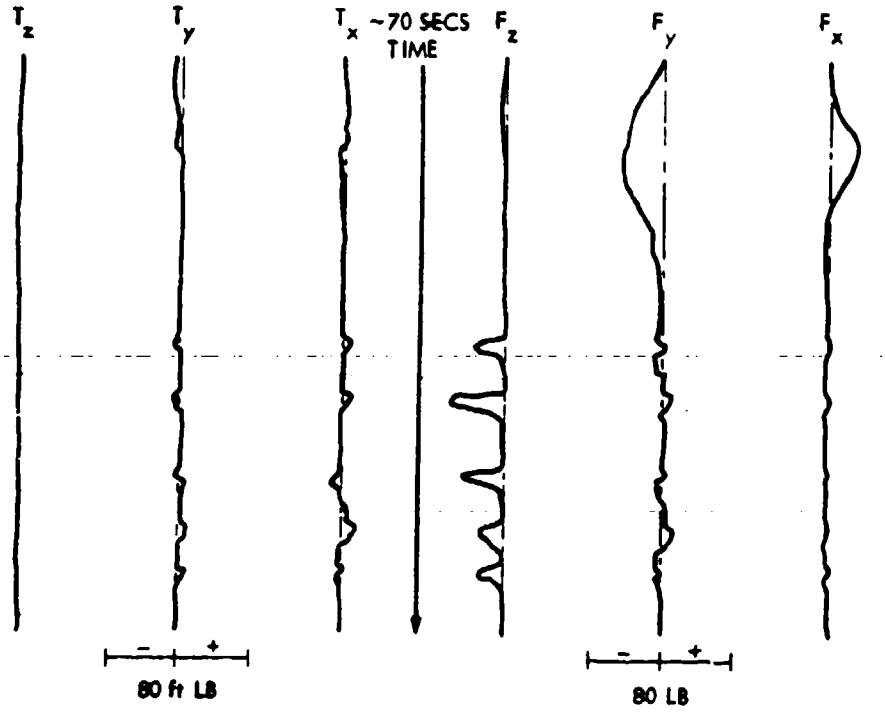
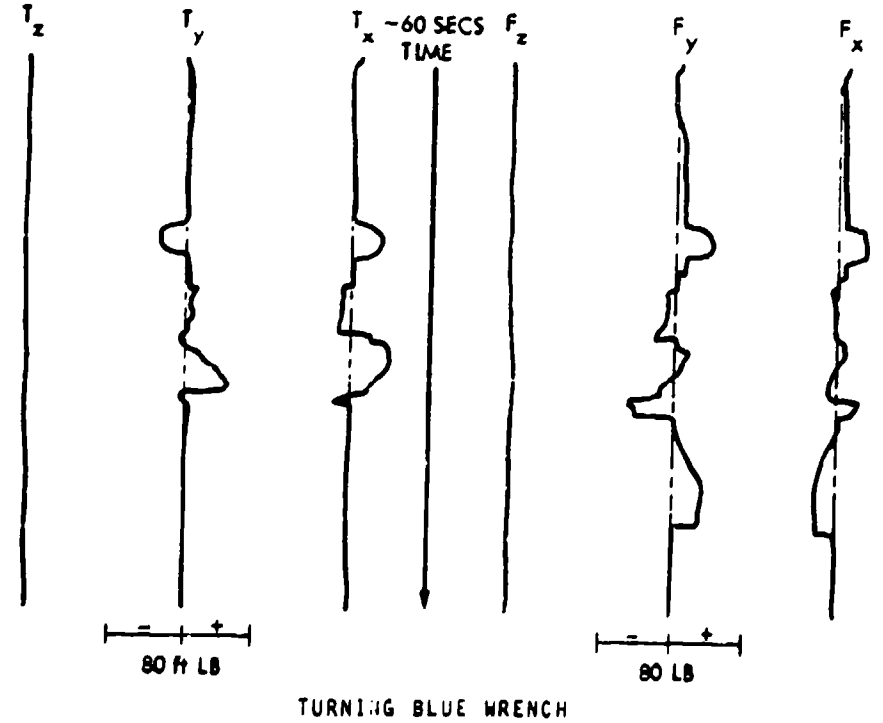


Figure 12. Typical Task Board Performance Data:
a) Cylinder Insertion (Upper Figure)
b) Box Insertion (Lower Figure)



TURNING & STORING RED SCREWDRIVER



TURNING BLUE WRENCH

Figure 13. Typical Task Board Performance Data:
a) Latch Turning with Flat Head Tool (Upper Figure)
b) Latch Turning with Square Head Tool (Lower Figure)

AD P001643

HELICOPTER INTEGRATED CONTROLLER RESEARCH

by
Sherry L. Dunbar
San Jose State University Foundation
San Jose, CA 95192

Abstract

The proposed research will examine the use of an integrated controller for the primary flight control of a helicopter. The primary controls of pitch, roll, and yaw will be integrated into a single, side-stick controller. Two types of integrated controllers will be studied: 1) an isotonic (displacement) controller; and 2) an isometric (force) controller. The two integrated control configurations will be compared to the conventional arrangement via performance on a 3-axis compensatory tracking task. RMS error will provide the dependent measure.

Introduction

Helicopters are being employed in an increasingly wide range of civil and military applications due to their unique flight capabilities. These are functions which fixed-wing aircraft either cannot do or cannot do very well. These functions include such things as low level contour flying, precision hovering, and landings in confined areas. These demanding maneuvers translate into high levels of pilot workload and fatigue. Advancements in helicopter technology will most likely result in increased automation of flight control functions and more on-board avionics which will provide some assistance in pilot decision making and navigation. Even with this added assistance, manual piloting will continue to be of utmost importance. Therefore, any improvements in helicopter cockpit or controls design will enhance the existing flight capabilities as well as maintain pilot workload and fatigue at a minimum.

In the present day helicopter, the control system consists of three primary elements; the cyclic, collective, and rudder (tail rotor). The cyclic is center-floor mounted and controls airspeed, pitch, and roll utilizing the right hand in a fore-aft and side-to-side motion. The collective is left-side floor mounted and directs changes in altitude, utilizing the left hand in an up-down motion. The rudder controls yaw movement, counteracting the effects of torque and controlling heading in hover manuevers, and is controlled by the feet in a pushing motion.

Although this represents the historical and current

configuration, it may not prove desirable, particularly in high demand flight situations. Both hands and feet are very active because of the frequent control activity required in these situations. In addition to the fact that the pilot is totally involved with the flying task, this activity is conducted from an awkward posture. The pilot is leaning left to actuate the collective, leaning forward for visibility, the arm is bent in front for the cyclic, and the knees are flexed to control the pedals; thus an uncomfortable posture to maintain for any length of time. A frequent complaint from helicopter pilots who have flown for a considerable time is severe back pain; the incidence of which is well documented. Other benefits resulting from the implementation of an integrated controller would be more efficient use of cockpit space, and a quicker escape from the cockpit if a situation should require it.

A need clearly presents itself to consider an alternative configuration in an attempt to alleviate the problems associated with this control arrangement. The proposed solution is an integrated controller, wherein the primary control functions (omitting altitude in the present case) are combined into a single side-stick controller. Two versions of the integrated controller will be compared with the conventional arrangement. One will be a moveable (isotonic) controller with output proportional to displacement, and the other will be a rigid (isometric) controller with output proportional to force. The rationale for considering two types of integrated control configurations is to gain information concerning both force and displacement characteristics to arrive at the optimal arrangement.

Some work has been conducted on the integrated controller concept. For instance, Boeing-Vertol has studied a variety of control types as well as stability and augmentation systems. Sikorsky has studied both displacement and force stick controllers with the addition of a secondary task. The object of the current research, and what has not been done to date, is to draw some conclusions about the use of conventional controls vs an integrated controller using objective performance measures, rather than subjective report.

It is assumed that use of a more efficient controller will result in improved tracking performance. This assumption will be tested by comparing performance between conventional controls and the integrated controller on a 3-axis compensatory tracking task.

The specific hypotheses to be tested are: 1)Tracking performance will be better with the integrated controllers than with conventional controls; 2) Tracking performance will be different with the isometric controller than with the isotonic integrated controller.

METHOD

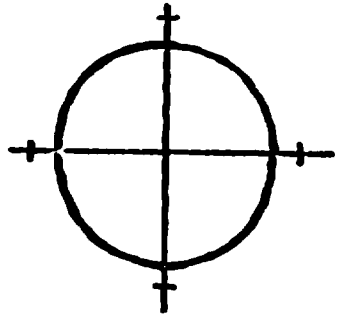
Subjects. Subjects will be non-helicopter pilots to eliminate effects of prior training. Subjects will be screened with the use of a pre- experimental tracking task to assure that they possess psycho-motor ability typical of that found in the pilot population. The pre-experimental tracking task will be Jex's (1966) "Critical Tracking Task" which proports to provide a low variability indicator of the operator's tracking ability. A matched group design will be employed with subjects matched on Jex Tracking Task performance, to guarantee some equality among groups. Subjects will then be randomly assigned into one of three experimental conditions; conventional control, isotonic integrated control, or isometric integrated control, resulting in three groups of six subjects each.

Design. A 3 X 3 mixed design will be employed with control configuration (conventional, isotonic, isometric) as the between factor, and task difficulty (low, medium, high) as the repeated measures factor. A randomized block design will be used with blocking on the Jex tracking task score.

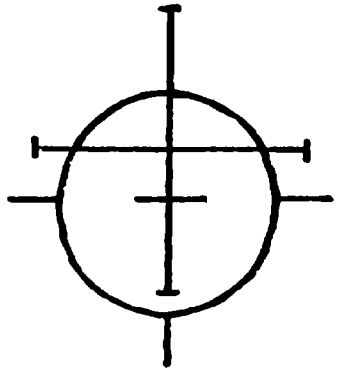
Apparatus. The pre-experimental (Jex) tracking task will be driven by an EAI 2000 computer, and displayed on a CRT. Subject input will be via a joystick mounted in a table and located directly in front of the scope.

The experimental station will resemble a helicopter cockpit. The station will consist of a platform upon which a seat, cyclic, collective, and rudder pedals typical of a helicopter will be mounted. The station will also be suited for the mounting of the integrated controllers on the right side of the seat. On the platform will be a 21-inch display oscilloscpe centered at eye level in front of the seat, and sloped back from the vertical by 4 degrees. A white noise generator and three first order filters will generate the input. The experimental task displays, recording of subject responses and analysis, will be accomplished by a DEC 11/34 computer.

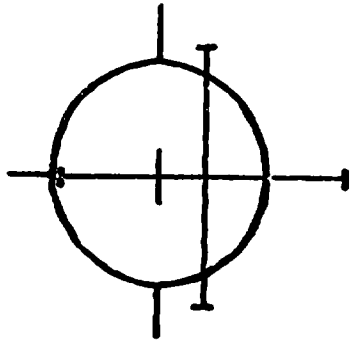
Procedure. The task will involve manipulation of the roll, pitch, and yaw axes, on a display like that shown in Figure 1. The object of the task will be to maintain the error at a minimum in all three axes using the appropriate controls by keeping the display as close to the null position (Figure 1a) as possible. The task requires the human operator to close three loops simultaneously as



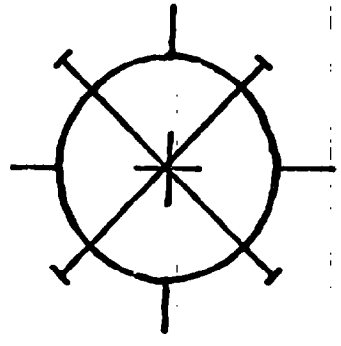
a) null



b) pitch



c) yaw



d) roll

FIGURE 1

illustrated in Figure 2. Subjects will by tracking a k/s plant in all axes.

In an off-line, non-linear helicopter simulation using a generic pilot model, Phatak (1980) demonstrated that system performance in the pitch axis will be severly degraded if the model remains open loop for seven seconds. Similarly, roll and yaw performance will be degraded in fifteen and twenty seconds respectively. This translates into a different amount of control activity required for the three axes. This relationship of control activity will be maintained in terms of the bandwidth of the forcing functions of the three axes. Specifically, the input forcing function will be filtered white noise, a separate level for each axis in accord with the frequency domain derived from Phatak's study. Then, three levels of the forcing function bandwidth will produce three levels of tracking difficulty while maintaining the bandwidth relationship among the axes. Figure 3 illustrates the bandwidth relationship among the axes and among the difficulty levels.

Error in the pitch axis will be displayed along the ordinate (Figure 1b); error in the yaw axis will be displayed along the abscissa (Figure 1c); and error in the roll axis will be displayed as angular deviation about the vertical (Figure 1d). Subjects will receive 36 two-minute trials per day, with a five-minute break after each set of nine trials. Subjects will participate until their individual performance reaches an asymptotic level.

Analysis. Data for analysis will be obtained from five sessions during which the asymptotic level has been maintained. Root Mean Squared (RMS) error will provide the dependent measure (Poulton, 1974). For a specific trial, the performance measure will be the sum of the mean RMS error scores from each axis. The trial error scores will then be averaged for each session (36 trials) providing an overall session performance measure. A 3 (control configuration) by 3 (difficulty level) analysis of variance will be performed on the data.

Under these controlled conditions, with a statistically well-defined task, the goal of the present research effort is to lay a sound foundation for further study and development of an optimal integrated controller.

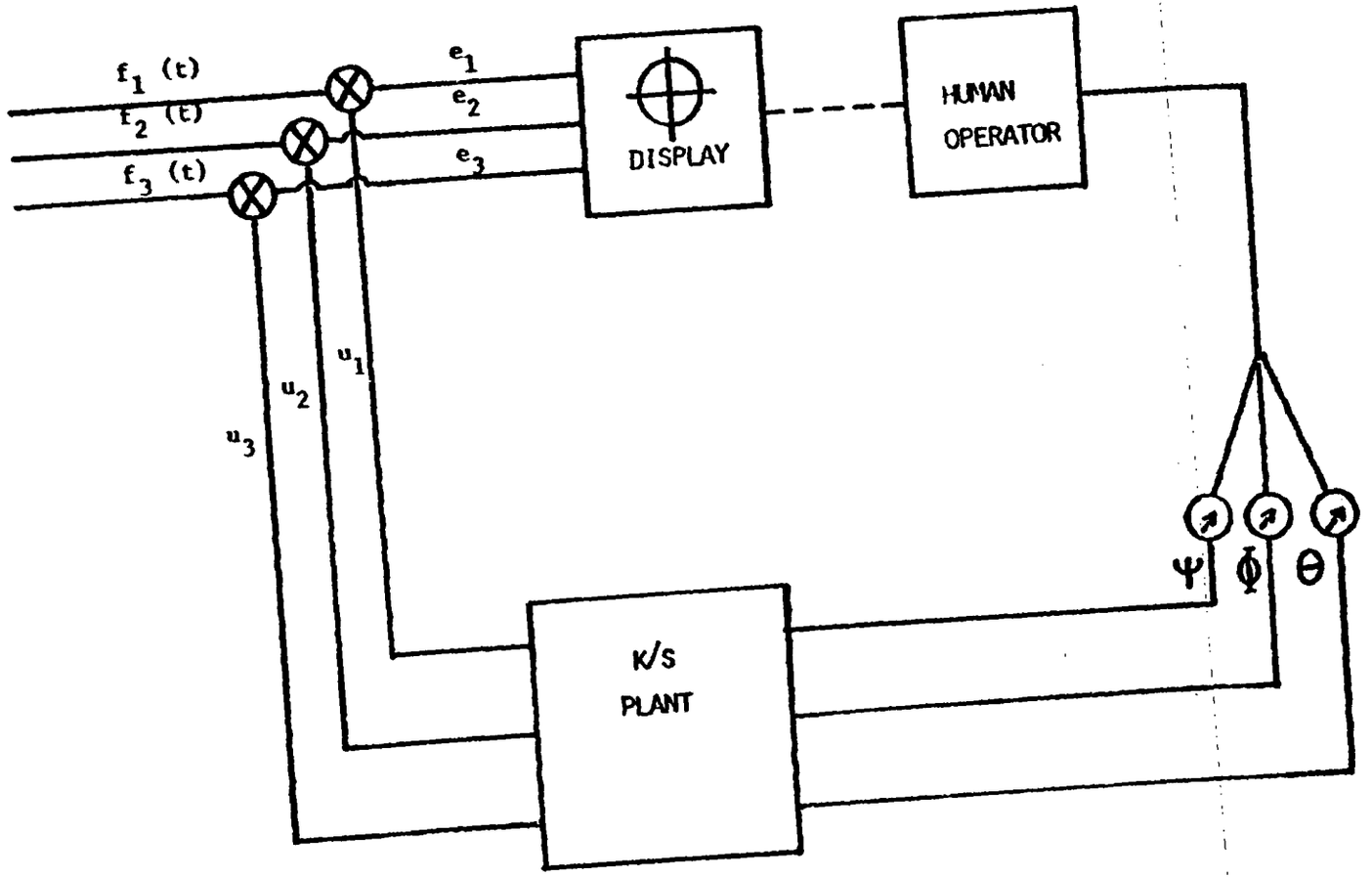

f_1 (t)
f_2 (t)
f_3 (t)
e_1
e_2
e_3
DISPLAY
HUMAN OPERATOR
u_1
u_2
u_3
K/S PLANT
Ψ
Φ
Θ

FIGURE 2

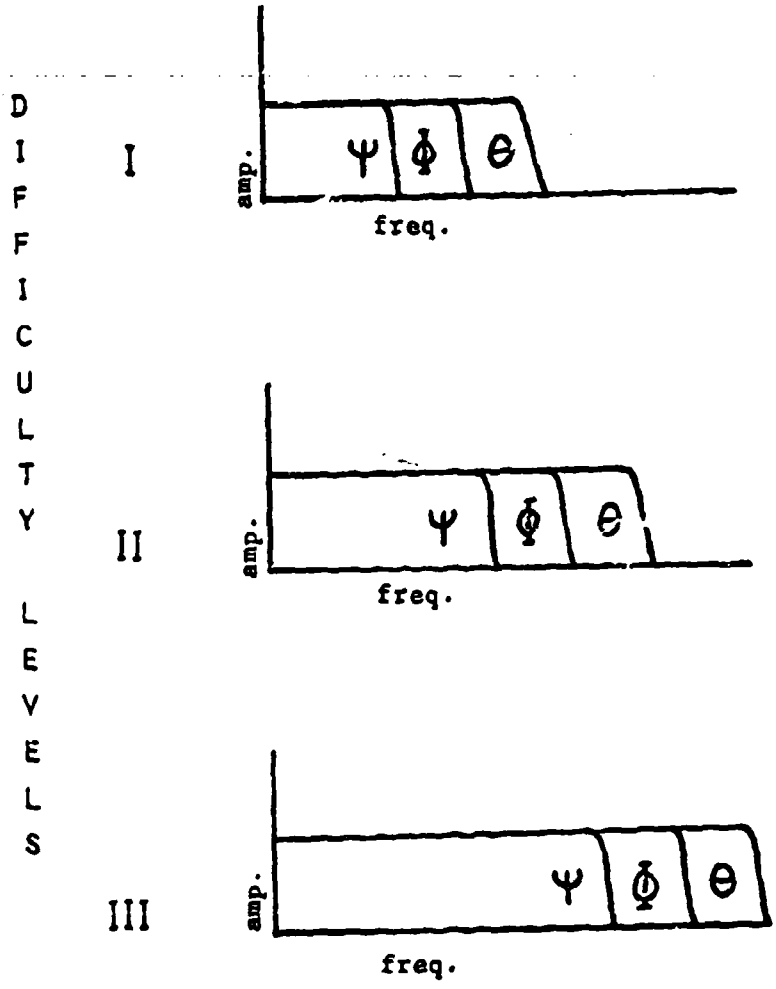


FIGURE 3

References

Jex, H. R., McDonnell, J. D., and Phatak, A. V. A "Critical" Tracking Task for Manual Control Research. IEEE Transactions on Human Factors in Electronics, 1966, 7, 138-145.

Phatak, A. V. Analytical Methodology for Determination of Helicopter IFR Precision Approach Requirements. NASA CR 152367, July 1980.

Poulton, E. C. Tracking Skill and Manual Control. New York: Harcourt, Brace, Jovanovich, 1974.

TENTATIVE CRITERIA FOR CONTROLLERS OF UNCOUPLED AIRCRAFT MOTION

JOHN M. EVANS
USAF

KEVIN CITURS
MCDONNELL AIRCRAFT CO

(PAPER NOT PROVIDED)

AD P001644

Application of Speech Synthesis Technology to Stinger Missile System

M. J. Crisp, Y. R. Yin and L. R. Perkin
GENERAL DYNAMICS, POMONA DIVISION

SUMMARY

Techniques for improved performance and reduced training cost of the U.S. Army/General Dynamics Stinger weapon system through improved man-machine interface and reduced operator work load have been evaluated. Two areas have been identified for possible improvement which include proper execution of operational procedures and recognition of audio cues. One of the possible solutions proposed is to use speech synthesis technology to provide the Stinger operator with a GO/NO GO type of voice cue. The voice module can also be used to correct operator procedural errors and/or cue the operator as to what function to perform next.

A prototype unit is currently under extensive evaluation by the U.S. Army users at Ft. Bliss, Texas. The research and development effort at General Dynamics, Pomona Division is continuing with emphasis on man-machine interface optimization including vocabulary selection, interface logic optimization, performance impact evaluation, power consumption and weight reduction, and optimal system packaging. Furthermore, the potential of this added voice capability on the man-portable weapon system with applications to command, control, and communications (C^3) interaction will be fully explored in the future.

INTRODUCTION

The effectiveness of any weapon system depends not only on the weapon itself, but also on how the weapon is used by its operator. This is especially true for a man-portable system like the U.S. Army Stinger surface-to-air guided missile. This shoulder launched, fire and forget weapon is a very effective and yet simple to operate system. There are only a few operating procedures to be executed properly in order to fire the missile round. However, humans do make errors which can cause a miss under normal field testing conditions, not to mention what can happen in a hostile battlefield environment.

Training helps to reduce operator error. Therefore, the U.S. Army has provided operators with several levels of training facilities from a portable Stinger trainer to a sophisticated half dome moving target

simulator. Most soldiers who complete this series of training exercise have demonstrated an operational effectiveness of the weapon close to its inherent effectiveness as a result of the training method.

But even the most extensive training cannot entirely eliminate the human related errors. Furthermore, the training cost is high and may be expected to go even higher in the future. Therefore, at General Dynamics, Pomona Division, a Stinger product improvement program has been underway to find ways to reduce the operator's work load and improve man-weapon interface. The ultimate goal is to minimize possible causes of human induced errors and also reduce the training cost.

Two areas have been identified for possible human factors/man-machine interface improvement. These are audio signal cues recognition and operating procedure execution. Among four different audio signals, the seeker acquisition tone discrimination requires skills that can only be acquired through tracking live targets. The high cost of operating live targets limits this type of tracking exercise to a minimum. Operating procedure execution, in general, does not constitute a major problem, but the idea of automated or even semi-automated operation of a man-portable weapon system is desirable.

PROBLEM IDENTIFICATION AND SOLUTION PROPOSITION

The normal sequence to launch the missile is shown in Figure 1.

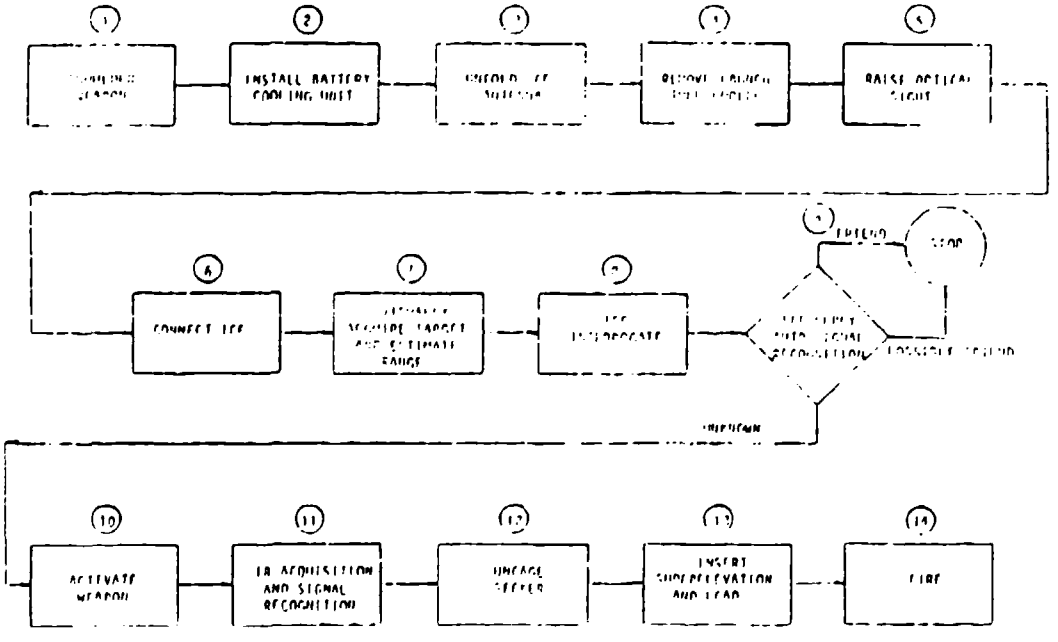


Fig.1 Launch Sequence of Events

There are roughly 14 steps in the launch sequence. Some of them are trivial like "shoulder weapon". Some of them are self-correcting, for example, if the operator does not follow Step 4 to remove the cover, his view would be blocked. Overall, Steps 1, 2, 3, 4, 5, 6, 10, and 14 are straight forward and in general contribute minimally to gunner errors. But steps 7, 9, 11 and 13 require skill, and hence training. Improper handling of any one or more of these procedures can cause the missile to miss the target. In particular, IFF (Identification of Friend or Foe) reply signals use three different audio tones to cue the operator with the information of a friend, a possible friend or an unknown target. The missile seeker acquisition signal also has to be recognized quickly by the operator to uncage the seeker head for continuing missile tracking. The reaction time is critical. In other words, the operator has to properly execute the procedure/sequence and firmly recognize the audio signal under any adverse battle field conditions in a very short time. Furthermore, there is neither additional direct cues regarding improper procedure execution nor verification of the correctness of the recognition of audio signals. It is this lack of communication between the operator and his weapon that prompted the U.S. Army Missile Command and General Dynamics to consider the addition of the digital voice capability to the Stinger missile system.

During the current phase, the task is to produce a field demonstration unit for further feasibility study. The design will strictly follow the guidelines of the product improvement program (PIP).

Fail Operable
Weight Conscious
Cost Effective

VOICE CUES SELECTION

Voice cue selection is a man-machine interface problem. There is, in general, no theoretical means to evaluate the effectiveness of the voice cue approach against the current audio tone method or one form of voice cue versus another. The ultimate judgement can only come from statistical results of field tests with real operators. Nevertheless, during the prototype design phase we have carefully studied many possible approaches which are feasible to implement. Four types of voice cues were singled out and described below.

(1) Echo/Confirmation: After the operator has performed a certain function, the voice module will respond with what has been done so that the operator will know whether he has followed the correct procedure. This is a leader and follower type of interface and perhaps the easiest for implementation. A major drawback with this approach is that the expected confirmation generally slows the gunner's reaction time.

(2) Error Correction: The voice module only responds when the operator

makes a mistake and is otherwise silent. This approach has the advantage of good reaction time. The voice module plus the control microprocessor plays a watchdog role.

(3) Command/Remind: This is a straight forward approach. The voice module will speak out what should be done next and will repeat the same message after a fixed time delay until the operator has performed the procedure. This is again roughly a leader and follower type of interface, but with the roles reversed from option (1).

(4) Warning: In case of anything unusual, the voice module will issue a warning to the operator. For example, when the battery voltage is low, the operator will be warned to replace the battery.

During the prototype design phase, a combination of all the above options was used except for the echo/confirmation approach. As part of the field test evaluation, more study will be done in the future so that the best form of man-weapon dialogue will be implemented.

Figure 2 depicts the voice cue logic sequence. The words within the double quatation marks are outputs from the voice modules.

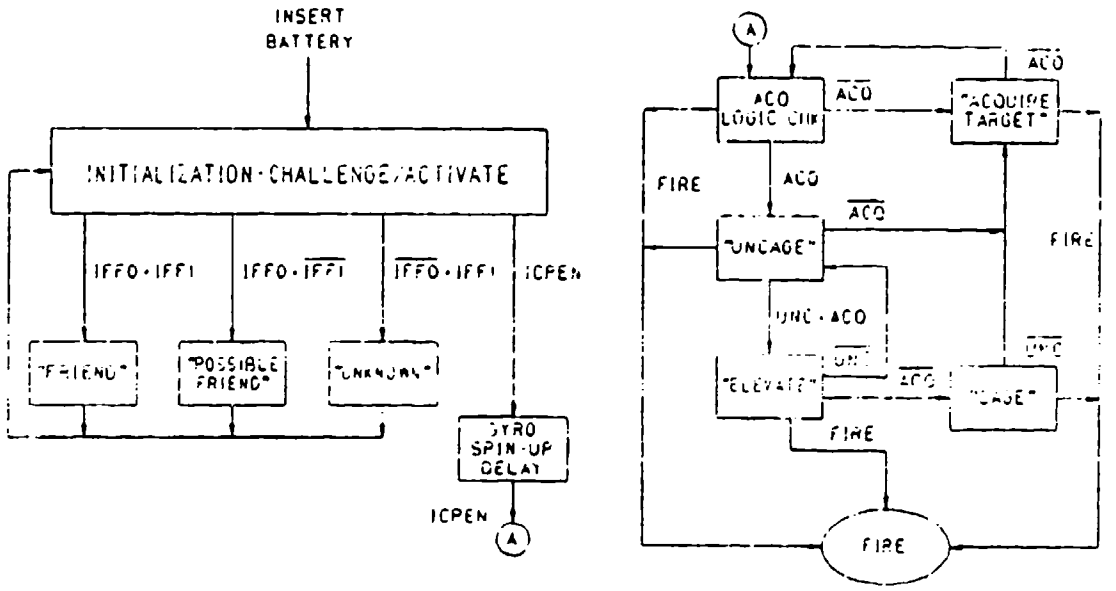


Fig.2 VOICE CUE LOGIC DIAGRAM

Basically, what has been done is to augment the modulated harmonic audio tones with the unambiguous voice in plain English to reduce the possible confusion in audio tone recognition. Furthermore, if a significant time delay occurs the voice module will remind the operator what is to be done next, correct him in case of procedural error, and finally warn him about anything unusual.

MAKING OF VOICE MODULE

Figure 3 depicts the interface between the voice module which consists of an Intel 8035 microprocessor and a General Instrument VSM2032 speech synthesis IC unit and various interacting components including the missile control electronics and the speaker.

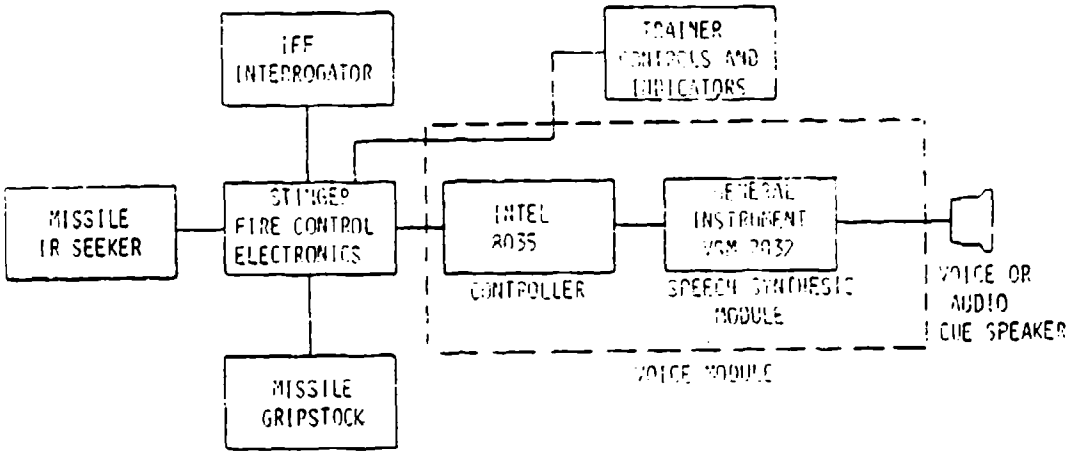


Fig.3 SYSTEM IMPLEMENTATION

Three kinds of speech synthesis chips or techniques are commercially available. These are Formant Synthesis (Phoneme/allophone), Linear Predictive Coding (LPC) and Wave Form Digitization. Speech synthesis IC units based on the first two methods have been available commercially for several years by quite a few chip manufacturers. The third one is a fairly new technique. Both the formant synthesis and LPC techniques units were considered;however we finally chose a LPC chip because of its high

quality human-like voice compared to the slight mechanical sound of formant synthesis chips. The need for large data storage is not critical in this application because only a limited number of vocabularies is currently required.

The whole module weights only 8.0 ozs and draws approximately 2.5 watts of power. The voice unit uses the same battery used by the missile launcher.

DISCUSSION AND CONCLUSION

The first demonstration unit was completed in October 1981 and was initially tested by General Dynamics engineers and training instructors. The results were good. The complete feasibility evaluation is presently underway by the U.S. Army operators.

As a tactical weapon, Stinger guided missile system's performance can best be described by its lethality and reaction time. Lethality is more a missile design problem although the proper execution of operating procedures is required to achieve the designed missile performance level. However, reaction time is as much a human factors problem as a design one for a man-portable system. It takes intelligence and experience to recognize the seeker lock on signal for a target masked by heavy background noise. The difference between a poorer operator and a better one can be either a long waiting period between action, missing the signal completely, or a clean shot. Since the voice cue or command is really a GO/NO GO type of signal, it has the potential to eliminate some uncertainty for some operators, and thus insures that all operators will be able to react within an acceptable time period. By retaining the modulated audio tones presently being used on Stinger, the fast reacting or more experienced gunners should not be adversely affected.

Through the addition of the voice module, it is believed that the goal of reduced operator's work load and minimal chance of human related error has been achieved. Because of these, hopefully the Army training costs can be reduced as well.

FUTURE WORK

The physical size, weight and power consumption of the voice module are acceptable for this phase. However, in the future, the Intel 8035 microprocessor will likely be replaced by a CMOS microprocessor to reduce the power consumption. The module size and weight will also be reduced using a General Instrument SP0256 speech synthesis chip and SPR16 memory ROM which replaces the VSM2032 module. All electronic circuitry and chips will then be optimally packaged as well.

The major future effort, other than the hardware change, will be the optimization of operator-weapon interface software. Vocabulary selection for easy audiobility, voice with different tempo, pitch and amplitude study and control logic optimization are all subjects for future work.

Meantime, the research and development effort is continuing to further explore the full potential of the voice capability to the Stinger weapon systems such as C^3 interaction.

ACKNOWLEDGEMENT

This work was supported in part by the U.S. Army Missile Command under contract No. DAAH01-81-C-A189. Here, the authors would also like to express their appreciation to Col. James E. Rambo for his many suggestions and comments.

SESSION 6: DISPLAY FACTORS

Moderator: Terry J. Emerson, Air Force Wright Aeronautical Laboratories

6

AD P001645

FACTORS AFFECTING IN-TRAIL FOLLOWING USING CDTI

by

James R. Kelly and David H. Williams
Flight Management Branch
Flight Control Systems Division
NASA Langley Research Center
Hampton, VA 23665

ABSTRACT

The Langley Research Center is participating in a joint NASA-FAA program designed to explore the potential benefits and liabilities of Cockpit Display of Traffic Information (CDTI) for a broad range of applications. As a part of this effort, piloted simulations have been conducted to determine the effect of various display parameters and separation criteria on terminal area in-trail following using CDTI. Both a conventional air carrier configuration, wherein the traffic data is presented on a centrally located weather radar scope, and an advanced configuration, where the data is superimposed on the primary navigation display, have been investigated. The scenarios employed have ranged from following a single lead aircraft to following multiple CDTI-equipped aircraft queues. This paper will summarize the results of these experiments and discuss their implications on the design and use of CDTI for in-trail separation.

INTRODUCTION

Cockpit Display of Traffic Information (CDTI) can, in the broadest sense, include devices ranging from a relatively simple panel mounted alphanumeric presentation of range and bearing to other aircraft, to pictorial electronic navigation displays, electronic primary flight displays and/or even head-up displays which incorporate traffic information. While any of these devices could be employed for self-spacing, this paper will focus on cathode-ray-tube (CRT) displays presenting a downward-looking picture (a plan view) of the traffic situation such as shown in figure 1. In addition, only the in-trail, self-spacing task will be discussed, although it is widely recognized that this is only one of many potential applications for CDTI.

This paper will summarize the results of some recent experiments conducted at the Langley Research Center as part of a joint NASA-FAA CDTI program. The effect of various factors, such as display size, and symbology will be discussed, along with self-spacing results obtained during piloted simulations using a variety of spacing criteria.

SYSTEM CONSIDERATIONS

System Elements

Conceptually, a CDTI can be divided into functional elements as illustrated in figure 2. There needs to be a sensor for detecting traffic, algorithms for processing the traffic, and a CRT (in this case) for displaying it. In implementing such a system, however, one is confronted with a multitude of choices and constraints which can affect the utilization of the CDTI. The following sections will address some of these issues and discuss the pros and cons of various choices based on the results of our research to date.

Traffic Sensor

The particular sensor used to obtain traffic information for CDTI will have a considerable influence on the accuracy, noise and, possibly, the update rate of the displayed traffic. While we have not specifically addressed all of these areas, we have examined the effect of noise on the displayed information for the in-trail following task.

The simulator study* was based on an aircraft-oriented, range-bearing device (such as a TCAS II) with the "noisy" traffic data displayed on a

*"Simulation Study of Traffic-Sensor Noise Effects on Utilization of Traffic Situation Display for Self-Spacing Task," by David H. Williams and Gene C. Moen, Proposed NASA TP.

centrally located CRT. The magnitude of the noise was parametrically varied by adjusting the standard deviation values for the range and bearing noise components. The results are shown in figure 3, and indicate the pilot acceptability boundary for range and bearing noise. This limit was based on the pilot's tolerance of the noise, as opposed to the spacing performance he could achieve. In fact, position errors, caused by sensor noise, with standard deviation values up to 0.3 n. mi. range and 8 degrees azimuth had a negligible effect on performance.

Traffic Selection

It is very unlikely that all of the traffic within a given area would be displayed on a CDTI simultaneously. To get an idea of what the CDTI might look like if it did portray all the traffic, consider the following: In developing the TCAS system, the FAA is using a traffic density of 0.4 aircraft per square nautical mile for the year 2000. This means that a CDTI with a 10-n. mi. radius could show up to 126 targets. In one of our early CDTI tests (ref. 1), display clutter was a major problem, even though no more than six other aircraft ever appeared on the screen at one time. Obviously, some of the traffic must be filtered out.

There are various filtering schemes that can be employed. Figure 4, which was generated during a contract study examining various CDTI traffic selection criteria*, shows the effect of altitude filtering in reducing the number of aircraft that would appear on a CDTI with a range of 10 n. mi. As suggested by Lockheed** intermediate altitudes (±1250, ±1750, etc.) should probably be employed for the filtering band (it may be asymmetrical) to keep aircraft at "even" altitudes, such as 4000, 4500, and 5000 feet, from intermittently tripping the threshold.

Another option for selecting traffic is to employ so-called TAU filtering. The time constant TAU is obtained by dividing the traffic's range by its range rate relative to ownship. As illustrated in figure 5, the TAU criteria eliminates most of the traffic, even at a fairly long time constant of 4 minutes.

Although any target selection criteria is going to involve compromise, it seems that limiting the number of aircraft displayed will not be too difficult. There is a question of how to treat a "pop-up" target, however. This is an aircraft which is within the horizontal range covered by the display, but had been filtered out using some criteria such as altitude. Once the target comes within the selection criteria, some appropriate method of bringing it up on the display must be employed.

*Work done under contract to NASA by the Systems and Measurements Division of Research Triangle Institute.

**From work done under contract to NASA by Lockheed-California Company.

DISPLAY CONSIDERATIONS

Location

Our simulator studies to date have employed both conventional and advanced cockpit configurations, as shown in figures 6, and 7, respectively. In the conventional cockpit configuration, the traffic information is presented on a single CRT located in the center console where a weather radar is frequently located. In the advanced configuration, the traffic information is presented on the navigation displays located on the captain and first officer's primary display panels.

Although we have not conducted any studies aimed directly at location effects, there is evidence from several studies which indicate that when the surrounding traffic is updated discretely (typically every 4 seconds), the pilots wait for an update while scanning the CDTI. During the study reported in reference 2, for example, where the CDTI was centrally located, one of the test subjects commented that the physical location of the display caused him to "spend too much time away from his primary flight displays." (This update rate/location problem had not been noted by the same pilot when he flew the advanced cockpit configuration.) One possible solution for this problem is to predict the traffic's position between sensor updates, and update the CDTI continuously using estimated positions.

Hand-in-hand with location is whether there are one or two CDTI's in the cockpit. Generally, if it's centrally located, there will only be one; if the CDTI is integrated into the primary flight displays, there may be two of them. In all of our in-trail following studies, we have found that most pilots prefer to use the most sensitive map scale feasible. With two CDTI's, as in the advanced cockpit configuration, one pilot can use the sensitive scale for self-spacing, while the other pilot uses a less sensitive scale for monitoring the traffic situation. With a single centrally located CDTI, either a compromise map scale setting or a time-sharing scheme has to be utilized.

Display Size

One of the simulation studies at Langley was directed specifically at determining the effect of display size on the self-spacing task (ref. 2). This study was conducted in the conventional aircraft simulator (fig. 6), using the CDTI format shown in figure 8. The important feature of this display, with respect to self spacing, is that pilot selectable data blocks were available which contained traffic groundspeed. This allowed the pilot to set up a comfortable overtake velocity on the lead aircraft for rendezvous, and then maintain the prescribed spacing interval by matching

the lead aircraft's groundspeed. This technique tended to diminish the importance of using a sensitive map scale for accurately determining ownship's position relative to the target. Although the standard deviation of the spacing performance shows a consistent trend toward lower values (improved performance) with larger displays, the mean spacing performance shows only a small (±0.06 n. mi.) random variation. Display size, therefore, may not be too critical if appropriate information is contained in the traffic data blocks.

Display Symbology

The general display format shown in figure 8 has been employed in all of our in-trail following studies to date. The main features include ownship oriented map information, coded traffic information, traffic trend vectors, traffic history trails, and traffic data blocks. The major variation in the symbology between studies has been the type of spacing cue employed. The straight trend vector emanating from ownship in figure 8 illustrates the constant distance (CD) spacing cue used in reference 2. Two range arcs cross the trend vector at 3 and 5 n. mi., which were the spacing requirements employed in the study.

The next figure (9) shows the symbology used in an investigation* of time-based self-spacing techniques. In this study, two different spacing cues were employed; part (a) shows a constant time delay (CTD) configuration where the lead aircraft leaves an 80-second history trail composed of 20, 4-second past position dots. The purpose of the long history trail was to provide a path for own aircraft to follow when the lead was not flying a defined navigation route. The last position dot includes a line and a data tag to indicate where the lead aircraft was 80 seconds ago (the spacing criteria employed), and what its groundspeed and altitude were. Ownship has a constant time predictor trend vector in this case, which indicates where ownship will be in 80 seconds if the pilot maintains his present groundspeed.

This constant time predictor (CTP) is actually another type of spacing cue, as illustrated in figure 9(b). Here the line and data block have been deleted from the lead aircraft's trail and an arc has been added to ownship's time predictor to emphasize the 80-second prediction point.

It should be noted that these last two formats used a curved trend vector, which predicted ownship's position, even during turns.

---

*"Preliminary Evaluation of Time-Based Self-Spacing Techniques During Approach to Landing in a Terminal Area Vectoring Environment Utilizing a Cockpit Traffic Display," by David H. Williams, Proposed NASA TM.

The three types of spacing cues illustrated have also been employed in the advanced cockpit simulator. Figure 10 illustrates one advanced cockpit format which combines two criteria, the CTP and CD. The traffic in this case is not coded (since it was assumed that the selection criteria would eliminate all incidental aircraft) and only has a groundspeed tag. Ownships trend vector is composed of 30, 60, and 90 second time predicting segments, and a range tick which remains 3 n. mi. ahead of ownship centered on the trend vector.

The type of symbology used to represent the lead aircraft has been found to be of little importance for the in-trail following task. The primary use of the traffic symbol in this case is to provide a position reference for the spacing cue. For other uses of CDTI, however, the traffic symbology utilized can be an important consideration. In particular, for general situational awareness and CDTI-aided visual acquisition of traffic, coded traffic symbology has been found to be a significant factor.

The first study (ref. 1) conducted at Langley under the joint NASA/FAA CDTI program was a flight evaluation of the coded symbology proposed in reference 5. These tests were directed primarily toward situational awareness, and were conducted in the aft flight deck (fig. 7) of the TCV aircraft using the CDTI format shown in figure 11. The traffic was displayed both with and without coded symbology which distinguish the traffic's altitude relative to ownship, whether they were under ATC control, and whether or not they were CDTI equipped. The traffic symbols also included trend vectors, past position trails, and pilot selectable data blocks giving aircraft identification, altitude, and groundspeed.

The most advantageous coding of the symbology in this study was found to be the altitude. With coding, the pilot could eliminate the data blocks until he needed additional information on a proximate aircraft. Without altitude coding, the pilot had to call up the data blocks more frequently, aggravating a display clutter problem. It should be noted that for in-trail following tasks, the groundspeed information contained in the data tag has been found to be an important piece of information. It would, therefore, appear advantageous to separate groundspeed and altitude data tags rather than including them in the same data block.

Another indication of advantageous coding of traffic symbols was an observation made during a flight test involving CDTI*. The purpose of this study was to evaluate CDTI as an aid for visually locating traffic. The display format was a simple plan position presentation with range rings and single numbers to represent the traffic. The numbers were obtained by summing the last four bits of the transponder identification code of each aircraft, thus providing different numbers for different aircraft. This simple encoding of aircraft identification in the traffic symbols was found to be very beneficial. The pilots found they could easily track a particular aircraft with short glances at the display even in the high density Los Angeles basin with as many as 12 aircraft displayed at one time.

*Under contract to NASA by Lockheed-California Company.

An important point to be made about display symbology, and display considerations in general, is that the type and amount of information presented on the display is highly dependent on the task being performed using the CDTI. In-trail following tasks require spacing information not necessary for other CDTI applications. Similarly, the other applications may require information not necessary or appropriate for in-trail following. An optimal CDTI display must, therefore, be flexible enough to present the minimum necessary information in the form appropriate for a given application.

SPACING CRITERIA

Basic Types

Figure 12 illustrates the three basic spacing criteria employed in our studies; namely, the constant distance criteria (CD), the constant time predictor criteria (CTP), and the constant time delay (CTD) criteria. The figure also gives the equations for the spacing intervals in terms of ownship's groundspeed, V_o, and the lead aircraft's groundspeed, V_L, and time predictor and time delay time constants, τ_p and τ_d. In practice, the pilot satisfies these criteria by keeping the appropriate cue superimposed on the lead aircraft symbol (CD and CTP), or on his own aircraft symbol (CTD).

Advantages--Disadvantages

The advantages and disadvantages of these spacing criteria are not always immediately apparent. Consider two aircraft flying a descending, decelerating approach to a landing. Using the constant distance criteria the follower must match the lead aircraft's speed profile all the way down the approach. This is illustrated in figure 13, which shows the lead aircraft's groundspeed, and the groundspeed which must be flown by the follower, as a function of distance to the threshold for a 3-n. mi. separation requirement. (The ideal groundspeed for several different spacing criteria is presented in reference 4 as a function of the lead aircraft's groundspeed.) It is immediately apparent that this criteria forces the follower to slow up too far out. In addition, since gear and flap schedules are generally tied to specific speeds, this means that if the lead aircraft extended gear and flaps at an optimum point, then the follower would have to extend them too early (although at the correct speed) on the approach.

Application of the constant time predictor criteria yields a gradual reduction in spacing as the aircraft slow up coming down the approach. This reduction in spacing represents what is typically done in today's ATC system and, hence, this criteria has the advantage of being somewhat familiar to pilots and controllers. Another advantage of the CTP criteria is

that it tends to smooth out speed variations of the lead aircraft. This feature is illustrated in figure 14, which shows a follower using CTP spacing on the same lead aircraft as used in the previous figure.

One of the disadvantages of the CTP criteria, somewhat apparent on this figure, becomes obvious when a string of aircraft are following each other using the CTP. The problem is that the CTP criteria causes the follower to fly slower than the lead aircraft during the deceleration phase of the approach and then faster than the lead when the lead is no longer decelerating. In order to maintain the desired time interval spacing, the follower is thus required to fly a faster final approach speed than the lead aircraft. For a string of aircraft using CTP spacing, each successive aircraft would thus be required to fly progressively faster approach speeds than the aircraft preceding it. This situation is operationally unacceptable, since each aircraft must slow to specified final approach speeds as dictated by the landing procedures for the aircraft. This operational constraint results in a slow-down effect where each aircraft takes longer to fly the approach than the preceding aircraft.

The constant time delay criteria also yields a gradual reduction in spacing as the aircraft decelerate down the approach. The main advantage to CTD is that it allows the follower to fly the same speed profile as the lead aircraft (fig. 15), thus avoiding the slow-down effect present with CTP spacing. Of course, CTD spacing would only be appropriate for similar types of aircraft. A possible disadvantage to CTD, as implemented in our studies, is the narrowing of pilot attention to the area of the display directly surrounding the ownship symbol. As a consequence, the pilot will possibly lose some of the predictive information concerning path changes of the lead aircraft, which may result in some undesirable lateral errors during turning maneuvers. This situation could be improved by projecting the spacing cue for CTD further up on the display allowing the pilot to better monitor the ground track of the lead aircraft.

OPERATIONAL FACTORS

Pilot Technique

In most of our investigations to date, we have observed that the pilots appear to adopt their own strategy in satisfying the spacing criteria. Some pilots like to close up (rendezvous) on a lead aircraft much slower than others. Some pilots attempt to hold position accurately, while others accept much more variation in their relative position. In addition, the pilots also tend to hold various position biases. However, in most cases they hold a positive position bias; that is, they almost always maintain more separation than that specified.

While our studies have indicated that different pilots adopt different techniques, the studies have also shown that the pilots are very consistent when it comes to applying their own particular technique. Figure 16 provides a graphic illustration of this point. The figure shows the mean interarrival times over the outer marker achieved by four different pilots during a recent study.* The study was conducted in the advanced cockpit simulator and involved in-trail, self-spacing on a lead aircraft performing a profile descent to an ILS approach and landing. The test matrix was based on a full-factorial, analysis of variance design and, thus, it was possible to obtain the effect means for each test subject. The two sets of data points represent the effect means from two test series separated by 4 months.

Both of the data sets in figure 15 illustrate the variation between pilots in arriving at the outer marker. By comparing the two curves, however, one can see that the pilot's performance was relatively consistent between the two test series which, as noted, were 4 months apart. Another noticeable feature of the data is that all the pilots maintained a greater separation during the second test series than they did during the first one. This could be a learning curve effect, since the pilots tended to get too close to the lead aircraft during the first simulation series.

Groundspeed

Self-spacing results indicate that the groundspeed of the lead aircraft is an important piece of information. When spacing is via the constant distance criteria, the pilots can match their own groundspeed to the lead aircraft's groundspeed on a knot-for-knot basis. The same matching can be applied with the constant time delay criteria if the lead aircraft's previous groundspeed is shown. Even in the case of the constant time predictor criteria, where the pilots cannot match the lead aircraft's speed, the ground speed tag aids the pilot in determining when the lead aircraft is decelerating and how fast.

Given that a CDTI will display the groundspeed of the lead aircraft (and possibly other aircraft as well), then the question arises as to how accurately this information needs to be displayed. The groundspeed study referred to earlier was directed specifically at answering this question. In addition, the study of time-based self-spacing techniques, conducted in the conventional cockpit simulator, and mentioned previously in the display

---

*"The Effect of Lead Aircraft Groundspeed Resolution on Self-Spacing Performance Using a CDTI," by James R. Kelly, proposed NASA TP.

symbology section, can also shed some light on this area, since the tests included approaches with and without groundspeed information.

Figure 17 presents the effect means for the IATS at the threshold as a function of groundspeed resolution. This data comes from the same set used to illustrate pilot technique in the previous section. Again, two sets of data are shown, one from each of the two test series conducted.

It is immediately obvious that the separation between the aircraft tends to be less (lower IATS) at coarse resolution levels compared with fine ones. It is also apparent that during the first test series the mean IATS at the threshold for the 10- and 20-knot resolution cases were less than the desired separation. During the second test series, however, the pilots employed a spacing buffer such that even though they tended to be closer to the lead aircraft at coarse resolutions, they typically crossed the threshold at or above the specified separation.

Based on our observations to date, it appears that coarse groundspeed readouts, or no readout whatsoever, mask the lead aircraft's decelerations, causing the pilots to inadvertently lose separation. Once separation is "lost" it is very difficult to increase it, especially when the lead aircraft is continuously decelerating, or when ownship has reached final approach speed. With fine resolutions, on the other hand, the pilots can readily detect when the lead aircraft starts to decelerate and take appropriate action to maintain separation.

The results from the conventional cockpit simulator further reinforce the influence of target groundspeed knowledge on the pilot's spacing performance. The mean IAT at the runway threshold was found to be significantly less (3½ sec) without a groundspeed tag, compared with a lead aircraft with a 10-knot resolution tag. In addition, the spacing strategy of the pilots using CD spacing criteria was significantly affected by the presence of the target groundspeed tag. Without the groundspeed information, the pilots would hold back and maintain greater spacing as they turned onto final approach. Despite this added spacing buffer, the pilots would still overshoot the desired spacing when the target decelerated, and the spacing when the target crossed runway threshold would be too close. Once again, a given pilot exhibited highly repeatable performance, with significant differences noted between pilots.

Lead Aircraft Characteristics

One of the primary factors which influences the pilot's self-spacing performance has been found to be the behavior of the lead aircraft (it's not unlike formation flying where a smooth leader makes the wingman's job easier). At one end of the spectrum is the case where the lead aircraft is flying at

constant speed and constant altitude. Under these conditions the pilots can achieve very consistent spacing with standard deviations of less than 0.1 n. mi. (refs. 5 and 6), which equate to an interarrival time standard deviation of less than 3 seconds.

As the lead aircraft begins to maneuver (decelerate), even mildly, the self-spacing performance as reflected by the standard deviations is adversely affected. For example, reference 2 presents spacing data for two constant distance spacing segments during which the lead aircraft is performing a very docile approach. The standard deviations ranged between 0.13 and 0.28 n. mi., depending on the display size employed. These deviations are equivalent to interarrival times between about 3 to 7 seconds. Another study using a similar task also produced standard deviations in the same 3 to 7 second range.

The most recent studies conducted at Langley have involved relatively active lead aircraft. In the study of time-based self-spacing techniques, the simulated lead aircraft was being vectored by ATC onto the final approach. In the groundspeed resolution study, the lead aircraft was performing an idle thrust profile descent with a close in turn to final. In both of these studies the standard deviation of the interarrival times was found to be 7 to 8 seconds. Figure 18 summarizes our experience to date in terms of the lead aircraft's effects on self spacing performance. Keep in mind that these are only "ball park" numbers; the test results from which they are derived have invariably involved lead-followers of the same type aircraft. This probably leads to better spacing performance than could be obtained with pairs of different types, which is to say different performance characteristics. On the other hand, the data used in deriving the "ball park" numbers have not been massaged to eliminate particularly poor performance producing constraints such as not having target groundspeed information. As such, the data base is not necessarily biased, since there are offsetting features incorporated in it.

Potential Payoff

Assuming it is possible to achieve a standard deviation of 8 seconds for interarrival time at the threshold using CDTI in the real world, then we can estimate the potential payoff with the aid of figure 19.

This figure illustrates the probability density curves for interarrival times using two different spacing techniques. The standard deviation value of 18 seconds used for the curve labeled "ground controlled" was taken from reference 7 and represents the accuracy of today's ATC system. The mean for the curve in figure 19 has been established by applying a ground rule that only 5 percent of the arrivals can have an IAT less than 50 seconds. (This actually results in a mean of about 79.6 seconds.) The basis for selecting a specific IAT and buffer size are explained in reference 7. In this paper, the somewhat arbitrary choice of 50 seconds and 5 percent will serve as an anchor point for comparing CDTI to present day techniques.

If we apply the 50-second, 5-percent ground rule to our "worst-case" CDTI spacing results (1σ = 8 seconds), the resulting mean is found to be about 63 seconds. This means that arriving aircraft could be spaced 17 seconds closer together which, in turn, could up runway throughput 10 to 25 percent or more, depending on the mean IAT in effect.

In addition, CDTI has occasionally been referred to in terms of being a "tail-cutter" of the IAT distributions. What this means is that, whereas a buffer is incorporated in the present ground controlled system to allow for the occasional 3 or 4σ IAT case, CDTI may eliminate such cases entirely and thus reduce the requirements for a buffer. It has certainly been our experience that the pilots are continuously aware of their separation from the lead aircraft and could, if directed, take some positive action (such as a go-around) to eliminate an excessive IAT error. In any event, the pilot using CDTI for in-trail following is always aware of his spacing and would probably anticipate an intervention by an ATC controller if the situation became critical. This anticipation feature could possibly lead to a reduction in the buffer size also.

Multiple Aircraft Queues

One final aspect of the in-trail following task this paper will address deals with multiple aircraft queues (sometimes referred to as a "daisy-chain"). There has been some concern that a queue of CDTI-equipped aircraft, all of which are self-spacing on their respective lead aircraft, would lead to dynamic oscillations similar to the accordian effect seen with a queue of automobiles in stop-and-go traffic. In order to examine this potential problem, we utilized the CDTI-equipped advanced cockpit simulator, and built a queue of aircraft by recording successive self-spacing approaches and appending them to a common traffic tape. The first approach flown was a profile descent to landing without any traffic. The second approach included self-spacing on the lead aircraft while flying the profile descent. For each successive approach, ownship was added to the end of the queue and self-spaced on the preceding aircraft. Two aircraft crews were utilized for these tests in which the pilots alternated captain and first officer duties, and the crews flew alternate runs. As such, the pilots had no a-priori knowledge of the lead aircraft they would be following.

Using the procedure described above, queues of up to nine aircraft were built during the simulation sessions. Figure 20 is a photograph showing a case where ownship is the ninth aircraft in the string. The photo was taken during a session using a 60-second constant time predictor spacing criteria. Other sessions were conducted using a 60-second constant time delay criteria. The results of these tests are shown in figure 2 in the form of aircraft groundspeed as a function of distance to the threshold: Only the odd numbered aircraft are shown to avoid clutter.

As seen from the figure, there is no evidence of any oscillatory tendencies with either spacing criteria. However, the slow down effect associated with the CTP criteria and discussed in a previous section, is readily apparent. Whereas the first aircraft took about 12 minutes to fly the approach, the ninth aircraft, starting from the same initial condition, took about 13½ minutes to fly the same approach. In the case of the CTD criteria, on the other hand, there is no slow down effect and, thus, all the aircraft fly the approach in about the same amount of time.

CONCLUDING REMARKS

This paper has attempted to explore some of the factors associated with in-trail following using CDTI. We indicated that one of the primary factors affecting self-spacing performance is the behavior of the lead aircraft. Due to this finding, there is an obvious need for data dealing with self-spacing between dissimilar aircraft types.

Knowledge of the lead aircraft's groundspeed has been found to be a significant factor in self-spacing performance also. Without this information, the pilot has a tendency to overshoot the desired spacing and get too close to the lead aircraft. Providing groundspeed information leads to more accurate spacing, and tends to compensate for a less than desired display sensitivity in the case of small displays.

Test subjects used in our simulation studies were found to develop individual techniques for performing in-trail following tasks with a CDTI. In some cases, these techniques were significantly different between pilots; however, the individual pilots exhibited consistency in applying their own techniques.

Small levels of traffic sensor noise, as exhibited by noise in the displayed traffic positions, were found to be more of a pilot acceptance problem than a spacing performance problem. Pilots objected to noise levels greater than approximately 2^{0} azimuth and 0.1 nautical mile range; however, their spacing performance was unaffected with noise levels of 8^{0} azimuth and 0.3 nautical mile range.

We have also shown that self-spacing within CDTI-equipped aircraft queues do not appear to be oscillatory, as had been previously hypothesized. Here again, though, these results are based on limited data with similar aircraft types following one another.

In summary, we have learned a great deal about CDTI, in general, and about in-trail following, in particular. In spite of this, however, numerous questions still need to be answered.

REFERENCES

1. Abbott, Terence S.; Moen, Gene C.; Person, Lee H., Jr.; Keyser, Gerald C., Jr.; Yenni, Kenneth R.; and Garren, John F., Jr.: Flight Investigation of Cockpit Displayed Traffic Information Utilizing Coded Symbology in an Advaned Operational Environment. NASA TP 1684 (AVRADCOM Technical Report 80-B-4), July 1980.

2. Abbott, Terence S.; and Moen, Gene C.: Effect of Display Size on Utilization of Traffic Situation Display for Self-Spacing Task. NASA Technical Paper 1885, August 1981.

3. Hart, Sandra G.; and Wempe, Thomas E.: Cockpit Display of Traffic Information: Airline Pilots' Opinions About Content, Symbology, and Format. NASA TM-78601, 1979.

4. Sorensen, John A.; and Goka, Tsuyoshi: Analysis of In-Trail Following Dynamics of CDTI-Equipped Aircraft. AMA Report No. 81-9, Analytical Mechanics Associates, Inc., June 1981.

5. Imrich, T.: Concept Development and Evaluation of Airborne Traffic Displays, MIT Flight Transportation and Laboratory, FTL-R71-2, June 1971.

6. Kelly, James R.: Preliminary Evaluation of Time and Distance Spacing Cues Using a Cockpit Displayed Target. NASA Technical Memorandum 81794, August 1980.

7. Swedish, William J.: Evaluation of the Potential for Reduced Longitudinal Spacing on Final Approach. Report No. FAA-EM-79-7, August 1979.

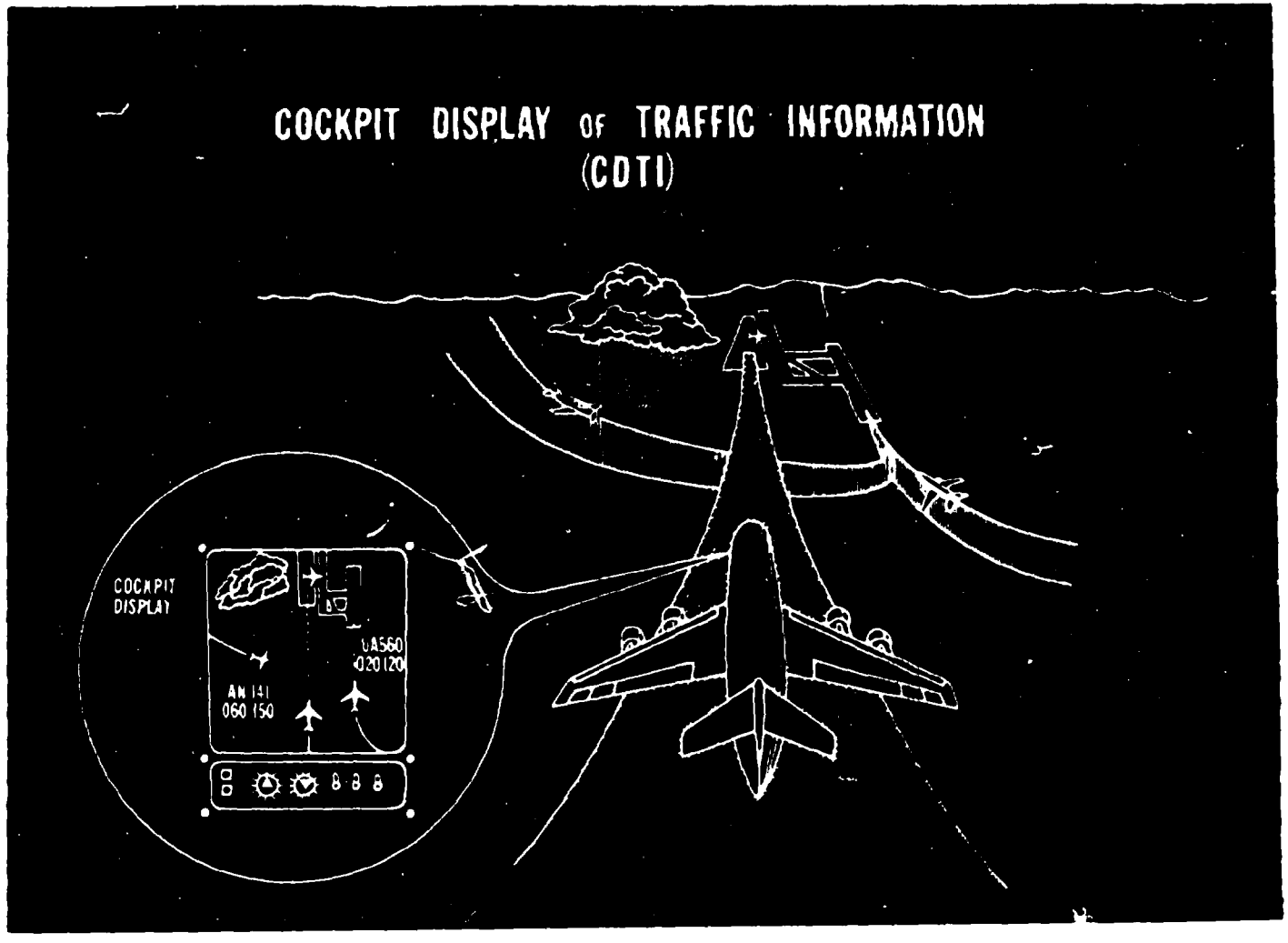


Figure 1.- CDTI concept.

Figure 2.- CDTI system elements.

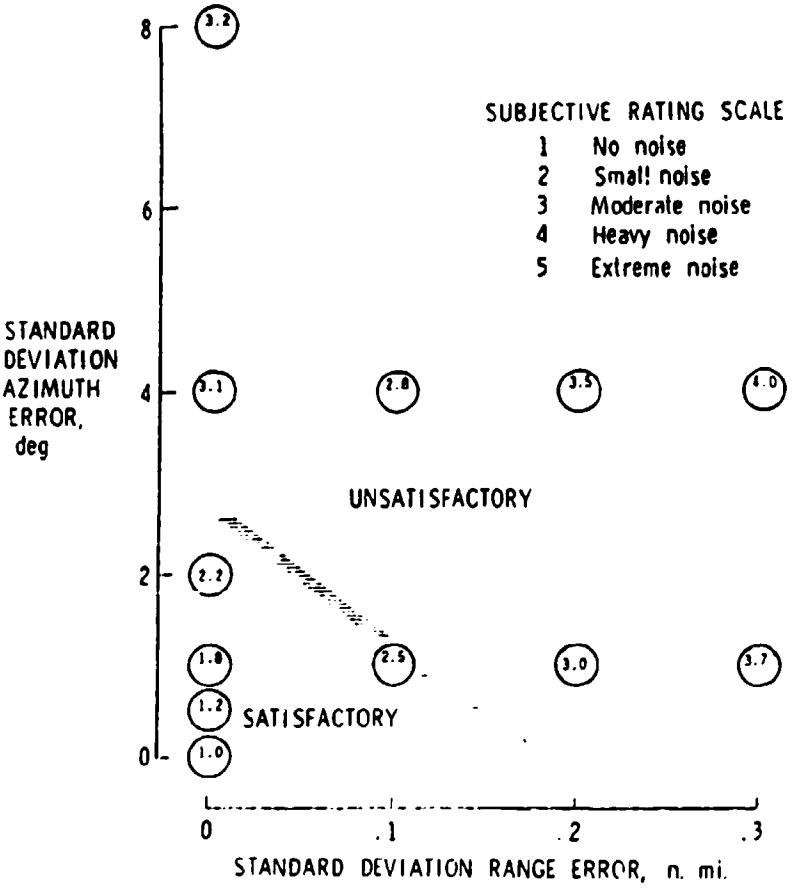


Figure 3.- Pilot rating of range and bearing noise.

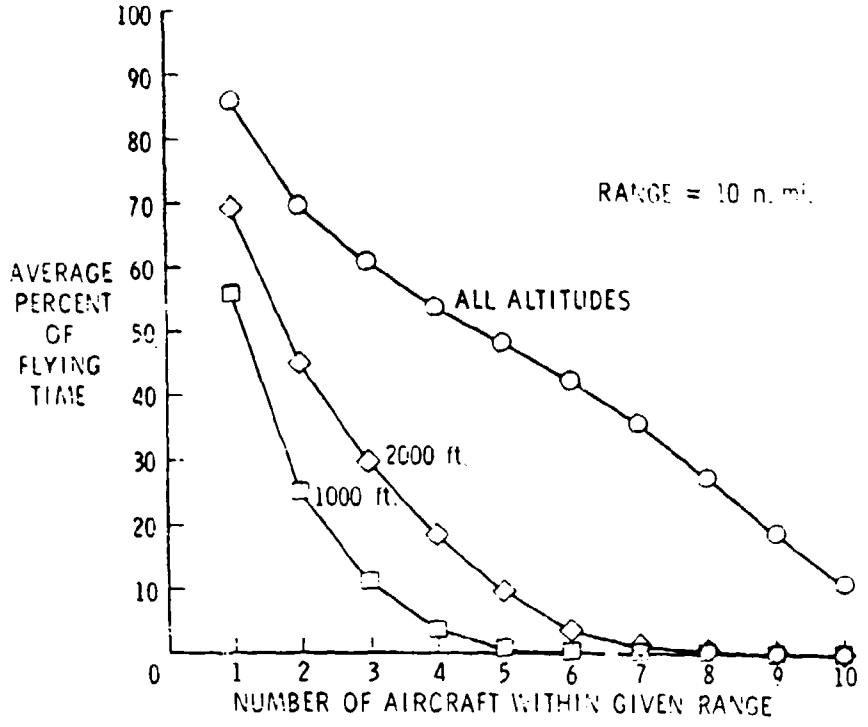


Figure 4.- Effect of altitude filtering on number of aircraft displayed.

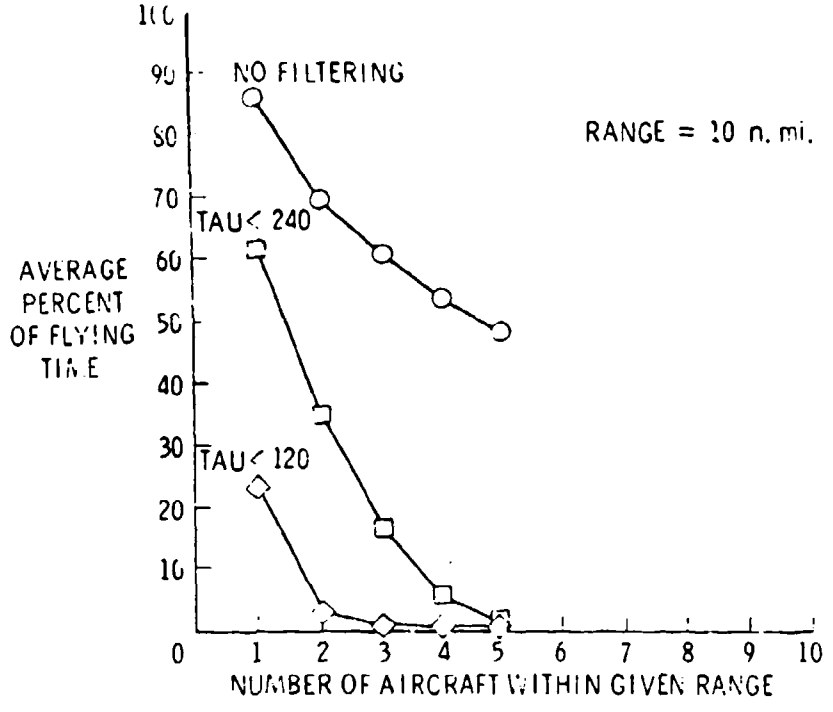


Figure 5.- Effect of tau filtering on number of aircraft displayed.

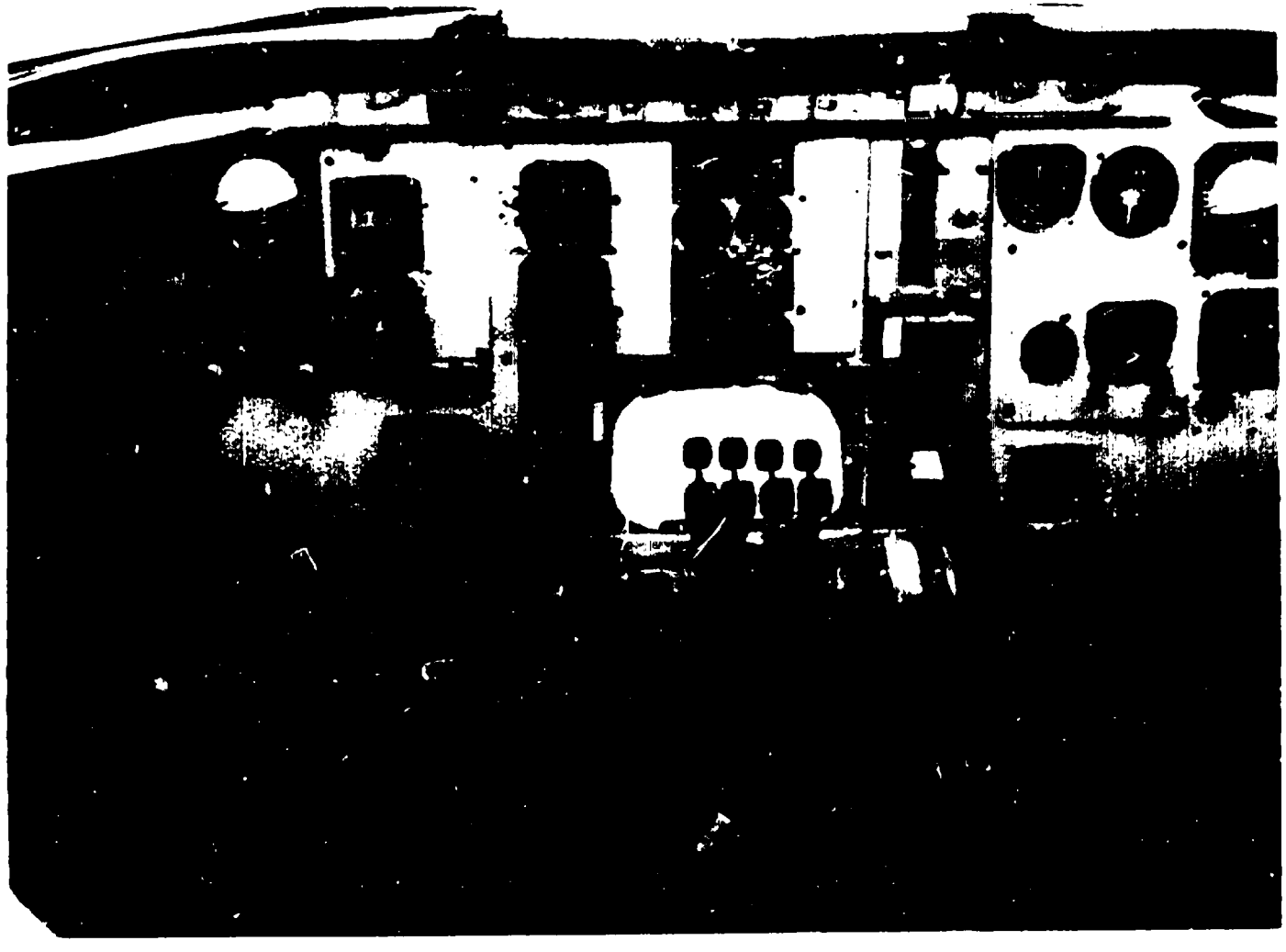
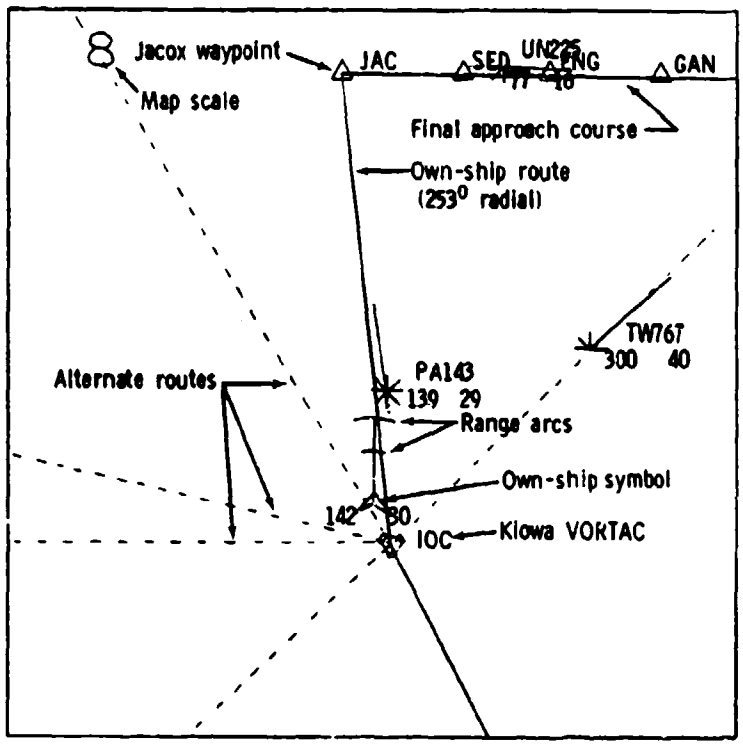
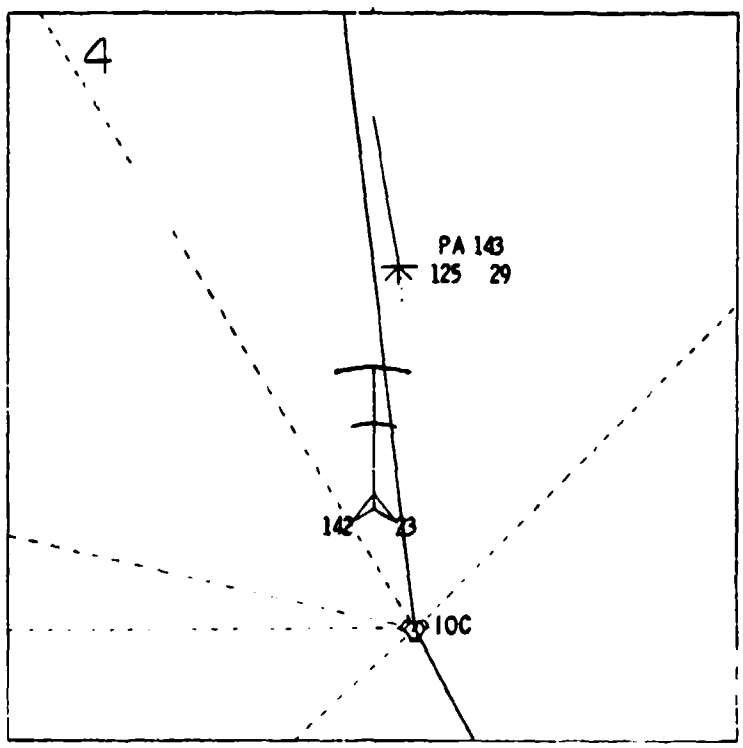


Figure 6.- Conventional cockpit simulator.

Figure 7.- Advanced cockpit (TCV aft-flight-deck) simulator.

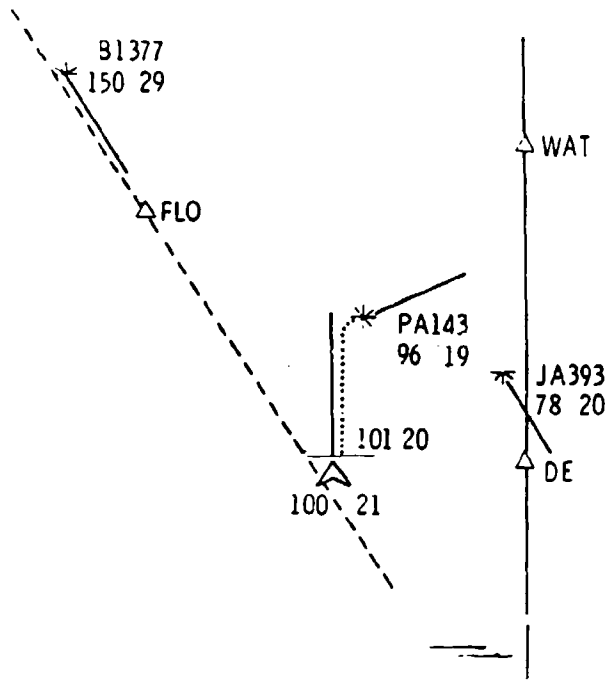


(a) Map scale 8.

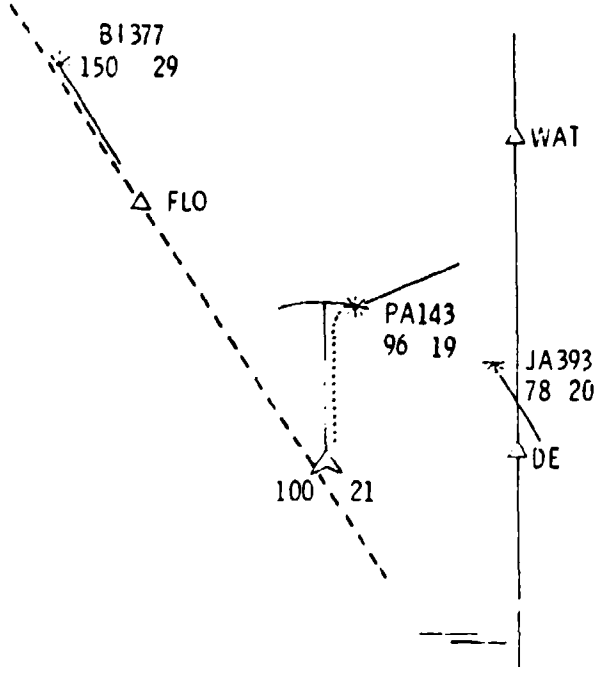


(b) Map scale 4.

Figure 8.- CDTI format from reference 4.



(a) Constant time delay cue.



(b) Constant time predictor cue.

Figure 9.- CDTI format from time-based self-spacing techniques study.

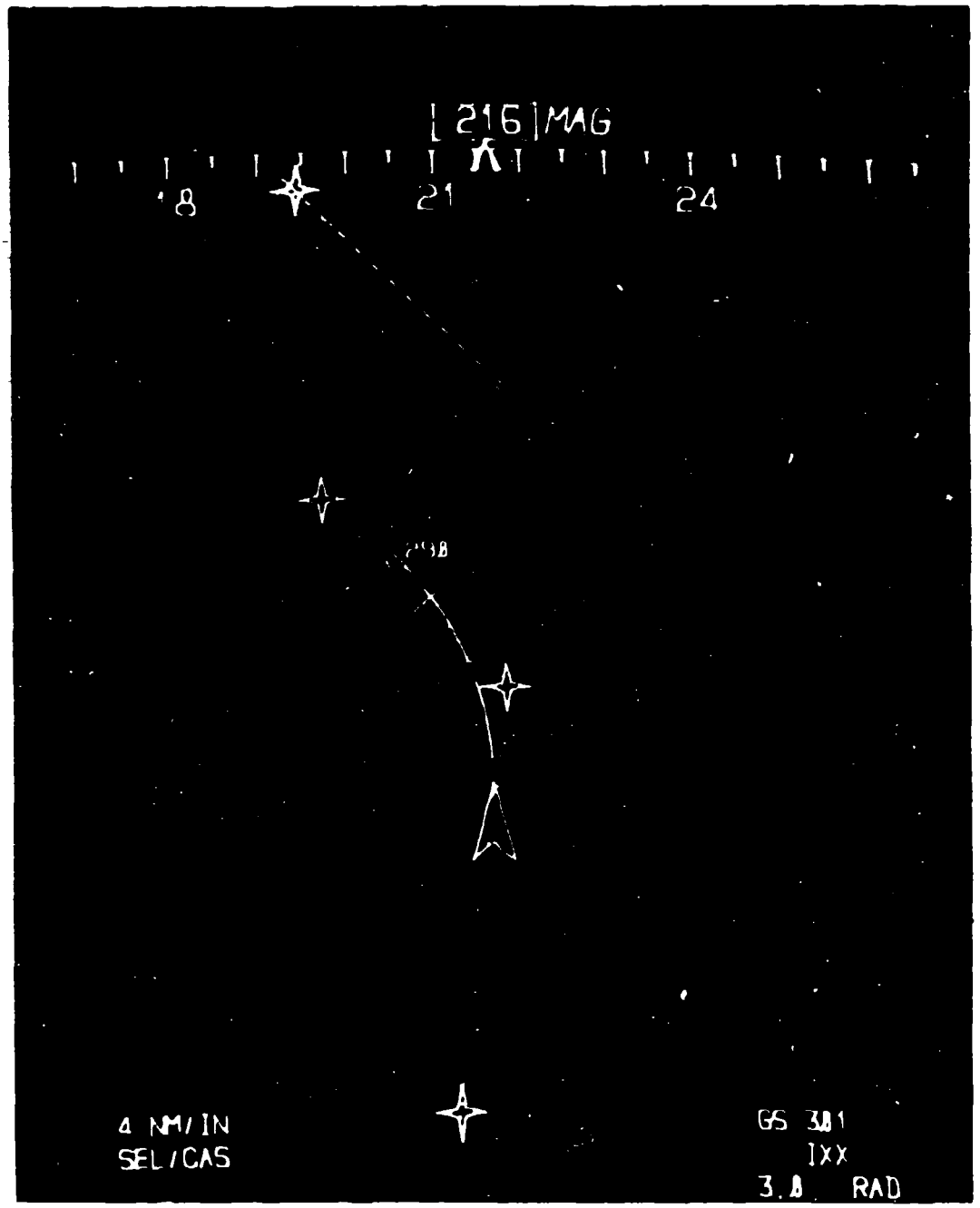


Figure 10.- Advanced cockpit CDTI format.

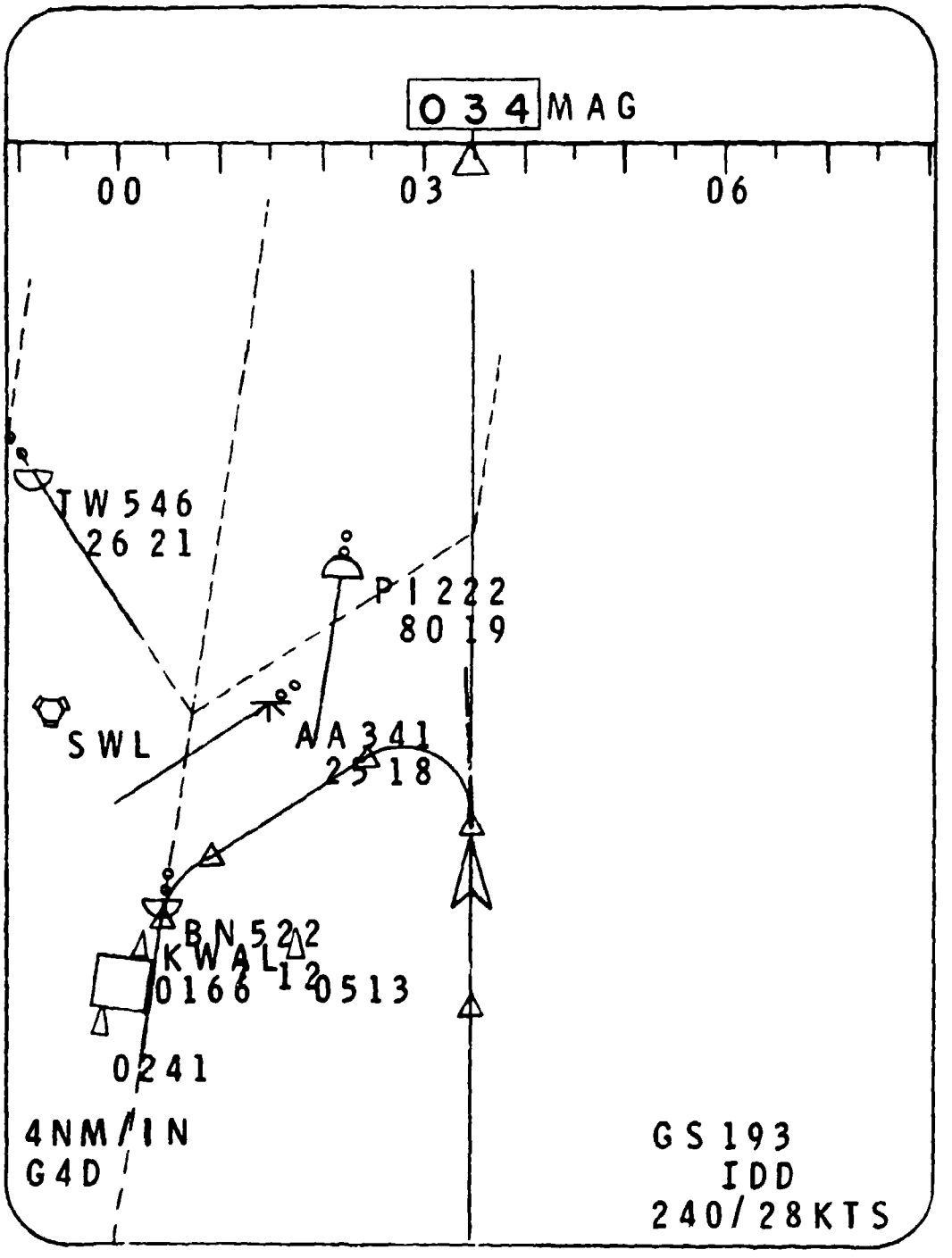


Figure 11.- CDTI format used in TCV flight tests.

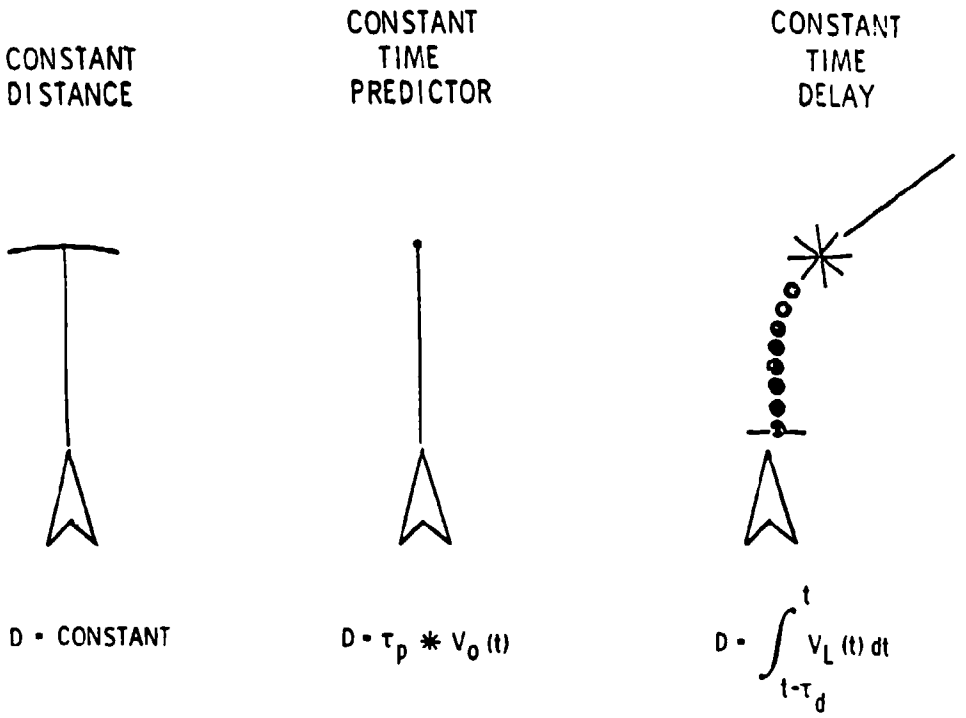


Figure 12.- Spacing criteria.

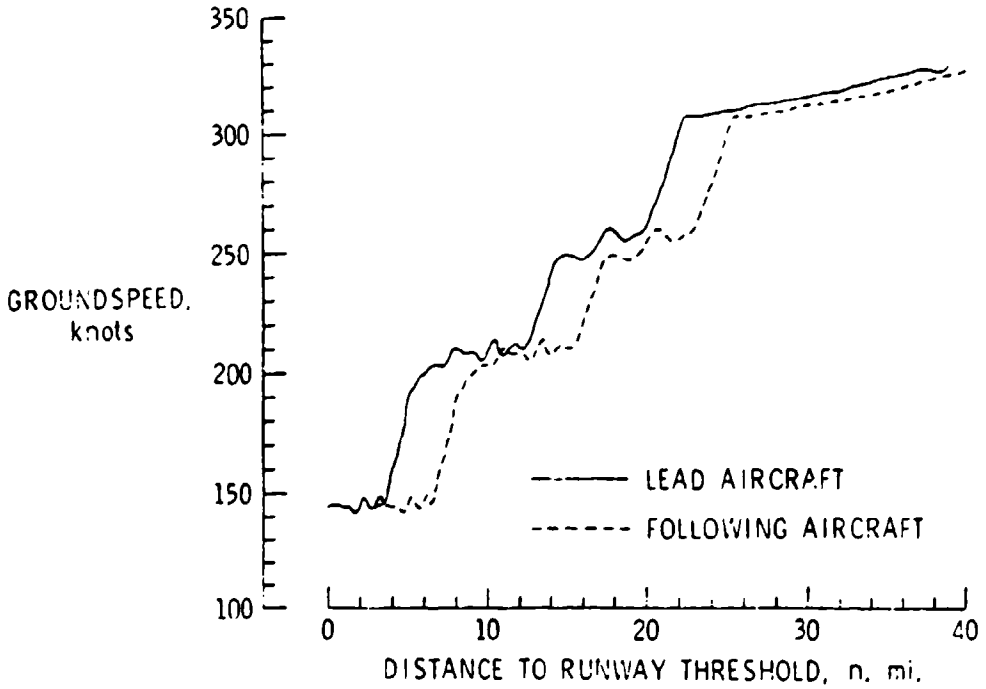


Figure 13.- Groundspeed profiles for constant distance spacing.

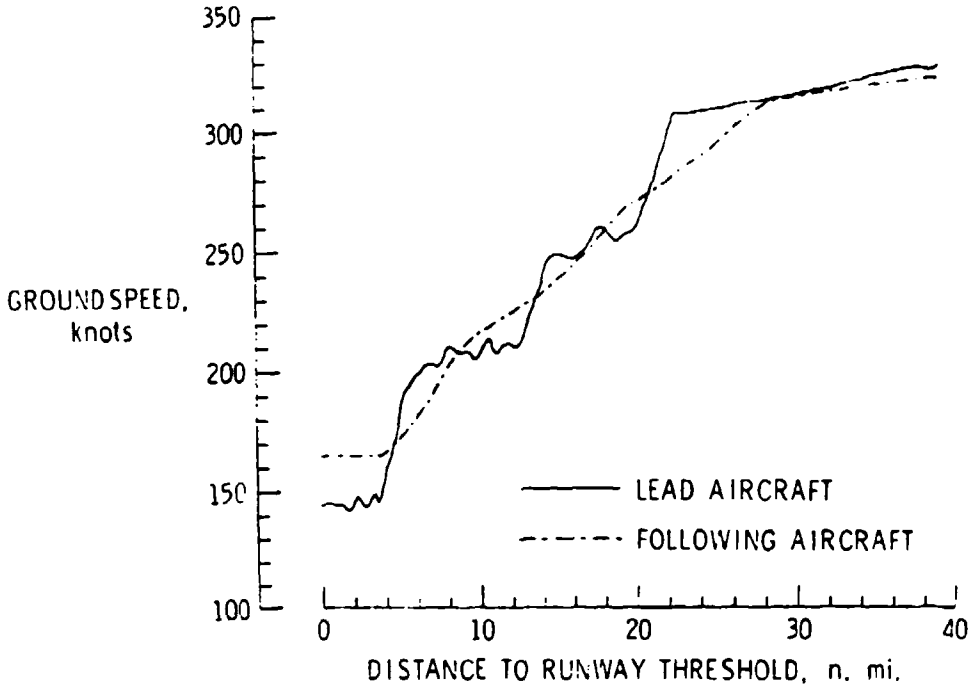


Figure 14.- Groundspeed profiles for time predictor spacing.

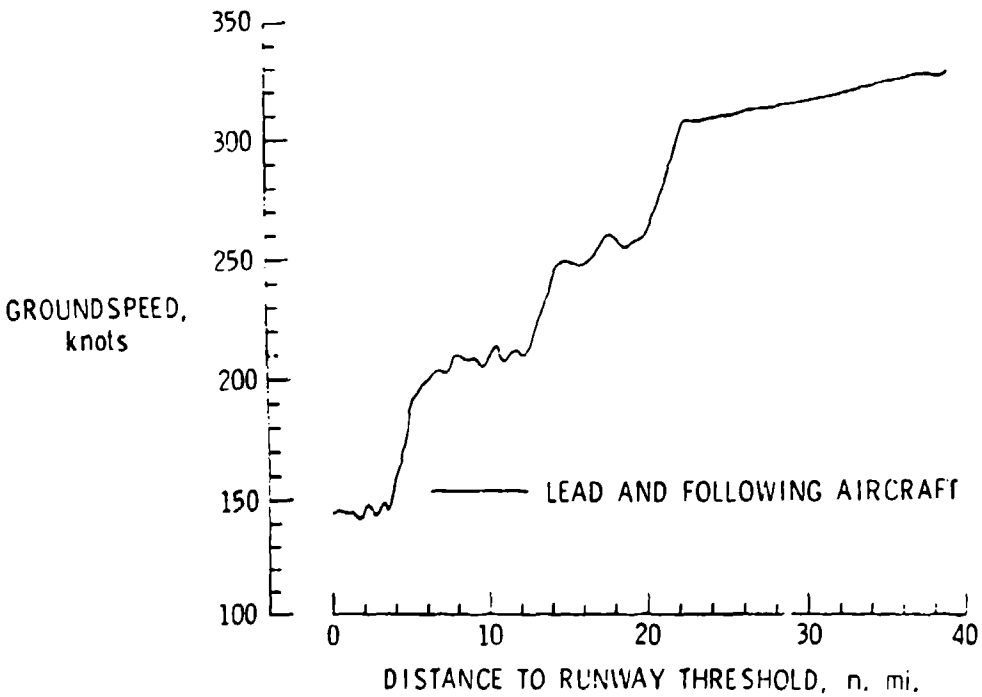


Figure 15.- Groundspeed profiles for time delay spacing.

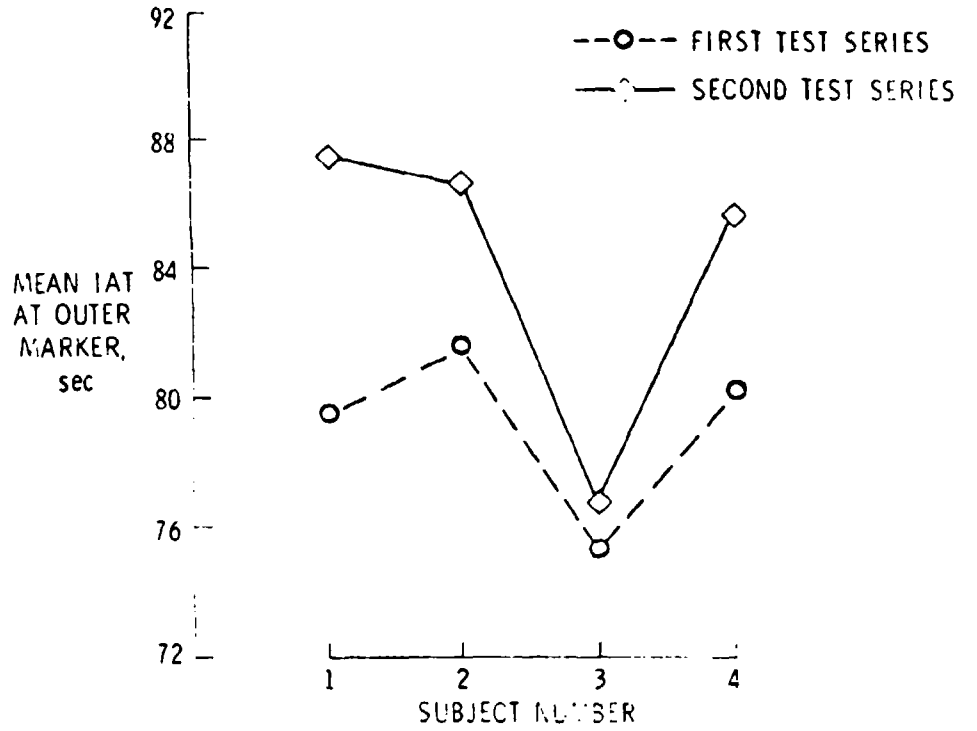


Figure 16.- Effect of pilot spacing technique on interarrival time at the outer marker.

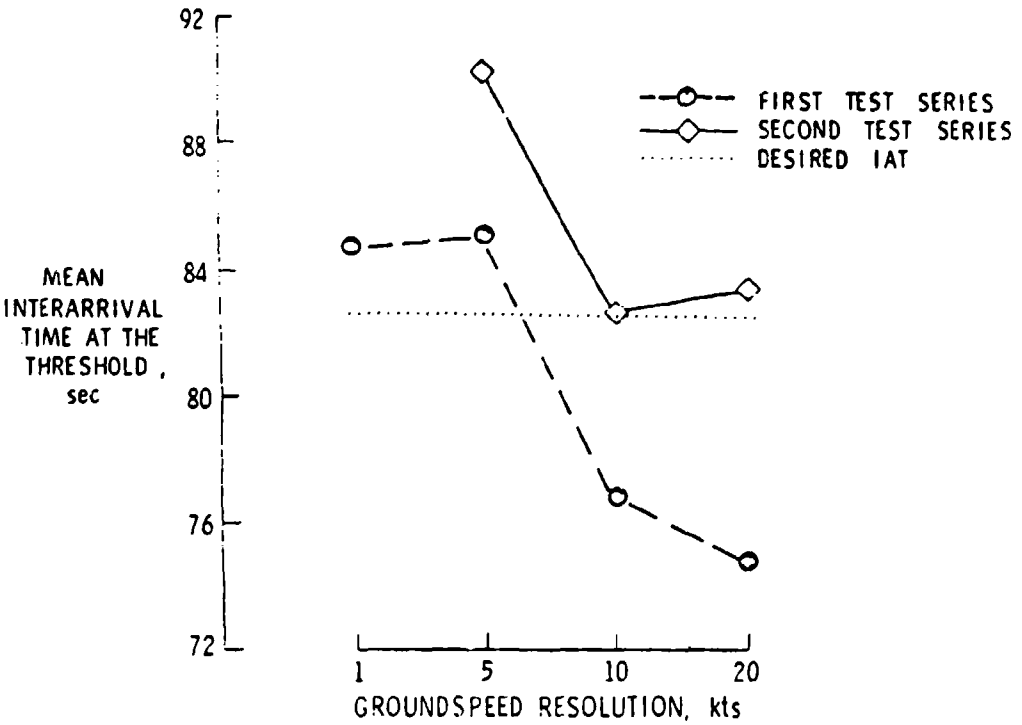


Figure 17.- Effect of groundspeed resolution on interarrival time at the threshold.

CHARACTERISTICS OF LEAD AIRCRAFT		
STEADY STATE	DECELERATING	MANEUVERING-DECELERATING
STANDARD DEVIATION OF IAT, sec		
<3	3-7	8

Figure 18.- Lead aircraft effect on interarrival time at the threshold.

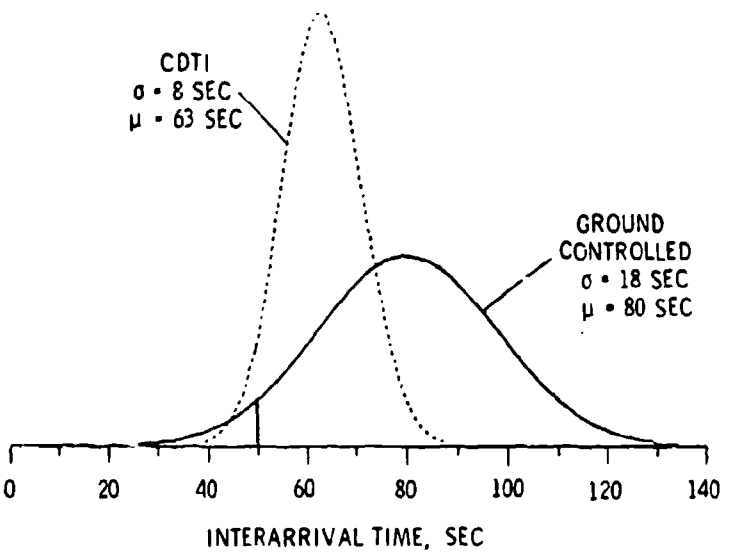


Figure 19.- Probability density curves for interarrival time.

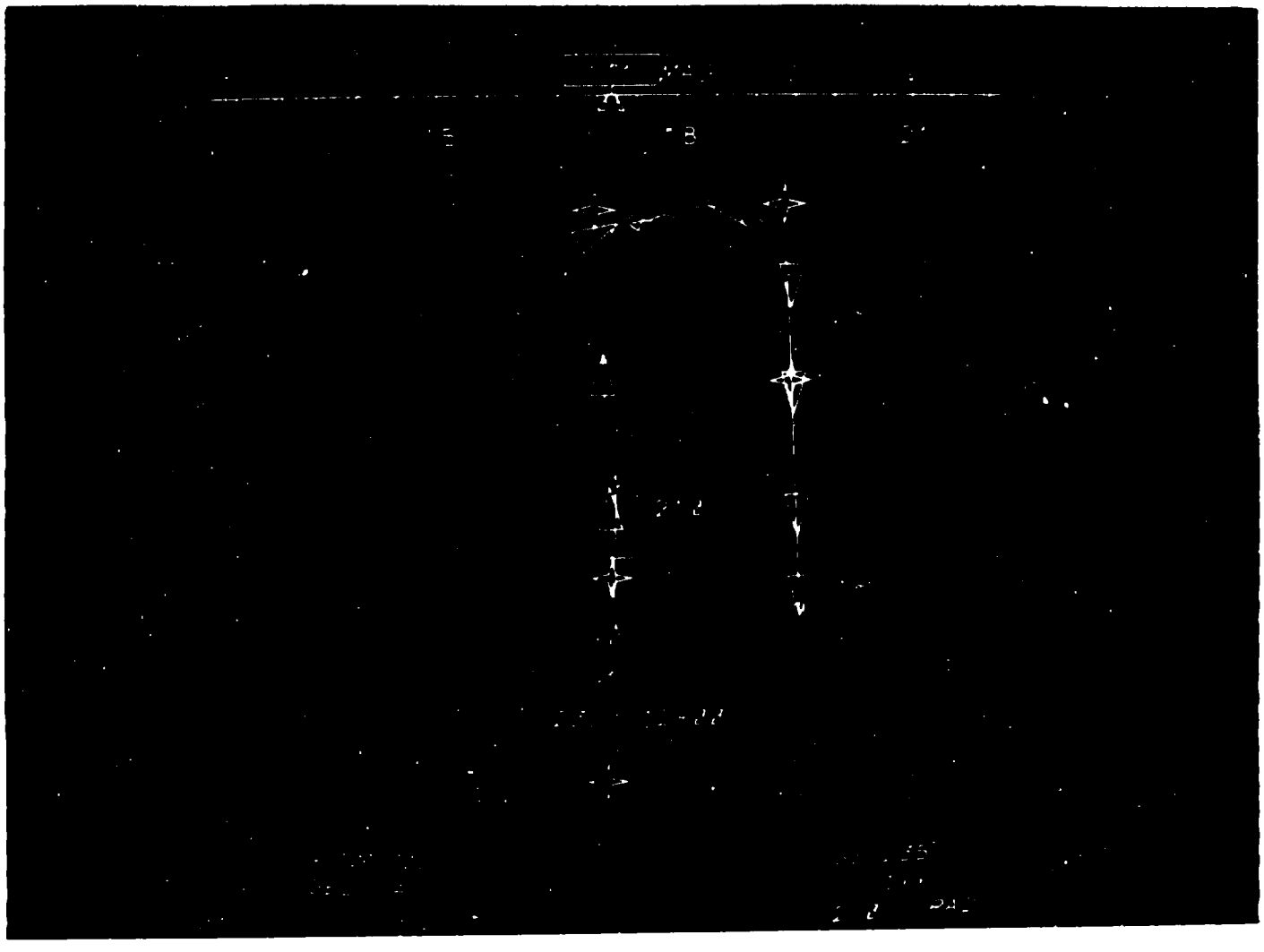


Figure 20.- CDTI format for multiple aircraft queues.

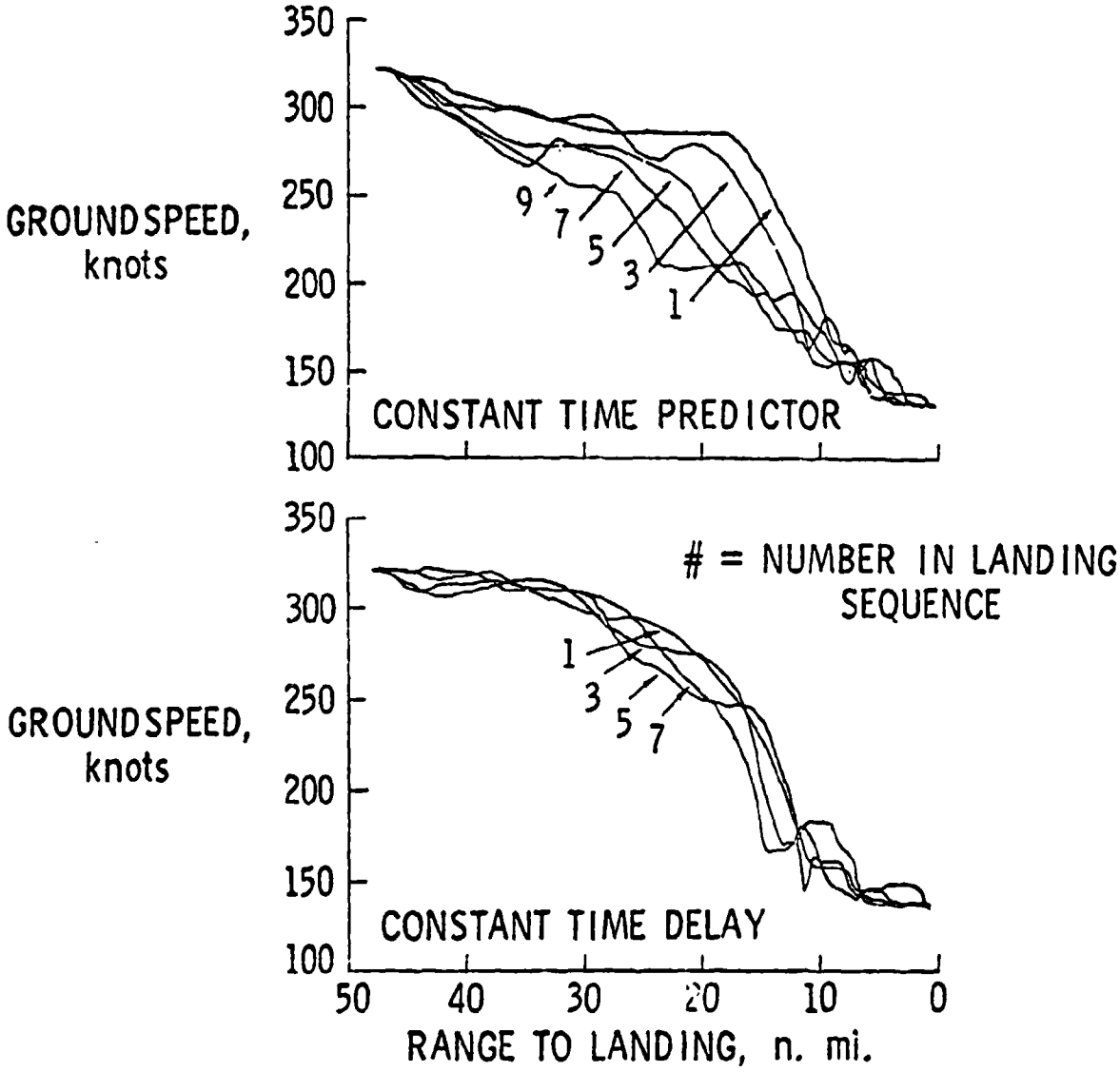


Figure 21.- Approach speeds in aircraft queues.

AD P001646

A Perspective Display of Air Traffic for the Cockpit

Michael Wallace McGreevy
Department of Electrical Engineering
and Computer Science
University of California, Berkeley
Berkeley, CA 94720

ABSTRACT

The typical format of the cockpit display of traffic information (CDTI) has been the plan-view [1-4]. Altitude has been represented numerically or by shape-encoding. [5][6] It is possible that manuever decisions are unfavorably biased and pilot workload increased by this kind of non-integrated format.

As an alternative, a perspective display has been developed. This traffic display presents a more natural view of the airspace around an aircraft. Perspective depth cues provide a sense of volume. Reference metrics enhance separation judgements. Particular emphasis is placed on vertical separation. A flexible computer graphics program drives the dynamic vector graphics display in real-time, allowing interactive development of the optimum viewing geometry.

The perspective and plan-view CDTI displays will be compared in experiments to be conducted in the summer of 1982. Further experiments will examine the ability of pilots to derive accurate mental models of spatial configurations from perspective displays, as a function of viewing geometry.

INTRODUCTION

For the last several years, NASA and the FAA have been studying the concept of a cockpit display of traffic information (CDTI). The study of CDTI display formats is a part of the human factors research of cockpit technology at NASA's Ames Research Center. The traditional format of CDTI has been a plan-view, with altitude represented numerically in data tags. Some recent efforts have been made to encode altitude by shape.

Experimenters at Ames who are studying pilot maneuver decisions using plan-view displays have become concerned that the plan-view may bias pilots' manuever decisions. A better representation of altitude was sought.

PERSPECTIVE DISPLAY

An informal analysis of projective display techniques suggested that orthographic projection might be preferable to perspective. The vertical dimension could represent altitude, and the horizontal dimension could be either along-track, or perpendicular to it. This, however, still required encoding of the third, collapsed dimension. Other positions for the plane of projection would negate much of the metric benefit of orthographic projection.

Perspective was then considered more seriously. Its advantage is that it is more natural, and may have some benefit for spatial judgements. Its disadvantage is that metric information is difficult to derive from such a display. To overcome this problem, a key element was the development of uncluttered metric symbology.

As an attempt to provide a better representation of the airspace for CDTI experiments, an interactive perspective display program was developed.

EQUIPMENT

The use of vector graphic equipment imposes a limitation on the representation options available, but also provides real-time performance. The vector display program that has been developed allows not only real-time traffic dynamics, but also real-time interaction with the display parameters. This allows display development to be done interactively, minimizing the number of iterations through the cycle of specification and programming. It also provides immediate feedback when varying the perspective parameters so that the contributions of each one to our particular needs can be assessed.

The display program is written in GMSI Pascal and runs on a PDP-11/70 mini-computer under the RSX-11M operating system. The display device is an Evans and Sutherland Picture System 2 (PS-2), which provides Fortran subroutines that are called from the Pascal program. Interaction with the program is via a DEC VT-100 video display terminal, PS-2 function switches (toggle and momentary action), and data tablet. Hardcopies may be made on a Versatec electro-static printer/plotter.

The display samples in this paper were done on the Versatec. Line qualities of the actual PS-2 display and the hardcopies differ in several ways. Vectors drawn by the Versatec are black on white, while vectors in the actual display are drawn in a range of intensities on a dark background. The judicious use of symbol intensities and intensity gradients is useful on the actual display for reducing visual clutter, but our hardcopy device does not support this necessary feature. Also, the vertical alignment of dashed lines in the hardcopies can cause some effects not present in the actual display.

DISPLAY CONTENT

The display presents a perspective view of the traffic surrounding ownship, with the ownship symbol at the center of the screen (fig. 1a-c,e). The exact position of each aircraft is represented by the nose point of the symbol.

Aircraft symbols are aircraft-shaped, and can represent various a/c types. The size of the ownship symbol is specified as a fraction of the display width and remains constant regardless of viewing geometry. The sizes of the other symbols vary as a function of perspective and distance from the eye. Thus, aircraft symbols on the near side of ownship appear slightly larger; those on the far side appear slightly smaller. The size difference depends upon the degree of perspective distortion established for the particular scene.

A horizontal reference grid is centered beneath ownship and is oriented so that one axis of the grid is aligned with ownship's heading. Grid lines are separated by a standard distance, typically 2 or 3 nautical miles. Grid lines perpendicular to the flight path move underneath ownship at ownship's forward speed, providing a sense of forward motion. The grid vectors are brightest near ownship and become dimmer with depth into the scene. This eliminates the clutter near the horizon that is evident in the hardcopies.

Metric lines, orthogonal to the grid, connect three points of interest: the aircraft position in three dimensions, the horizontal position on the ruled grid, and the point that corresponds to ownship's altitude, at that position (fig. 1d). The ownship-altitude intersection at each other-aircraft position is marked with an 'x', and the vertical separation distance is ruled with thousand foot reference tics. The aircraft symbols, tics and 'x' are drawn with bright vectors.

A data tag is associated with each other-aircraft and may contain any combination of information required, typically identity, altitude and vertical speed. The position of each tag within the three dimensional scene is solved in two dimensions on the plane of projection by a priority placement algorithm (fig. 2) which ensures that the tag is in a readable (non-overlapping) position. The algorithm considers the projected positions of all aircraft symbols, metric lines, and other data tags. It then selects the best free screen position near the appropriate aircraft symbol, based on a priority scheme. The priorities rank the preferred positions of the tag relative to the aircraft symbol, and may be changed interactively. Tags are drawn with bright vectors and stand out well against the background.

One version of the traffic display program incorporates sixty-second trajectory predictors and history dots. The latest addition to this program puts tics and the ownship-altitude 'x'

on the metric line at the end of the predictor, as an aid to judgements of future vertical position (fig. 1d).

VIEWING GEOMETRY

Viewing geometry, which determines the appearance of the display, is the specification of the position of the eye within the space, and the 3D space to be projected onto the 2D screen.

The position of the eye is variable in real-time. It may be positioned anywhere within a hemisphere centered on ownship, at or above ownship's altitude plane. The surface of a hemisphere of selected radius is mapped onto the data pad, which allows movement of the eyepoint over this surface by moving the stylus over the pad. Ownship is always the point of regard. A forward-looking view which provides the best horizontal and vertical resolution is preferred.

Specification of the viewing geometry (fig. 3) requires that an approximation to a cone of vision, called a viewing frustrum, be established. Three parameters of this frustrum are considered here:

fov: field of view or viewing angle; determines the amount of convergence of parallel lines;

we: the distance between the eye and ownship; The frustrum is a 3D "window", and the eye-distance is one of its parameters, hence "we" in PS-2 jargon.

range: the minimum distance between ownship and the nearest clipping plane (side of the frustrum), which is the closest that other objects can be to ownship without entering the displayed region.

The data pad may be used to vary the viewing geometry in real-time. By selecting appropriate modes with the function switches, the fov, we or range may be held constant while manipulating the others via the pad.

A constant viewing angle allows the convergence due to perspective to be held constant while changing the amount of area displayed around ownship. A constant eye-to-ownship distance allows the size gradient of symbols with respect to depth to be held relatively constant while changing the convergence and range. A constant range allows a constant minimum local region to be displayed around ownship while changing the perspective distortion.

Thus, the viewing geometry of the eye point and the perspective is easy to manipulate, allowing the spatial relationships in the display to be arranged as desired. This is important because the information in the display must be accurately perceived, and the effects of the various perspective parameters must be considered.

For example, the perceived proximity of an intruder varies

as a function of the viewing geometry, and the global scaling of the symbols. At any particular scale setting, the geometry may be manipulated to obtain the desired perception of threat (fig. 1abc).

When a particular viewing geometry is found to be of interest, one of a set of eight "style" switches may be defined so that all of the parameters may be set with a single switch. Any number of style sets may be defined and stored as files for later use.

DYNAMICS

The program has a primitive trajectory generator which drives all of the aircraft symbols. Once a file of initial positions, headings and rates are input, motions of the symbols are generated for as long as needed. Trajectories may be straight or variably curved, with selectable vertical speeds. The dynamics may be stopped and restarted at any time so that different viewing geometries may be assessed in different situations.

OTHER INTERACTIVE FEATURES

A snapshot switch allows hardcopies to be made as the display program runs. Since the essential viewing geometry parameter values may be displayed on the screen by switch selection, hardcopies may record both the appearance and the underlying geometry of any display. Other switches allow elements of the display to be erased.

Static interaction via the video display terminal (VDT) can be selected with a function switch. This mode of interaction allows: direct numerical specification of all parameters of the viewing geometry; output of the matrices that embody the geometric transformations; output of the 3D coordinates of each aircraft; adjustment of the brightness of elements in the display; change of the symbol scale; selection of traces and dumps. VDT interaction mode also allows redefinition of the following: the viewing geometry "style" switches; grid distances, orientation and motion; aircraft initial conditions and trajectories; tag placement priorities.

FUTURE PLANS

The perspective and plan-view CDTI displays will be compared in experiments to be conducted in the summer of 1982. Further experiments will examine the ability of pilots to derive accurate mental models of spatial configurations from perspective displays, as a function of viewing geometry. This should allow us to further evaluate the utility of perspective displays for spatial information transfer.

ACKNOWLEDGEMENTS

Special thanks to Fumei "Amy" Wu of Informatics, Inc. for her patient assistance with the intricacies of RSX-11M, for her expert help in converting the demonstration program to its multi-tasked form for running experiments, and for her continuing efforts to maintain and perfect that version.

REFERENCES

1. Herbert. G. Weiss and Richard W. Bush, "The Role of an Airborne Traffic Situation Display in an Evolving ATC Environment," Aviation Advisory Commission report PB-215 714, Lincoln Laboratory, Massachusetts Institute of Technology (May 1, 1972).

2. Boeing Commercial Airplane Company, "Cockpit Displayed Traffic Information Study," D6-42968, Seattle, Washington (September 1977). Sponsored by Federal Aviation Administration, Advanced Concepts Staff, Office of Systems Engineering Management

3. F. A. Palmer, S. J. Jago, D. L. Baty, and S. L. O'Conner, "Perception of Horizontal Aircraft Separation on a Cockpit Display of Traffic Information," Human Factors Vol. 22(5), pp.605-620 (October 1980).

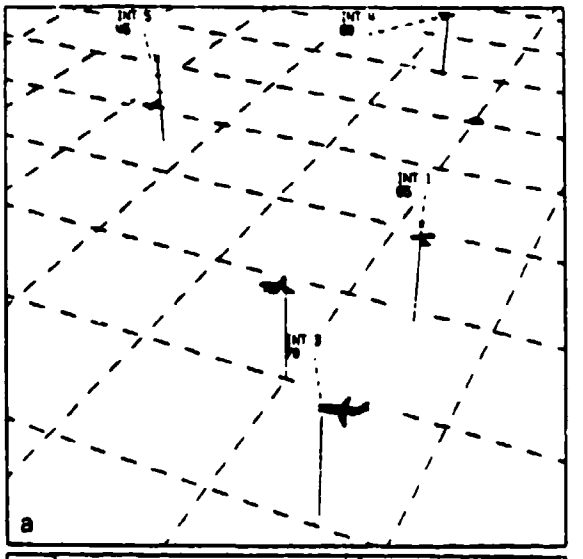
4. S. G. Hart and L. L. Loomis, "Evaluation of the Potential Format and Content of a Cockpit Display of Traffic Information," Human Factors Vol. 22(5), pp.591-604 (October 1980).

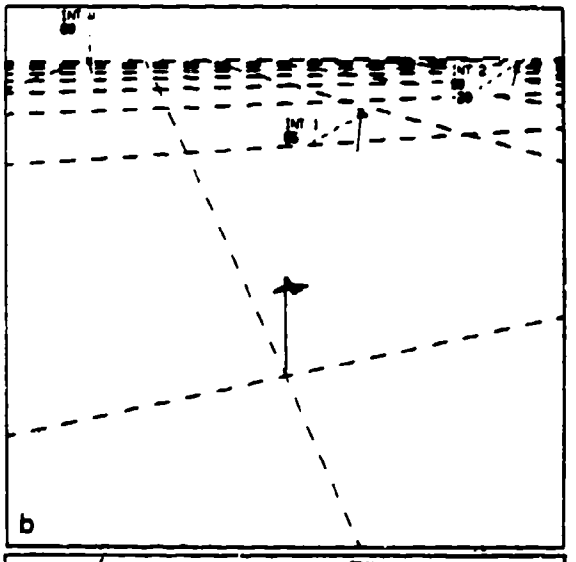
5. S. J. Hart, "Relative Altitude Display Requirements for a Horizontal Traffic Situation Display," Proceedings of the 17th Annual Conference on Manual Control (June 1981).

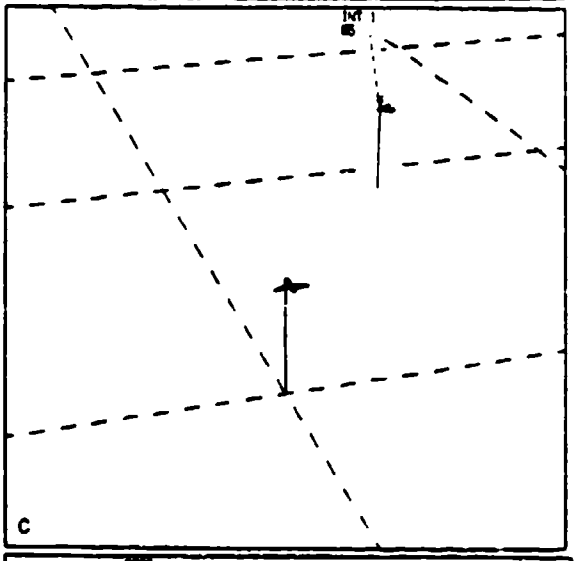
This work was partially sponsored by NASA grant NCC 2-86. Stephen R. Ellis, research psychologist at NASA's Ames Research Center, monitored the work and generously offered assistance and advice.

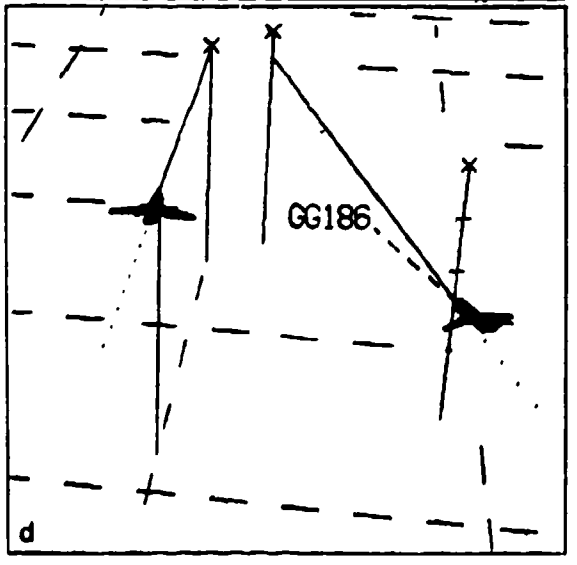
fig. 1 Examples

(a-c: same a/c positions)
a. compressed space
increases sense of proximity
b. expanded space
with strong convergence
reduces sense of proximity
c. expanded space
with weak convergence
moderates sense of proximity
d. full-sized detail showing
fully implemented metrics:
altitude change easy to read
e. scene for possible use
in experiments









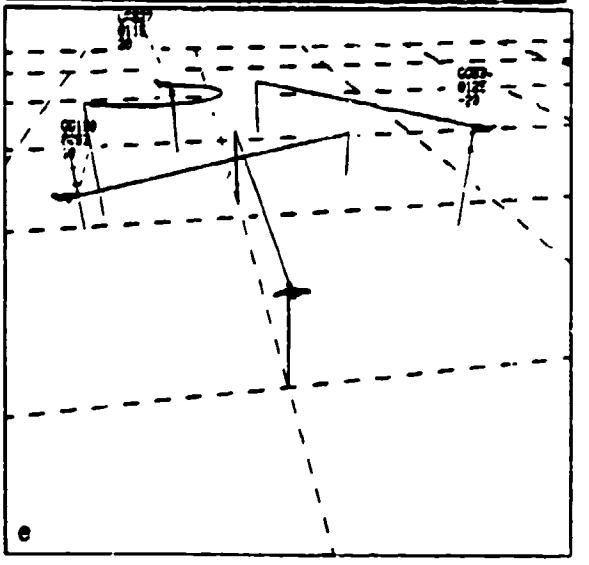
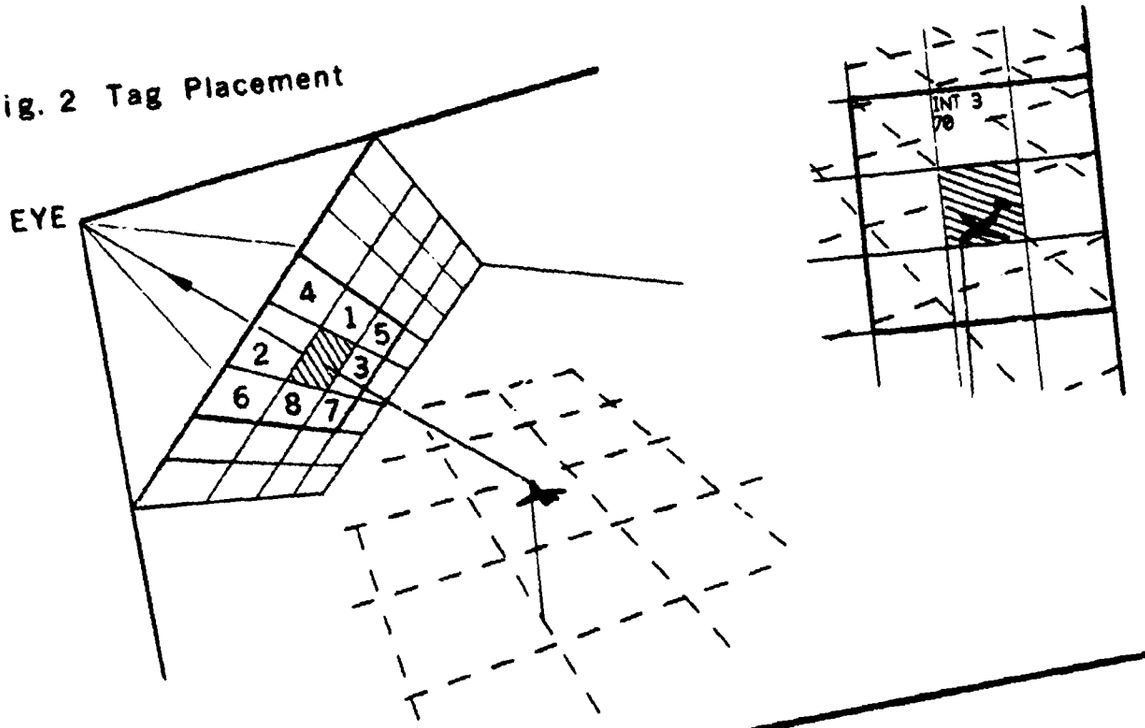


fig. 2 Tag Placement



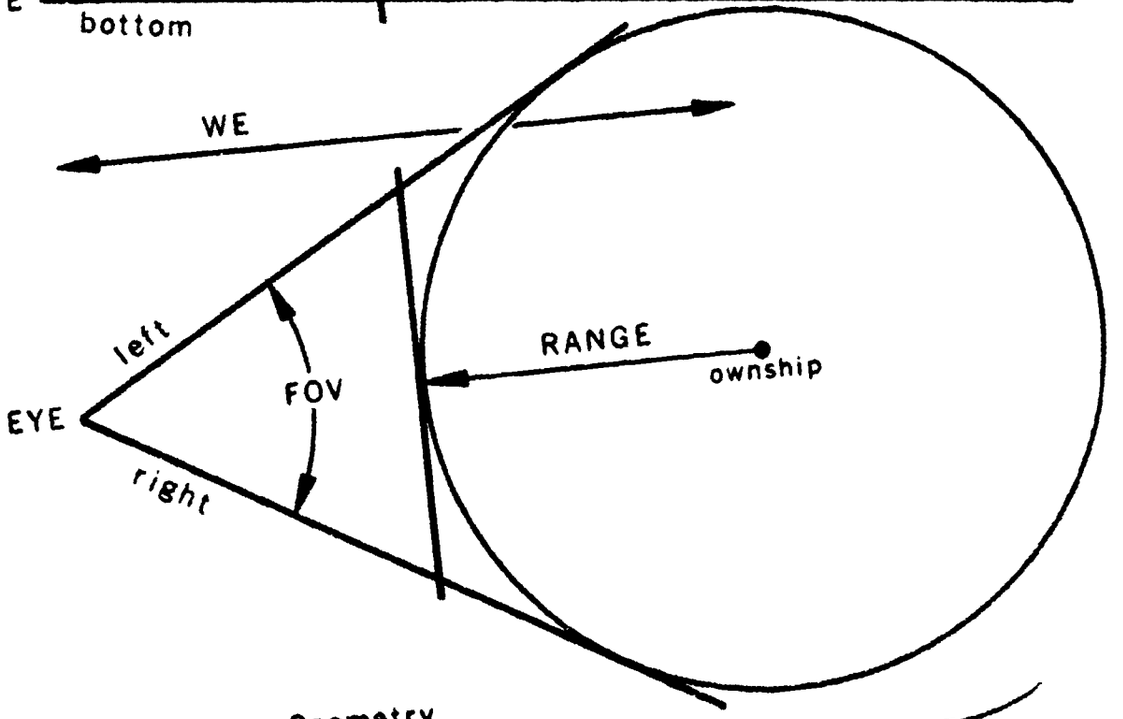


fig. 3 Viewing Geometry

AD P001647

EFFECT OF VFR AIRCRAFT ON APPROACH TRAFFIC WITH AND WITHOUT COCKPIT DISPLAYS OF TRAFFIC INFORMATION

Sandra G. Hart
Man-Vehicle Systems Research Division
NASA Ames Research Center
Moffett Field, CA 94035

ABSTRACT

This study investigated the impact of cockpit displays of traffic information (CDTI) on the flow of approach traffic. A mix of aircraft type, CDTI-equippage, and type of air traffic control (ATC) were included in the simulation. In addition, the practical issue of simulator fidelity in conducting such experiments was studied. Seven piloted simulators that represented a mix of general aviation-type and transport-type aircraft were simulated with two levels of control fidelity. They were flown by four teams of seven pilots each under ATC. A computer-generated target flying a predetermined flight path was also included to represent aircraft not in contact with ATC. The results indicate that aircraft type and not simulator fidelity influenced pilot and system performance. The frequency and content of communications, several measures of system performance, and pilot ratings also reflected pilot willingness to accept closer spacing and clearances to follow aircraft seen on a CDTI.

INTRODUCTION

The impact of providing aircraft with cockpit displays of traffic information (CDTI) on the national airspace could be as limited as providing pilot assurance or as far reaching as allowing pilots to assume responsibility for maintaining separation, terminal area merging, and detecting air traffic control (ATC) system failures. The impact of CDTI-equipped aircraft on the division of responsibility between the air and ground has been the object of research for years. However, in order to make meaningful decisons about the advisability of including CDTI in the national air system, additional research is needed to assess the potential benefits and liabilities of such a system. The benefits might include increased airport capacity, enhanced safety, improved fuel management, and pilot assurance. The liabilities might include increased pilot and controller workload, increased pilot attention in the cockpit rather than looking out the window and the possiblity that pilots might take unilateral actions based on information from a CDTI.

Much of the information that might be displayed on a CDTI is already available in the cockpit from a variety of sources inclu-

ding ATC, charts, maps, weather radar, transmissions of other aircraft, and looking out the window. Although this information is available, it is not integrated or presented in a convenient form. Thus, CDTI could provide, in addition to traffic information, relevant navigation, terrain, and weather information presented on a single multi-function display with a common reference, format and scale. This should assist pilots in forming mental representations of their environment and improving their situation awareness.

Because there is so much information that could be displayed on a CDTI, and so many ways it might be presented, elimination of obviously unacceptable alternatives in advance of simulation and inflight research was undertaken by Hart and Wempe (ref. 1) with a pilot opinion survey. Numerous candidate displays that incorporated different categories of information presented with varying levels of complexity, symbology, and format were presented to groups of pilots. They were asked to evaluate the display features and to specify a CDTI that incorporated essential information presented with minimal clutter and confusion. Subsequent part- task simulations were conducted (refs. 2, 3) to determine whether features preferred in the opinion survey would contribute to accurate and rapid assessment of the spatial relationship between a pilot's own and another aircraft. Finally, a simulation that included multiple piloted simulators and ATC controllers was conducted to validate the earlier experimental results in a more realistic environment (ref. 4). The experimental variables included: presence or absence of CDTI, symbolic encoding of the relative altitude of other aircraft, and presence or absence of aircraft flying under visual flight rules (VFR) intermixed with aircraft flying under instrument flight rules (IFR). The CDTI features that were simulated included digital readout of ground speed, altitude, and identification of other aircraft, flight path predictors and flight path histories for own and other aircraft, a navigation display, and pilot selectable map scale and orientation.

The purpose of the current study was to investigate further the impact of VFR aircraft on the ATC system in a terminal area environment in which all or some of the IFR aircraft were equipped with a CDTI. Some of the other variables considered were: (1) whether or not the altitude of VFR aircraft was known; (2) the effect of traffic mix (single- or twin-engine light aircraft or a jet); and (3) the impact of simulator fidelity on the behavior of the controllers and pilots. Three levels of simuation fidelity were provided from pilot-controlled simulators to computer-generated targets flying fixed flight paths. This issue was introduced to determine whether the differences found in the earlier simulation (ref. 4) between the number of clearance changes and vectors given to simulators with limited pilot control as compared to those with complete pilot control, occurred because of differences in simulation fidelity or because the simulators represented general aviation or transport aircraft respectively.

METHOD

Simulation Facility

The experiment was conducted in the multi-cockpit facility at Ames Research Center. The computational equipment consisted of a Digital Equipment Corporation PDP 11/70 computer and an Evans and Sutherland Picture System II. The research facility was separated into four areas: (1) three enclosed simulator cabs; (2) four "pseudo pilot" stations; (3) an ATC station; and (4) an experimenter's station. Communications were recorded for later transcription and experimental runs were videotaped. The computer update rate was 10 times per second, although relevant aircraft positions and pilot inputs were recorded once per second.

Simulator Cabs

The simulator pilot (SP) cabs contained an adjustable aircraft seat, throttle, and flap levers, communication radio tuner, head set, map scale selection switch, and a manual control stick with trim switch. The basic information necessary for IFR flight was presented graphically and digitally on a 57.7 cm Xytron display centered in front of the pilot. (Figure 1) A vertical situation display presented information about altitude, vertical speed, heading, indicated airspeed, ground speed, fuel flow, map scale, pitch attitude, bank angle, raw glideslope and localizer deviation, and distance from the destination airport (DME). All measures were computed and displayed in nautical miles, knots, and feet, as these are the units of measurement used in aviation.

A 12.5 cm display located beneath the vertical situation display showed significant navigation features for a southern approach to San Jose Airport. It translated and rotated beneath the symbol for ownship (a chevron) which was centered laterally and displaced toward the bottom of the display with a heading-up map orientation. Five discrete map scales ranging from 118 to 7 km and an automatic mode in which map scale varied continuously with altitude from 2 to 118 km were provided. When SPs were given a traffic display, all aircraft within map range and a pilot-selectable range above or below the pilot's own altitude were displayed. Digital data tags (identification, ground speed, altitude, and direction of vertical flight) were provided for IFR aircraft on all CDTI trials and for VFR aircraft as well when they were simulated with altitude encoding transponders. The shape of the symbols that represented other aircraft depicted the relative altitude of those aircraft: a full hexagon indicated that an aircraft was within +/-150 m of a pilot's own altitude and the upper or lower half of a hexagon indicated that it was more than 150 m above or below respectively. This discrete relative altitude coding had been preferred in the earlier studies (ref. 1, 4). If altitude for a VFR aircraft was not known, this aircraft was displayed at all times as though it was at the same altitude

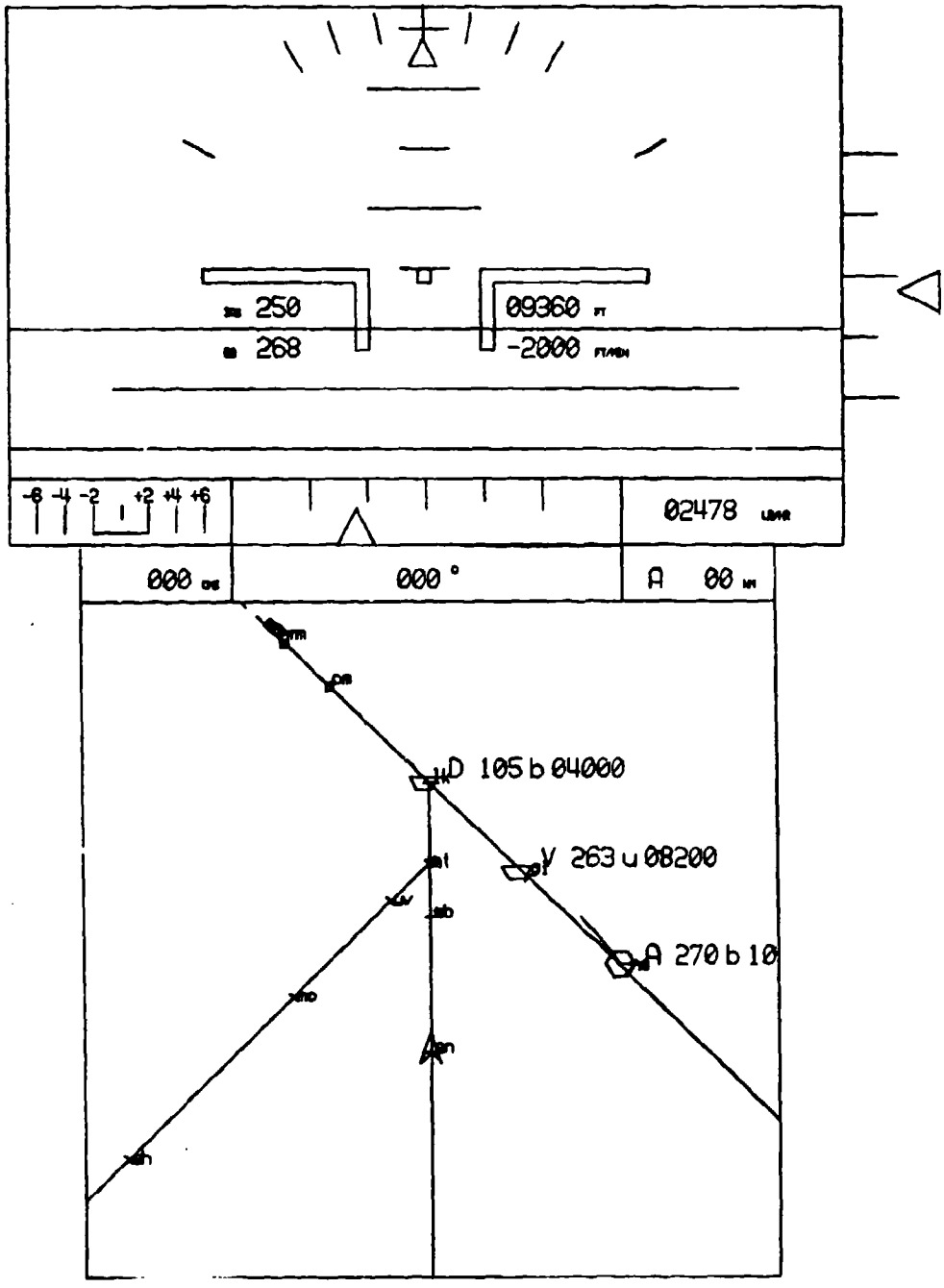


Figure 1: Simulator pilot display including traffic information

as ownship, assuming the worst case. Otherwise, it was displayed according to the same algorithms as the IFR aircraft. A two-segment, 30-sec flight path predictor and flight path history showing four previous positions each 8 sec apart were added to the symbols for own and other aircraft.

Pseudo pilot station

The term "pseudo pilot" does not refer to the qualifications of the operators (all were instrument-rated general aviation pilots) but rather to the simple controls, displays, and aircraft models provided for these pilots. The movement of these aircraft across the pilots' and controllers' traffic displays was, however, identical to that of the more relistically modeled and controlled simulator cabs. The communications and procedures followed were identical. The pseudo-pilots (PP) were included in the simulation, as were the controllers, to increase the realism of the simulation for the SPs who were a primary focus of the experiment.

The four PP stations shared a 57 cm Xytron display that presented a north-up fixed-scale traffic situation display. (Figure 2) All active aircraft within the 92 km range of the map were displayed at all times represented by triangles oriented north-up. The positions of the targets were updated every 4 sec, simulating the radar update rate in the terminal area. Graphic flight path history, digital ground speed, identification, and altitude were shown for all IFR targets and for VFR targets with altitude encoding transponders. The display represented a rudimentary traffic display as it was north-up, fixed reference and map scale, with limited symbolic information about the direction of flight and relative altitude of other aircraft.

Current and commanded speed, heading, and altitude for each PP aircraft were presented in four tables at the peripheries of the display. Commands were entered by the pilots of each aircraft by moving a joystick (heading and altitude) or a slide switch (speed). Aircraft targets achieved commanded values at realistic rates for the type of aircraft simulated. These targets were preprogrammed to complete a nominal approach and landing with no input from a pilot; however, once a pilot assumed control over one or more functions, responsibility for that function was transferred to the pilot for the remainder of the approach. The information displayed was not adequate for PPs to land their aircraft; thus the final approach and landing were accomplished by an automatic landing algorithmn that became active at the outer marker.

ATC Station

Stations were provided for an approach controller and a tower controller; no departures were simulated and the problem was initated by a written "handoff" from a simulated enroute control facility. The controllers shared a map display that depicted an

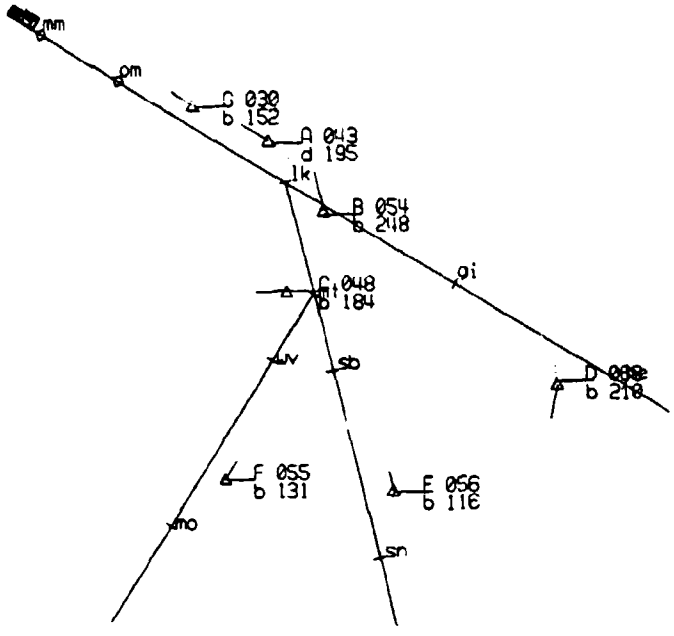


Figure 2: ATC and pseudo pilot traffic display.

area south of the San Francisco Bay Area controlled airspace which was identical to the map portion of the PP display. (Figure 2) Each controller was given a headset, microphone, and radio frequency selector.

Scenarios

The simulation was based on a southern approach to runway 30L at San Jose Airport. The location and names of some of the waypoints were modified to create three converging routes with reporting points located at equal distances from the airport along the routes. Aircraft entered the simulation at one of seven waypoints located 65, 46, or 35 km. from the airport. The descent from the LICKE intersection to the airport (DME 16) was based on a standard approach plate for runway 30L. The localizer and glideslope were intercepted at LICKE for a simulated instrument landing system (ILS) approach.

Appropriate initial altitudes and starting speeds were programmed for the three aircraft types and the different starting locations. Entry into the system was scheduled to simulate routine handoffs from an enroute control facility. Longitudinal separation of at least 9 km for aircraft at the same altitude or a vertical separation of at least 1200 m for aircraft at the same position was provided as an initial condition. The end of each run occurred when the last aircraft landed.

Twenty different scenarios were developed that represented four levels of VFR aircraft intrusion. Four were used for practice

presented with each of the four experimental conditions and the remaining 16 were used twice: once when all aircraft were CDTI-equipped and again when only PPs were given a traffic display. No scenario was used more than one time for any group. The different combinations of task difficulties generated by initial condition, VFR aircraft path, and traffic mix were presented in a balanced order between and within groups.

VFR flight paths were created in the simulator cabs in advance of the simulation and played back at predetermined times to allow controlled recreation of situations typical of the San Jose area. These aircraft were not in contact with ATC as they were flying under VFR outside controlled airspace.

Meteorological Conditions

Eight different Airport Terminal Information Service (ATIS) messages were recorded to provide information appropriate for each approach. Marginal meteorological conditions with wind from the west or east at 46 km/hr decreasing to 28 km/hr across the runway with low turbulence and no shear were presented in a balanced order across experimental conditions.

Aircraft Models

Two cabs were configured as medium jet transports and one as a twin-engine light aircraft. The simulator cab models (ref 5) were based on principles of fundamental fluid dynamics with parameters that represented particular aircraft features or mechanics. The goal was to balance simplicity (for manageability) against complexity (for realism) to provide for significant, observable effects to pilots rather than to produce a faithful reproduction of the aerodynamics involved. The model provided a full operating range from cruise to landing for many aircraft simultaneously. PP aircraft were altitude-configured to fly like one single-engine or two twin-engine light aircraft or a medium jet. Commanded changes were achieved by rule-of-thumb algorithms, turns were standard rate and coordinated, climb and descent rates were constant, and nominal profiles were appropriate for the aircraft type (ref.6).

Experimental Variables

All aircraft were equipped with CDTI displays on one half of the trials (allCDTI). PPs but not SPs were given a traffic display (someCDTI) on the rest. On half of the allCDTI or someCDTI trials, VFR aircraft without altitude encoding transponders were simulated, and thus no altitude information was displayed (noTAG). On the remaining trials, altitude information for VFR traffic was displayed to whomever was given a traffic display for that trial (TAG). Two replications of each of the four experimental conditions were obtained; one with minimal VFR traffic interference and one with considerable interference.

Subjects

Four instrument-rated general aviation pilots served as PPs for each group. One general aviation pilot and two current transport pilots flew the simulators. A recently retired air traffic controller and a specially trained instrument-rated pilot served as the approach and tower controllers respectively. Four groups of seven pilots and two controllers each served as paid participants.

Procedure

Each of the four experimental groups was familiarized with the purpose of the experiment, their role in it, and given several hours of practice flying the simulators on the day before the experiment. Four additional practice runs were conducted on the morning of the experiment followed by eight experimental runs in the afternoon. Each practice or experimental run lasted 25-30 min and consisted of a VFR flight and seven approaches by four PPs and three SPs. SPs completed a questionnnaire at the conclusion of each approach and again at the end of the entire experiment to obtain their opinions about the scenarios, displays, CDTI, the simulation, their own performance, and the workload experienced.

Written Materials

All participants were given an area map and an approach chart for an ILS approach to runway 30L at San Jose Airport. PPs and SPs were given information about the operating characteristics of their aircraft and a nominal descent profile. The controllers were given "flight data " strips for each aricraft before it entered the terminal area and the pilots were given written "clearances" with the time and place of their entry into the simulation and their initial speed and altitude. SPs were also given a set of rating scales to completed after each approach. Fourteen of the scales presented task-related, operator-related, and performance-related dimensions in a bipolar format. Three additional scales developed by Airbus Industrie for use in certification (ref. 7) were tested that asked SPs to rate the frequency of interruptions, amount of free time and likelihood of errors.

RESULTS

Measures of pilot effort, pilot performance, system efficiency, and safety were analyzed to determine the effect of information about VFR aircraft and presence of absence of CDTIs in all aircraft on a variety of performance indicators. In addition, pilot responses to several forms of subjective rating scales were obtained to evaluate the impact of these parameters on the pilots. All measures were examined initially for differences among experimental groups with a series of one-way analyses of variance for

independent groups. No significant group differences were found for any measure, so the remaining analyses were performed across individual SPs (n = 12) or PPs (n = 16). The effects of display conditions and replications on each of the measures were examined with three-way analyses of variance for repeated measures. Individual pilot performance measures such as localizer and glideslope deviations were not examined in great detail because the quality of the simulation did not warrant such an analysis and it had been shown previously (ref. 4) that such measures did not reflect variations in display and system parameters such as those under investigation.

System Performance

Four measures of system performance were studied: (1) separation violations; (2) intercrossing times; (3) duration of approaches; and (4) communications. For these analyses as well as others, the effect of the experimental variables was examined for the system as a whole, for the different aircraft types (single or twin engine light aircraft or medium jet) and for the two levels of simulator fidelity (SP and PP).

Separation Violations

Four levels of separation violation severity were selected in advance of the simulation to reflect four levels of system failure. The levels ranged from slightly less than standard IFR separation (Class 1) to a collision (Class 4) The frequencies of separation violations by experimental condition, simulator type and aircraft type were computed excluding any violations that occurred between a PP aircraft and any other aircraft after the automatic landing system had taken control of the PP aircraft. As no control of those aircraft was possible after that point, separation violations did occur, but as a consequence of simulation limitations rather than system failures. Of the remaining violations, more instances of less than standard separation occurred with CDTI in all aircraft (on 6% of allCDTI approaches) than when only some airraft had CDTIs (on 2% of these approaches), possibly indicating pilot willingness to accept closer spacing with CDTI than without. (Table 1) More separation violations occurred with at least one aircraft that was not under ATC control (on 19% of VFR approaches) than between two aircraft that were (on 3% of all IFR approaches). Fewer violations occurred between two PP aircraft (on 3% of PP approaches) than with one or two simulators (on 8% of SP approaches). No collisions occurred during the simulation.

Intercrossing Times

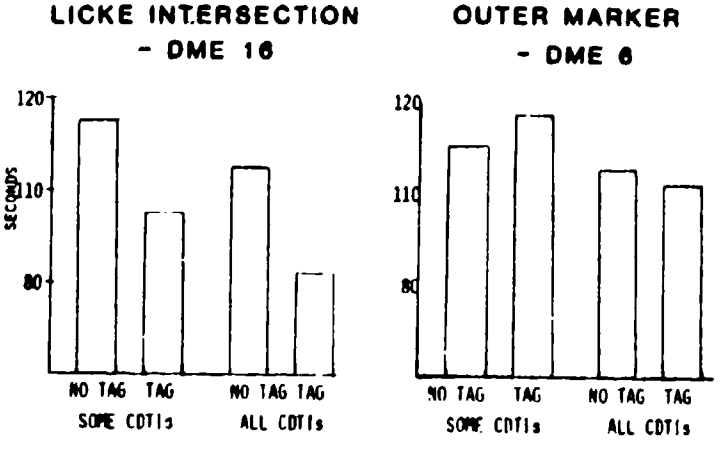
The mean intercrossing times between aircraft at the LICKE intersection and the outer marker were computed as measures of system efficiency and safety under different experimental conditions.

TABLE 1: Separation Violations that occurred between two or more aircraft during 256 approaches (Class 1 to 4 respectively denotes an increasing severity of loss of separation from less than standard separation to a collision)

	IMPACT OF CDTI		IMPACT OF VFR AIRCRAFT		SIMULATOR FIDELITY	
	Some	All	VFR/IFR	IFR/IFR	Pseudo	Simulator
CLASS 1	1	2	2	1	2	2
CLASS 2	1	4	3	2	1	4
CLASS 3	0	2	1	4	1	2
CLASS 4	0	0	0	0	0	0

Two three-way analyses of variance for repeated measures were performed on the mean group intercrossing times. (Figure 3) Aircraft spacing was less with allCDTIs than with someCDTIs (105 vs. 94 sec) and less with altitude information for VFR aircraft than without (110 vs 89 sec) at the LICKE intersection. These differences were not statistically significant, however. All differences related to displays were eliminated by the outer marker. Intercrossing times were 12 sec longer, on the average, at the outer marker than at the LICKE intersection, as all aircraft were established on the ILS, and had achieved approximately 9.25 km (5 n mi) separation by that point to allow for

Figure 3: **AVERAGE INTERCROSSING TIMES (sec)**



differences in approach speeds among the aircraft. There were no significant replication effects at either location nor any interactions among experimental conditions.

Approach Durations

Average approach durations, in time and distance, were analyzed with two three-way analyses of variance for repeated measures. The simulation scenarios had been set up so that all jets and most twin-engine light aircraft were initialized at a distance of 65 km (35 n mi)along the routes from the airport. Additional miles flown might thus serve as a rough indicator of path stretching and delaying maneuvers given by ATC. The average distance flown by the twin-engine aircraft was greater than the jets, indicating, as did the communications analysis that these aircraft received more vectors from ATC. There was a reduction of 3 n mi in the distance flown by light twins with the addition of information about the altitude of VFR aircraft and a similar, but smaller, reduction with the addition of CDTIs for all aircraft. There was no such difference in distance flown for jets as a function of display condition, indicating that ATC focused on the general aviation aircraft rather than the jets when flight path modifications were required for separation. The single engine aircraft flew the shortest distances, primarily because they typically entered the system from a point much closer to the airport than the rest. These aircraft were typically vectored off of the final approach course to expedite faster traffic and then brought back on later. The addition of CDTIs in all simulators resulted in shorter approaches for these aircraft, as it did for the others, but no such difference was found when altitude information was given for VFR aircraft. The reduction of distance traveled per approach as a function of some (36 n mi) or allCDTIs (35 n mi) was significant across all aircraft types ($F = 6.23$ $(1,6)$ $p < .05$). No significant differences were found for VFR data tag conditions or replications and there were no significant interactions.

As one would expect, the length of time taken to complete an approach was not the same for the three aircraft types due to differences in approach and landing speeds. PP flight times were longer, on the average, as three of the four PP aircraft were simulated as slower general-aviation-type aircraft. There were no differences in flight times as a function of either display conditions, nor any significant interactions. There was a significant decrease in average flight time from 14.0 to 13.6, min as the experiment progressed ($F = 9.73$ $(1,6)$ $p < .01$) possibly reflecting increased system efficiency with practice.

Communications

All communication between pilots and controllers were recorded, however four trials were lost for Group 4 due to faulty equipment. Consequently, the communications were not subjected to

statistical analysis. Averages for the remaining data and the trends that were observed will be described, however, because of the important role that communications plays in this type of simulation.

Communications were conducted realistically by all participants, providing not only an important element in creating a realistic environment but also a valuable source of information about the effects of the experimental variables on the ATC system. A content analysis was performed by playing back the videotapes and scoring the number, types, source, and destination of each message using a modification of the categorization scheme developed by Chappell and Kreifeldt (ref. 8). Since many transmissions conveyed several different types of information, the number of messages transmitted was greater than the number of separate transmissions. The categories included: specific types of traffic-related messages, types of ATC commands, vectors or questions about them, position reports, requests to maintain current status, and acknowledgements.

Table 2: Mean number of messages transmitted by pilots or controllers on each approach as a function of display condition.

TOPIC OF MESSAGE	FROM PILOTS					FROM CONTROLLERS				
	CDTI		VFR			CDTI		VFR		
	some	all	noTAG	TAG	TOTAL	some	all	noTAG	TAG	TOTAL
TRAFFIC	4.1	4.4	3.8	4.8	8.5	3.3	6.8	3.9	6.3	10.1
ALTITUDE	7.3	7.2	6.9	9.7	14.5	13.7	15.4	13.5	15.5	29.1
SPEED	2.2	1.7	1.8	1.9	3.9	14.3	11.1	13.4	12.0	25.4
CLEARANCE	.7	.7	1.8	1.9	1.4	20.0	22.5	21.7	20.8	42.5
VECTOR	-	-	-	-	-	15.1	12.7	13.8	14.0	27.8
POSITION	22.1	17.5	21.1	18.6	39.6	10.1	7.1	10.2	7.0	17.2
MAINTAIN	.1	1.7	.6	1.2	1.8	9.6	9.7	10.4	9.0	19.4
ROGER/ READBACK	63.3	57.9	59.0	62.4	121.2	14.5	13.2	14.7	13.0	37.7
TOTAL	99.8	91.1	93.2	97.3	190.5	100.6	98.5	101.6	97.6	199.2

The total number of messages transmitted between pilots and controllers and messages specifically related to traffic may be seen in Table 2. On the average, nearly 200 messages were transmitted by the two controllers and seven pilots during each set of approaches. Although there were an approximately equal number of messages originating with the air and ground, the information content was quite different. More than twice as many altitude, speed, clearance, and vector commands, queries, and responses originated with ATC as from the pilots. Twice as many position-related messages originated with the pilots as with the controllers, and four times as many acknowledgements. All of the controller acknowledgements were simple "rogers", where as the majority of the pilots' involved a complete or partial readback.

In general, there were fewer transmission with all CDTIs than someCDTIs (200 versus 189). This was attributable to fewer vectors issued from ATC and fewer position-related transmissions from pilots (Table 2). For those messages specifically related to traffic, however, the reverse was true. There were more traffic-related messages when all pilots had traffic displays (7.4) than when only some pilots had them (11.2). It was apparent that ATC did use the pilot's ability to see other aircraft on the CDTI when it was available to involve pilots in maintianing "visual" separation by giving traffic advisories and clearances to follow other aircraft, and directing pilots to maintain separation from aircraft. In general, the additon of VFR data tag information, had less impact, although the number of position-related messages did decrease from 31 to 25, on the average, with its addition.

Table 3 presents the communications analysis by experimental condition, aircraft type, and topic, with messages to and from pilots combined. Pilots of single engine PP aircraft were given and initiated the fewest communications. This is consistent with the observation that ATC vectored these aircraft out of the way of the faster traffic and ignored then until other faster aircraft had passed. Their most frequent type of communication was a command to maintain present status or a readback.

Both PP and SP twin-engine aircraft were given more vectors than any other commmand, and received and provided a number of position queries. Both types of communications were most evident when information about the altitude of the VFR aircraft was not known and there were no traffic displays in the simulators. In addition, there were a large number of "maintain..." commands given to these aircraft in the noTAG, allCDTI condition.

The pilots of the jets participated in relatively fewer communications than did the pilots. In general, the presence or absence of altitude information for VFR aircraft affected the frequency of responses made by jet pilots differently depending on whether or not they had a CDTI. They were given more altitude and speed changes than were the pilots of the twin-engine aircraft, but half as many vectors. There were more messages on the whole with CDTI

than without for the pilots of the jets, whereas the reverse was true for the pilots of the light aircraft.

No differences in communications types or frequencies were found that could be attributed to the fidelity of the simulator. All of the differences found related to aircraft type or experimental condition. In general, ATC gave all pilots few speed commands when they were able to see the other aircraft on their displays, giving them instead clearances to follow an aircraft and to maintain separation "visually" using the CDTI. Position requests and responses were considerably reduced by the addition of CDTIs and VFR altitude information. Altitude changes, vectors, and clearances were not affected by display conditions across aircraft.

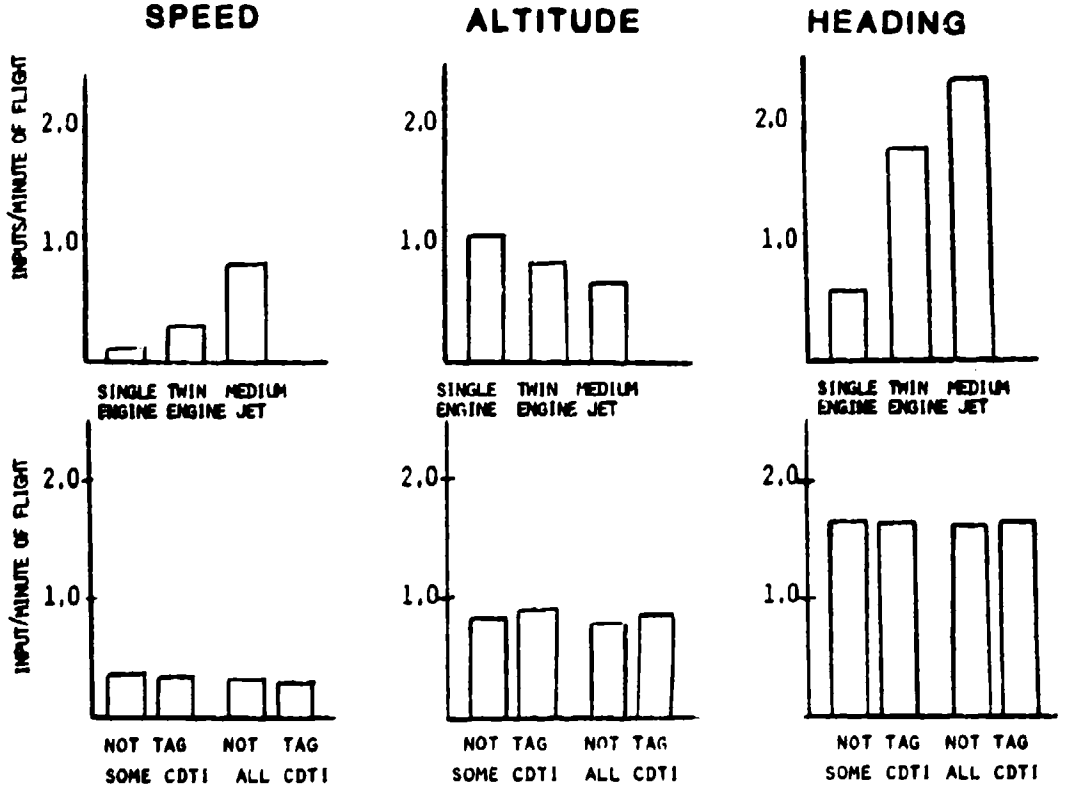
TABLE 3: Communications: Average number of messages transmitted to and from pilots of each aircraft type per group, per approach by aircraft type and experimental condition

	SINGLE ENGINE A/C (1 pilot/group)				TWIN ENGINE A/C (3 pilots/group)				MEDIUM JET (3 pilots/group)			
	SOMECDTI		ALLCDTI		SOMECDTI		ALLCDTI		SOMECDTI		ALLCDTI	
TOPIC OF MESSAGE	NO TAG	TAG	NO TAG	TAG	NO TAG	TAG	NO TAG	TAG	NO TAG	TAG	NO TAG	TAG
TRAFFIC	0.5	1.5	2.1	1.4	2.5	3.9	4.5	5.8	2.0	4.5	4.4	5.0
ALTITUDE	1.7	0.2	2.1	3.0	9.0	8.0	8.6	9.6	8.9	12.8	10.5	8.3
SPEED	0.5	1.7	0.2	0.6	6.7	8.6	8.6	6.6	8.2	8.7	8.7	8.5
CLEARANCE	1.7	2.2	3.0	3.2	10.5	8.4	8.9	11.2	9.4	10.7	9.1	9.4
VECTORS	1.2	2.5	1.6	3.0	10.1	7.7	6.9	7.9	3.1	5.3	4.1	2.0
POSITION	2.5	0.2	2.9	3.0	27.9	10.5	8.9	11.1	11.0	8.6	9.9	13.7
MAINTAIN	1.0	3.7	2.0	1.1	5.1	4.7	16.5	4.2	4.9	3.7	6.0	4.2
ROGER & READBACK	12.5	3.8	6.5	10.0	29.2	35.5	28.8	32.4	22.0	34.8	31.0	23.0
AVERAGE/ PILOT	21.6	15.8	20.4	25.3	33.6	29.1	30.5	29.6	23.1	26.3	27.9	25.7

Pilot Effort

Several measures were used to assess the amount of effort exerted by the pilots under different experimental conditions: number of SP throttle, flap, and trim changes and root mean squared aileron and elevator activity and PP speed, heading, and altitude changes. Each of the eight effort measures were analyzed with a three-way analysis of variance for repeated measures. A combined effort measure was then created by summing the measures transformed so the mean and standard deviations of each pilots' distributions of scores within each measure were zero and one respectively. The average effort scores were computer for SPs and PPs separately due to differences in the types of controls available to them. Other actions, such as radio frequency change, map scale or altitude range selection, and gear down that occurred only once or twice per approach were examined, but were not included in the combined measure because they contributed little to pilot workload and did not vary from trial to trial.

Figure 4: **PSEUDO-PILOT EFFORT: AVERAGE NUMBER OF CONTROL INPUTS PER MINUTE**



Pseudo Pilot Effort

Figure 4 depicts the frequency of speed, heading, and altitude changes made by PPs per minute of flight. These inputs occurred partly as a function of vectors and clearances given by ATC and partly as a function of the pilot's own initiative to achieve desired timing, rates, positions, and separation. There were no significant differences in number or type of control inputs for any of the experimental conditions for the raw frequencies, nor any significant interactions. The results were virtually identical for the standardized summary measure of PP effort. These results are not surprising as the PPs were only indirectly affected by the display conditions.

There was, however, a distinct difference in the control strategies used for the three aircraft types piloted by PPs. Few speed changes were made by any of the PPs, however, twice as many were made by the pilot of the jet (1.0 per minute of flight on the average) than by the pilots of the two general aviation-type aircraft (0.3 per minute of flight). Pilots of the single-engine aircraft made the greatest number of altitude changes (1.0 per minute of flight), but the fewest number of heading changes (0.5 per minute of flight). The pilot of the jet, on the other hand, made the fewest altitude changes (0.7 per minute of flight), and the greatest number of heading changes (2.3 per minute of flight). The many heading changes made by jet SPs represented control inputs initiated by the pilot to achieve the original flight plan, rather than numerous vectors or clearance changes from ATC.

Simulator Pilot Effort

One control activity that was of particular interest was the frequency and type of changes made by SPs in the altitude filter that controlled the proportion of other aircraft within the horizontal map range that were displayed. Pilots selected the range of +/-640 m (2000 ft) which had been identified as the optimal range in earlier studies (ref. 1) 60% of the time. They added an additional 610 or 1220 m to the range below their own aircraft the remaining 40% of the time. This increase was necessitated when they were given a clearance to follow an aircraft that had descended out of range of the narrower filter and had thus disappeared from view. On the average, SPs changed the relative altitude filter 1.5 times per approach. Somewhat more changes were made on trials in which VFR aircraft were presented with altitude information because these aircraft, like the IFR aircraft, were only displayed as long as they were within the selected range if their altitude was known. SPs rarely increased the range above their own altitude because aircraft that were above them were typically behind, and were thus of less interest. It is likely that increasing the range above would be of use in a departure environment, however departures were not simulated.

There were no significant differences among the measures of SP effort as a function of the two display conditions, although there was more aileron and throttle activity with a CDTI than without (Table 4), possibly reflecting fine tuning to maintain separation assumed by SPs when given clearances to follow an aircraft displayed on a CDTI.

Each of the effort scores for individual SPs were transformed and summed as described previously to produce a single measure of performance. A three-way analysis of variance for repeated measures was performed on these ratings. A significant interaction was found between presence or absence of VFR data tag information and replications ($F = 10.42$, $(1,11)$, $p<.05$). On earlier runs, fewer pilot inputs occurred with a VFR tag than without, whereas later in the experiment, more inputs were given. This could indicate that the pilots had learned how to accomplish the task with the different types of information and controls available to them and were using the information in flying their aircraft. Overall, fewer pilot inputs were recorded with VFR data tags than without, whereas significantly more inputs were recorded with CDTI than without ($F = 8.72$, $(1,11)$, $p<.05$), reflecting fine tuning by SPs with traffic displays.

TABLE 4: Individual measures of SP effort and a combined standardized measure of effort (Mean = 0; St. Dev. = 1)

Measure	Replications		CDTI		VFR Data Tag	
	1st	2nd	Some	All	noTAG	TAG
Aileron (rms)	.016	.017	.014	.018	.015	.018
Elevator (rms)	.016	.016	.016	.016	.016	.016
Throttle changes (number)	18.7	18.1	16.5	20.3	20.6	16.2
Trim actuations (number)	14.2	13.9	14.4	13.8	14.0	13.9
Flap changes	5.6	4.8	5.2	5.2	5.8	4.5
Combined standard effort score	-.009	-.003	-.073	.078	.001	.005

Simulator Pilot Opinion

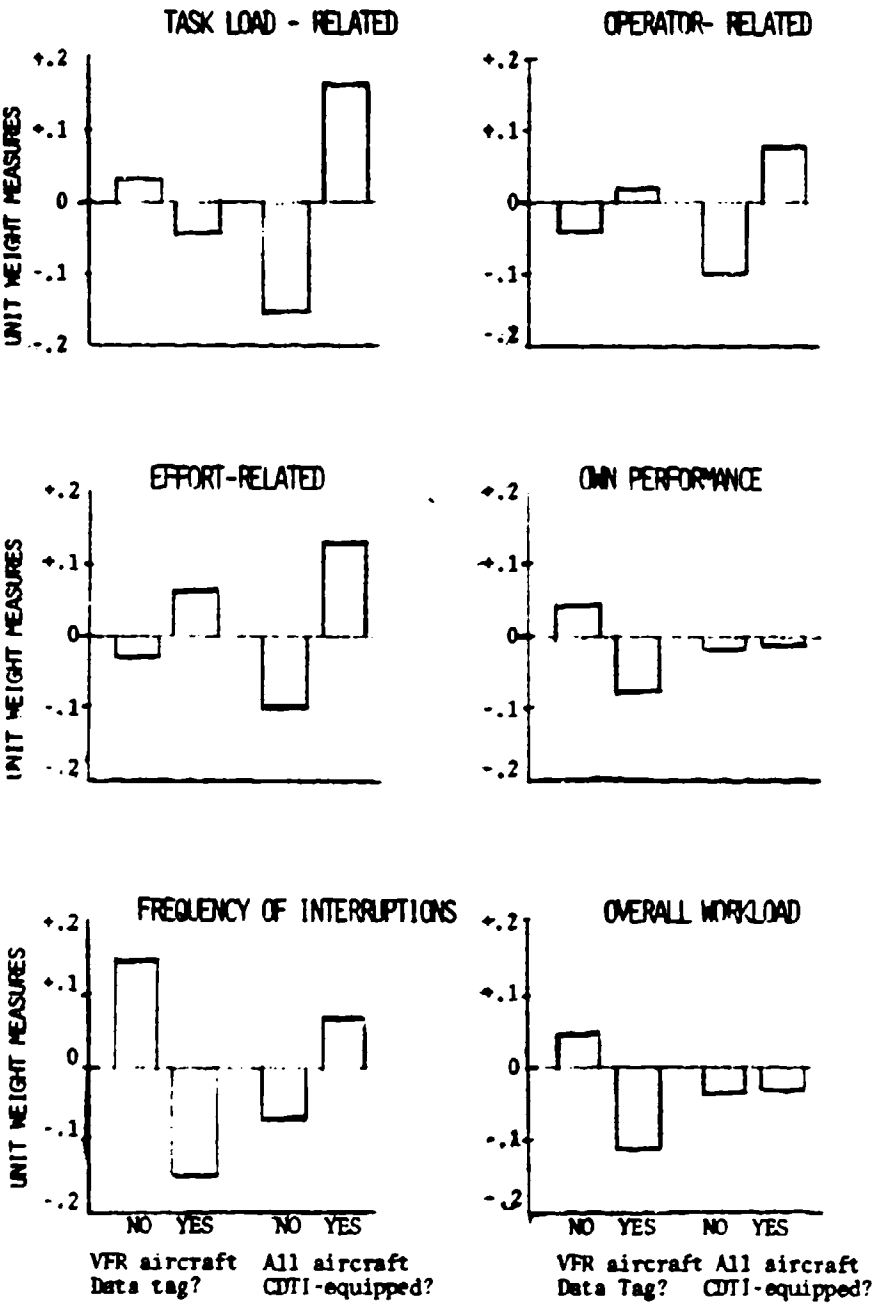
At the conclusion of each approach, SPs completed 17 rating scales. The first 14 scales represented relevant dimensions, such as fatigue level, perceived task demands, workload, etc. presented with labels at the extremes (e.g. lo/hi). Responses were given on a scale from one to seven, one representing a rating of "low" and seven a rating of "high" on whatever dimension was being rated. (Table 5) Fourteen three-way analyses of variance for repeated measures were performed to determine the effects of replications and the two display conditions. There were no significant differences in responses related to replicaitons. Significant increases in rated task difficulty ($F = 9.51$, $(1,11)$, $p < .05$), activity level ($F = 8.05$, $(1,11)$, $p<.05$), and required attention ($F = 7.77$, $(1,11)$, $p<.05$) were found when SPs were given traffic displays. This may reflect the increase in pilot effort and

TABLE 5: Responses to bipolar adjective scales by 12 simulator pilots. (A rating of 1 was very low, 7 very high)

	PRESENCE OF VFR ALTITUDE INFO.		CDTI EQUIPPAGE	
ITEM RATED	noTAG	TAG	someCDTI	allCDTI
TASK-RELATED				
Task difficulty	4.0	4.0	3.7	3.9
Task demands	4.3	4.1	4.0	4.4
Predictability	3.1	3.0	3.0	3.0
Task complexity	3.9	3.7	3.7	3.9
OPERATOR-RELATED				
Fatigue	3.2	4.1	3.1	3.2
Emotional stress	3.5	3.5	3.4	3.9
EFFORT-RELATED				
Activity	4.3	4.4	4.2	4.5
Attention	4.5	4.5	4.4	4.6
Monitoring	4.8	5.1	4.9	5.0
Planning	4.2	4.2	4.2	4.2
PERFORMANCE				
Skill required	4.2	4.2	4.1	4.3
Own performance	4.4	4.4	4.3	4.5
INTERRUPTIONS	2.8	2.5	2.5	2.7
OVERALL WORKLOAD	4.0	3.9	4.0	4.0

Figure 5: **SUMMARY OF PILOT RATINGS**

(STANDARDIZED SCORES: X=0; SD=1.0)



communications with the addition of CDTI. A significant increase in rated fatigue ($F = 6.34$, $(1,11)$, $p<.05$) and the amount of monitoring required ($F= 6.04$, $(1,11)$, $p<.05$) was found with the addition of VFR altitude information.

Responses to the 14 scales were grouped into six categories on the basis of meaning and obtained scores (Table 5). The ratings were then standardized so that individual subject's distributions of ratings within each scale had means of zero and standard deviations of one. These scores were then combined for the six categories. (Figure 5) Task-related, operator-related, and effort-related scales increased significantly ($F = 6.06$, 5.13, and 7.56 respectively, $(1,11)$, $p<.05$) with the addition of CDTI, however pilots' estimates of their own performance and workload were not significantly affected by the addition of a CDTI. Frequency of interruptions and overall workload were rated as decreasing non-significantly with the addition of VFR altitude information.

The last three scales required the pilot to estimate the number of interrupted tasks, the amount of visual, mental, and auditory free time available during the approach, and the likelihood of error. Three, three-way analyses of variance for repeated measures were performed and showed no significant differences among the reponses as a function of display condition or replications. All of the ratings were between 3.4 and 3.6 for the interruption scale, between 4.2 and 4.4 for the rated amount of visual, auditory , and mental free time, and between 3.7 and 3.9 for estimated likelihood of error out of a total range of 1 to 7.

At the conclusion of the experiment, pilots completed a debriefing similar to that used in earlier studies. They reported that CDTI assisted them in planning ahead and that they were more aware of the distance between their own and other aircraft with CDTI than during typical IFR flights with visual or radar separation. They reported that their workload was equivalent to that experienced in a typical flight and was not excessive. They considered the simulation to be adequately realistic for the type of research conducted, and that the types of tasks performed in the simulator would be about as difficult to perform in the aircraft they were accustomed to flying. The pilots responded that they would be willing to accept aircraft-following types of clearances, as received in this simulation, and to take responsibility for maintaining separation in line operations if the lead aircraft was in sight on a CDTI.

DISCUSSION

Many of the most interesting results of this simulation related to practical considerations associated with conducting research with multiple piloted simulators. A major concern in earlier studies (refs. 4, 8), was that ATC control strategies differed as a function of the level of realism of the simulator. This question could not be resolved, however, because aircraft type

(general aviation or transport) covaried with the the fidelity of the simulators. In the current study, different aircraft types were simulated with full pilot control (SP) or with more limited control (PP). The results indicated that control differences found in earlier studies occurred as a consequence of the type of aircraft simulated rather than the realism of the simulator, as long as the simulator was manned by a qualified pilot and communications with ATC were conducted realistically. ATC expedited the flow of transport traffic into the simulated airport in a manner representative of current operations, using speed and altitude control and giving vectors to single- and twin engine general-aviation-type aircraft to achieve desired sequencing and separation. Transport-type aircraft were rarely vectored off the route except to expedite their approach. The primary method of control used for them was the issuance of speed changes and clearances to follow or maintain separation behind aircraft in view on a CDTI.

This simulation provides further information about the importance of providing a realistic environment when analyzing simulator pilots' interactions with various CDTI configurations and the impact of such configuraitons on the ATC system. The PPs and controllers were included in this simulation to increase its realism for the SPs by providing communciations, strategies, errors, and decisions made by human operators. It appears that relatively simple displays, controls and aircraft models can be used to provide this dimension as long as the aircraft thus simulated move realistically on whatever display medium is used and are controlled by qualified pilots who sound authentic and react appropriately to ATC. The use of qualified controllers is another relatively simple but important way to place a simulation in an appropriate context for simulator pilots and to thus increase the validity of results.

As was found in the earlier study (ref. 4), no pilot ever took unilateral action as a consequence of information observed or inferred from the CDTI. SPs were willing to accept closer separation with a CDTI, but not to the point that the safety of flight was in jeopardy. They complied with ATC clearances to follow a lead aircraft or to maintain separation from an aircraft that was displayed on the CDTI. Because the CDTI provided more information, the possibility of pilot participation in some aspects of planning and achievement of control strategies developed by ATC was possible. Communications, particularly those related to traffic, increased with CDTI, and high frequency, low amplitude control activity to accomplish fine tuning increased. The addition of CDTI was also associated with an increase in the pilots' subjective ratings of task difficulty and activity, but not their perception of workload or the quality of their own performance. Even though the pilots experienced increased task demands and effort with CDTI, they rated their overall workload in this simulation as approximately the same as in a typical flight in response to questions on the debriefing. Furthermore, they indi-

cated that the addition of CDTI into line operations would ultimately result in fewer communications with ATC. They would be willing to accept aircraft-following types of clearances with a CDTI and to take responsibility for spacing between their own and a lead aircraft in sight on a traffic display "most of the time".

The results of this simulation provide additional verification of the opinions expressed by pilots in the earlier studies (refs. 1, 2) about content, format, and symbology for a CDTI after they had had a chance to experience many of the different features while conducting simulated approaches. Further research is clearly required, however, to address the remaining conceptual and practical issues. These issues include: (1) the impact of CDTI on procedures and regulations under different mixes of CDTI-equipped aircraft in controlled and uncontrolled airspace; (2) the relative quality of information obtained, retained, and used from various CDTI configurations; (3) pilot assurance, situation awareness, and ability to plan ahead with CDTI; and (4) practical considerations such as the effect of CDTI on fuel economy, airport throughput, and system safety. Answers to these and other questions are prerequisites of a decision to provide visual displays of traffic information in cockpits.

REFERENCES

1. Hart, S. G. & Wempe, T. E. Cockpit Display of Traffic Information: Airline Pilots' Opinions about Content, Symbology, and Format NASA TM-78601, (August, 1979)

2. Hart, S. G. & Loomis, L. L. Evaluation of the Potential Format and Content of a Cockpit Display of Traffic Information. Human Factors, 1980, 22(5), 591-604.

3. Palmer, E. A., Jago, S. J., Baty, D. L., & O'Connor, S. L. Perception of Horizontal Aircraft Separation on a Cockpit Display of Traffic Information. Human Factors, 1980, 22(5), 605-620.

4. Hart, S. G. Relative Altitude Information Display Requirements for a Horizontal Traffic Situation Display. Paper presented at the Seventeenth Annual Conference on Manual Control, University of California, Los Angeles, June 1981.

5. Heffley, R. K. & Schulman, T. M. Mathematical Model Descriptions for a Medium Jet Transport and a Light Twin. Mountain View, CA: Systems Technology Inc. Working Paper No. 1164-2R, June 1981. (Work performed under NASA Contract NAS 2-10611)

6. Heffley, R. K. & Schulman, T. M. Pseudo-pilot Mathematical Model Description for Basic Heading, Altitude, and Airspeed

Command. Mountain View, CA: Systems Technology Inc. Working Paper 1164-3, May 1981. (Work performed under NASA Contract NAS2-10611)

7. Airbus Industrie. Airbus Industrie Approach for Workload Analysis and Crew Complement Certification: Attachment 3. Toulouse, France: Airbus Industrie, 1981.

8. Chappell, S. L. Air Traffic Control of Simulated Aircraft With and Without Cockpit Traffic Displays. Paper presented at the Eighteenth Conference on Manual Control. Dayton, OH, May 1982.

AD P001648

AIR TRAFFIC CONTROL OF SIMULATED AIRCRAFT WITH AND WITHOUT COCKPIT TRAFFIC DISPLAYS*

Sheryl L. Chappell
Computer Sciences Corp.
NASA Ames Research Center
Moffett Field, CA

John G. Kreifeldt
Dept. of Engineering Design
Tufts University
Medford, MA

ABSTRACT

This study examines the air traffic control of air carrier and general aviation simulated aircraft with and without a cockpit display of traffic information (CDTI). Air traffic controllers provided spacing and sequencing for air carrier type simulators flown by airline pilots and small twin engine/single engine lower-fidelity simulators flown by general aviation pilots on an approach to landing. Some of these aircraft had CDTI, a graphic representation of the position of other aircraft relative to his/her aircraft and the navigational aids.

Differences in the air traffic control were found due to aircraft type and the presence of a CDTI. There was no difference due to CDTI in aircraft separation violations. Two air carrier aircraft were within one nautical mile lateral separation and 950 feet vertically for a longer period of time than two general aviation (GA) aircraft or an air carrier and GA aircraft. Two GA aircraft were within one mile longer than an air carrier and a GA aircraft.

The variance of intercrossing times at the middle marker between an aircraft and the previous aircraft was greater for aircraft with CDTI and for air carrier aircraft.

The number and type of controller/pilot communications were different for the conditions of simulator type and presence of CDTI. Of those communication types showing differences in frequency due to the presence of a CDTI, all except heading changes occured more often to and from an aircraft with a CDTI. For example, the controller gave more traffic advisories and had more questions for an aircraft with a CDTI. Those aircraft with CDTI had more communications of all types than those without CDTI. The airline aircraft in general, had more communications but the smaller aircraft had more communications concerning heading changes.

Air traffic control of the aircraft in this simulation showed similar characteristics to the current aviation system. For example, the air carrier aircraft were given speed and altitude changes while the GA aircraft were given

* Supported by NASA Grant NSG 2156 to Tufts University.

heading changes. The CDTI effected communications between the controller and the pilots reflecting the use of the CDTI information for aircraft separation.

INTRODUCTION

A cockpit display of traffic information is a display showing the pilot the position of other aircraft relative to his or her aircraft position. The display may include other information such as, ground speed and altitude of the other aircraft and navigational information. This and other studies explore the effects on safe and efficient air traffic flow associated with the use of a CDTI.

When the CDTI becomes available, it is likely that air carrier operators will be among the first to make use of this traffic information. General aviation (GA) aircraft will typically be much less likely to invest in this type of display. This creates a variety of capabilities within the air traffic control system. In this simulation there was a combination of aircraft types: general aviation single engine aircraft flown by GA pilots and air carrier aircraft flown by line pilots with a mix of CDTI capabilities. This study examines the effects of the aircraft type and the use of the cockpit display of traffic information on air traffic control.

METHODOLOGY

Subjects

There were three groups of subjects from the San Jose area. Each of the groups included one controller, three airline and four general aviation pilots. The subjects performing the air traffic control were former air traffic controllers from approach, tower, and center facilities. The nine line pilots were currently employed by major U.S. air carriers. The twelve general aviation, instrument-rated pilots had a variety of experience levels.

Equipment

The study was conducted in the Multicockpit Facility at NASA Ames Research Center. The facility had seven aircraft under air traffic control (ATC). There were three single-pilot air carrier simulators based on the Boeing 727 with simplified flight, navigation, and engine instruments and controls. The four general aviation simulators were modeled on the Cessna 310 and 172. The simulators consisted of a

joystick for the control yoke and a sliding potentiometer for the throttle. The GA pilots had digital information of their current and commanded speed, heading and altitude. An intercom was used for the radio transmissions between the pilots and the controller.

The cockpit displays of traffic information used in this study were presented on a cathode ray tube. The controller's station had a display similar to that found at an approach/departure facility (Figure 1).

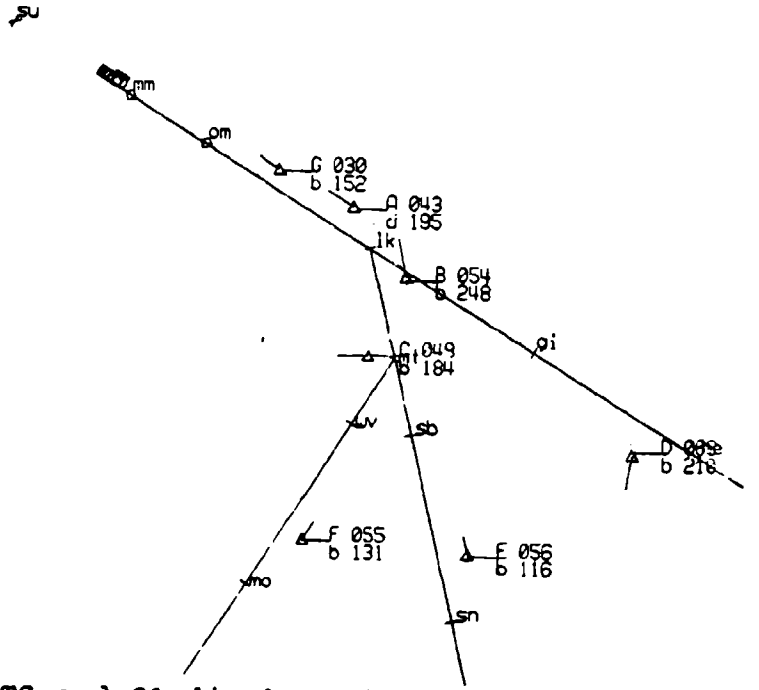


Figure 1. ATC and GA display of traffic and navigational information.

The air traffic controller had a north-up display of all aircraft at a 60 nautical mile map range. The position of the aircraft were shown by a triangle. The aircraft moved across the display relative to the navigational information in the San Jose area.

There was a two-segment flight path predictor line in front of each aircraft symbol which showed where the aircraft would be in 60 seconds, if the speed, heading, and rate of turn were maintained for that period of time. The length of the predictor varied with ground speed since it reflected the distance traveled in 60 seconds. There were history dots behind the aircraft showing the previous positions at four second intervals for the last thirty two seconds. Each aircraft had a data tag showing the aircraft identification, and the altitude (in hundreds of feet) on

the first line, and the ground speed (in knots) and an arrow which indicated whether the aircraft was climbing or descending (vertical velocity greater than 100 feet/minute) on the second line.

This display was identical to the display shared by the four general aviation pilots and was updated every four seconds. Since the GA pilots required the CDTI information to make their flight, the conditions of no CDTI consisted of them ignoring the other traffic on the display. They did so very effectively.

The display for the air carrier simulators was always heading up (Figure 2). The ownship symbol remained stationary on the display horizontally centered and vertically one third from the bottom. The airway appeared to pass under the ownship symbol.

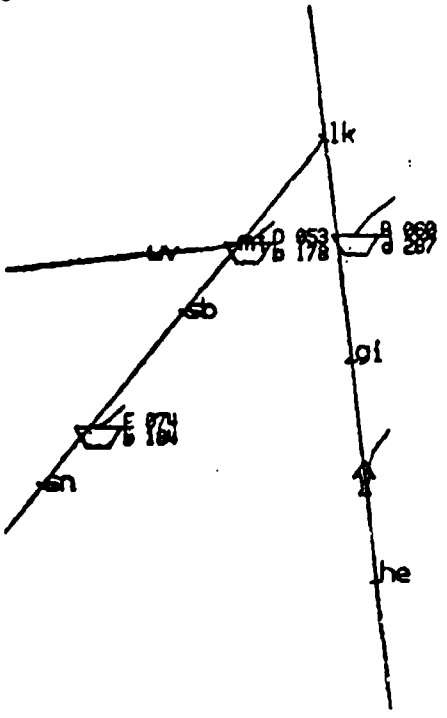


Figure 2. Air Carrier CDTI showing ownship and map information.

The map scale was pilot-selectable to 2, 4, 8, 16, 32 and 64 nautical miles (nm). It was also possible to select an automatic scale that varied from 64 to 2 miles with the aircraft's altitude. All aircraft +/- 2000 feet in altitude from the pilot's aircraft were shown, in addition the pilots were able to selectively include aircraft +/- 4000 and 6000 feet.

The traffic on the air carrier simulator CDTI's was

shown by a hexagon if the aircraft was within 500 feet of ownship. Only the top half of the hexagon showed if the aircraft was more than 500 feet above and the bottom if more than 500 feet below. For the no CDTI condition, the display showed the position of ownship only relative to the navigational map.

The flight and engine instrument information was updated every one tenth of a second. The ownship information on the CDTI including the navigational information, was updated five times a second. The information on the other aircraft was only updated every four seconds. No radar errors were simulated in this study.

Procedure

The subjects were paid for the two days they participated. Initially, the controller and pilots were briefed on the experiment and the simulation. The controller's task was to provide spacing and sequencing for all seven aircraft on the approach and landing. During the first few minutes of a run, all the aircraft became airborne (randomly placed) at the navigational fixes south of the San Jose Airport. From this position the pilots flew their simulators under ATC control according to the published ILS approach procedures and landed at San Jose Airport. After an aircraft landed it was repositioned (with equal probability) at one of the three outer most positions, 35 miles from the airport. All the aircraft remained on an instrument flight plan with normal radio communications carried out with the controller. The airline pilots flew the simulators manually. The GA pilots provided autopilot inputs to their aircraft until the outer marker where an autopilot took control of their aircraft for the landing.

Ten minutes after the run began, the airport was closed and the aircraft were put into holding patterns until the airport reopened five minutes later. Each run lasted thirty minutes. There was a 25 knot wind above 2000 feet decreasing linearly to 15 knots on the ground. No turbulence was simulated.

Training - The first three flights were flown without a CDTI to familiarize the pilots with the simulators. The forth and fifth flights were made with the CDTI available but pilots were not required to use the information for separation. For the sixth and seventh flights, only two air carrier aircraft had CDTI.

Experimental Flights - There were six experimental flights. The first two were made with no CDTI (i.e.

navigational information only) for the air carrier simulators. The GA pilots had traffic information but were asked to ignore the other aircraft and follow only the ATC commands for heading, speed, and altitude. For the next two flights, all seven aircraft had a CDTI. The final two flights were made with only two air carrier simulators having CDTI. Therefore, the number of CDTI aircraft varied from zero to seven. At the end of each flight, time was taken for discussion and questions.

RESULTS

There are several areas of interest when exploring the air traffic control of aircraft with a CDTI. The foremost question is safety. Does the implementation of the CDTI increase or decrease the overall safety of air travel? More specifically, is the separation of aircraft as good or better when the aircraft involved have a traffic display? Among other concerns is efficiency. If the aircraft are too far apart in landing, fewer aircraft can land in the same amount of time, therefore, there is a decrease in the capacity of the air traffic system. Another good indication of the characteristics of the ATC system is the communication between the controller and the pilots. The above measures were examined by 1) presence of CDTI, 2) type of simulator, and 3) number of CDTI-equipped aircraft being controlled.

Aircraft Spacing

There were two variables of aircraft spacing used to evaluate the air traffic control system: aircraft spacing violations and aircraft intercrossing times at the outer and middle marker on final approach.

Aircraft spacing deviations were measured in seconds during which two aircraft were too close, i.e., within 1, 2, and 2.5 nautical miles laterally with vertical separation of less than 950 feet. There were no effects on aircraft spacing violations due to the presence or absence of a CDTI in the simulator. No effects were found due to the number of aircraft with CDTI. There were differences in the separation errors by simulator type. Two air carrier simulators were within one mile separation for a longer duration than an air carrier and GA aircraft ($t=-4.93$, $df=18$, $p<.001$) or two GA aircraft ($t=3.46$, $df=18$, $p<.01$). GA aircraft were in violation at 1 nm separation for a longer time than an air carrier and GA aircraft ($t=2.20$, $df=18$, $p<.05$). At 2 nm there were no differences due to simulator type. Two GA aircraft were within 2.5 nm longer than an airline and GA aircraft ($t=2.02$, $df=18$, $p<.05$).

The number of seconds between an aircraft and the previous aircraft when crossing the outer or middle marker was not significant for any condition. However, the variance between the intercrossing times was significantly larger for aircraft with CDTI than aircraft without (Levene F=9.29, df=158, p<.01). Air carrier simulators had greater variance of intercrossing times with the aircraft ahead of them at the middle marker than GA aircraft (F=6.49, df=158, p<.01).

Communications

The communications between the controllers and pilots were categorized. An aircraft with CDTI had more total communications to and from ATC than an aircraft without a CDTI (Mann-Whitney z=4.61, df=2518, p<.001). ATC gave more clearances (z=-2.51, df=164, p<.01), traffic advisories (z=-8.43, df=164, p<.001), clearances to follow another aircraft (z=-10.84, df=164, p<.001) and had more questions (z=-2.63, df=164, p<.05) for aircraft with CDTI than aircraft without CDTI. The controller gave more heading changes to aircraft without CDTI (z=-2.51, df=164, p<.05). Pilots asked more questions (z=-2.54, df=164, p<.05) and made more acknowledgments (z=-6.85, df=164, p<.001) when they had a CDTI in their simulators.

The total number of communications went up proportionally with the number of CDTI-equipped aircraft (7,2 or 0) on an experimental run (F=4.12, df=1887, p<.05). ATC gave more traffic advisories (F=16.97, df=123, p<.001) and clearances to follow another aircraft (F=47.15, df=123, p<.001) as the number of CDTI-equipped aircraft increased. Pilot communications were also greater by the number of CDTI-equipped aircraft. Pilots asked more questions (F=3.65, df=123, p<.05) and made more acknowledgments (F=12.88, df=123, p<.001).

Air traffic control had more communications, of the following type, with the airline simulators: clearances (Mann-Whitney z=-7.45, df=124, p<.001), speed changes (z=-6.85, df=124, p<.001), altitude changes (z=-4.28, df=124, p<.001), questions (z=-4.17, df=124, p<.001), traffic advisories (z=-3.87, df=124, p<.001), and clearances to follow another aircraft (z=-3.20, df=124, p<.01). The controller gave more heading changes to the GA aircraft (z=-3.96, df=124, p< .001). The pilots of air carrier simulators had more acknowledgments (z=-4.41, df=124, p<.001), and readbacks (z=-4.36, df=124, p<.001). There were more total communications by or with air carrier simulators (z=-5.86, df=1888, p<.001). Communication categories showing no effect were position reports and readbacks.

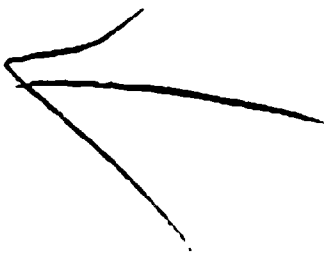
None of the dependent measures showed learning effects across trials.

DISCUSSION

The findings of this study are not surprising. Basically the air traffic controllers handled the simulators as they would real aircraft, i.e. the controllers gave altitude and speed changes to the air carrier aircraft and heading changes to the general aviation aircraft.

The presence of a CDTI yielded no effect on separation violations which demonstrates no effect on safety due to the CDTI in the aircraft. There is larger variance in intercrossing times at the middle marker therefore, there is less consistency with CDTI-equipped aircraft. This may be tolerable since there is more information in the system, i.e. CDTI-equipped aircraft have redundant ATC traffic information.

The differences in the types of communications due to the presence of a CDTI and the number of CDTI-equipped aircraft suggest the controllers are using the information in the traffic display to separate and sequence the aircraft. The clearances to follow another aircraft, for example, require the pilot to maintain visual contact with the aircraft and provide separation. These clearances were given on an average of 1.35 times per CDTI-equipped aircraft per flight. When an aircraft did not have a CDTI, there was not sufficient information to assume responsibility for separation from the preceding aircraft. As both controllers and pilots become more familiar with the CDTI, there will likely be additional differences in the communications which demonstrate other uses of the cockpit display of traffic information.



LIST OF AUTHORS

18TH ANNUAL CONFERENCE ON MANUAL CONTROL

AUTHORS

Minoru Abe 366

Gyan C. Agarwal 58, 77

Bart J. Bacon 256

Sheldon Baron 186

Arthur Beckman 267

A. K. Bejczy 440

R. Beveridge 345

Tom Bösser 5

J. W. Brown 440

Sheryl L. Chappell 545

Mary E. Childress 127

Kevin Citurs 474

Margaret M. Clarke 413

Edward M. Connelly 119

Sidney A. Connor 150

R. Cortilla 345

Kathleen M. Craig 336

D. Francis Crane 166

M. J. Crisp 475

V. Dea 225

Joseph C. De Maio 108

Ron S. Dotson 440

Sherry L. Dunbar 345, 466

Elliot E. Entin 263

John M. Evans 474

J. W. Frazier 427

Richard T. Gill 100

Gerald L. Gottlieb 58, 77

Sol W. Gully 268

Stephen Handel 413

Mark D. Hansen 320

Sandra G. Hart 127, 522

E. James Hartzell 345

Charles P. Hatsell 3

Jan R. Hauser 127

Ronald A. Hess 267

K. E. Hudson 427

Hideo Iguchi 366

Robert J. Jaeger 77

Richard J. Jagacinski 41

Richard S. Jensen 327

Andrew Junker 99

James R. Kelly 485

David L. Kleinman 263, 268

Kiyoshi Komoriya 366

Jonathon Korn 268

John G. Kreifeldt 389, 413, 545

William C. Krenz 42

Stephen H. Levine 389

William H. Levison 239, 419, 427

AUTHORS - CONCLUDED

J. L. Lewis 440

Jeffrey L. Maresh 327

Edward A. Martin 165

Michael W. McGreevy 514

Grant R. McMillan 165

Patricia R. Mineer 3

Christine M. Mitchell 293

Hitoshi Miura 58

Don L. Monk 41

Neville Moray 120

Ramal Muralidharan 419

Alex Netch 225

B. O'Lear 427

Roy E. Olsen 420

Krishna R. Pattipati 263

L. R. Perkin 475

Lee H. Person, Jr 365

Lloyd D. Reid 168

Daniel W. Repperger 427

Kanji Sahara 225

Yoshihiro Sakai 366

Nils R. Sandell, Jr 263

David K. Schmidt 19, 256

V. Skowronski 427

Eric N. Solowka 168

Lawrence W. Stark 42

Fuchuan Sun 42

Susumu Tachi 366

Masao Takabe 366

Kazuo Tanie 366

Blaza Toman 207

K. Waterton 120

Sharan L. Ward 207

Kwang C. Wei 207

Joseph Whitmeyer 207

Christopher Wickens 100

Walter W. Wierwille 150

David H. Williams 485

Glen F. Wilson 99

Joseph G. Wohl 263

Yung-Koh Yin 225, 475

*U.S.Government Printing Office: 1983 — 659-006/3031