Nineteenth Annual Conference on Manual Control

May 23-25, 1983
Massachusetts Institute of Technology
Cambridge, Massachusetts

Sponsored by
NASA Ames Research Center
Moffett Field, California

NCC2-34456

Proceedings of the

NINETEENTH ANNUAL CONFERENCE ON MANUAL CONTROL

held at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Cambridge, Massachusetts 02139

May 23-25, 1983

sponsored by

NASA AMES RESEARCH CENTER
Moffet Field, California

John Stewart, Project Manager

Conference General Chairman: William H. Levison
Bolt, Beranek and Newman, Inc.
Cambridge, Massachusetts

Conference Arrangements: Thomas Sheridan
Massachusetts Institute of Technology
Cambridge, Massachusetts

Conference Publications: Robert Kenyon
Massachusetts Institute of Technology
Cambridge, Massachusetts

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

Abstract of a formal paper submitted for presentation at the Annual Conference on Manual Control, May 23 - 25, 1983.

SUBJECTIVE SCALES FOR WORKLOAD EVALUATION: CRITICAL ASPECTS AND NEW DIRECTIONS FOR RESEARCH

Mary E. Childress*
San Jose State University

As aircraft and other mechanical systems increase in complexity and rely more heavily on computerization of function, and as the pilot or other operator assumes greater supervisory responsibility for system monitoring and control, need for evaluation of the workload associated with system changes increases. Many of the methods currently available, though helpful in specific situations and often necessary in promoting understanding of some basic processes, are often difficult and unwieldy to use in complex, practical situations. Subjective rating scales, however, are convenient instruments for evaluating this workload and for estimating the magnitude of changes in load as system changes occur. The use of such scales has historical precedent in the personnel literature, particularly in performance evaluation. Subjective scales also have been used to evaluate specific system characteristics, such as aircraft handling qualities. The utility of the method is clear; however, psychometric development of subjective scales for the evaluation of workload currently is in its infancy. Thus, though the literature is replete with examples of and recommendations for their use as well as with criticisms of their deficiencies, research directed towards examination of their properties, and evaluation of the conditions under which their use is appropriate and obtained results generalizable is just beginning. Several important works (e.g., Landy & Farr, 1980; Moray, 1979; Nisbett & Wilson, 1977; Wherry, 1950, 1952) have described the problems associated with subjective ratings, have detailed some of the situations in which they may be appropriate, and have recommended specific topics for future research. This paper presents a review of critical aspects of that literature which suggest directions for future research relative to self-ratings of subjective workload. It provides examples of some recent work at Ames Research Center which has suggested extending the basic input-processing-outcome model for examining workload to consider all input sources and the related outcomes, and it details current work based on that model.

---

*Research Associate at NASA/Ames Research Center under NASA Cooperative Agreement NCC 2-34 with San Jose State University Foundation, Ms. Sandra Hart, Grant Monitor.

REFERENCES

Landy, F. J., & Farr, J. L. Performance rating. Psychological Bulletin, 1980, 87, 72-107.

Moray, N. (Ed.). Mental workload: Its theory and measurement. (Proceedings of the NATO Symposium on Theory and Measurement of Mental Workload, Mati, Greece, August 30-September 6, 1977.) New York: Plenum Press, 1979.

Nisbett, R. E., & Wilson, T. D. Telling more than we can know: Verbal reports on mental processes. Psychological Review, 1977, 84, 231-259.

Wherry, R. J. Control of bias in rating: Survey of the literature. (Technical Report D-49-0853 OSA 69). Washington, D. C.: Department of the Army, Personnel Research Section, September 1950.

Wherry, R. J. The control of bias in rating: A theory of rating. (Personnel Research Board Report 922). Washington, D. C.: Department of the Army, Personnel Research Section, February 1952.

Subjective Workload Experienced During Pursuit Tracking as a Function of Available Information

Jan R. Hauser
Arizona State University
Tempe, AZ

Sandra G. Hart
NASA-Ames Research Center
Moffett Field, CA

ABSTRACT

Twelve general aviation pilots performed a pursuit tracking task where the objective was to "pilot" the pursuit vehicle (a simplified delta wing aircraft) after a target (represented by a cross). Successful acquisition of the target was always displayed on the screen. Each subject experienced twenty 10-minute experimental runs in a partially counterbalanced order. The experimental variables included: (1) number of dimensions (2 or 3), (2) target path complexity (low or high), and (3) availability of information (both target and pursuit vehicle were displayed for either 100%, 50%, 25% of the time, or at subject command). Subjects controlled the vehicle by pressing rocker arm switches for the functions of yaw (left and right), roll (left and right), speed (acceleration and deceleration), pitch (up and down), and screen illumination (for the one condition where subject control was given to display time). After every experimental trial, subjects rated the preceding experience using a set of 15 bi-polar adjective scales. Significant differences were found for the difficulty and display time variables on the majority of the scales, and for the dimension variable on eight of the scales. Strong relationships were found between actual time-on-target with ratings of overall workload, performance, and difficulty for almost all subjects, who were also able to estimate their time-on-target with a high degree of accuracy.

INFLUENCE OF PILOT WORKLOAD AND TRAFFIC INFORMATION ON PILOT'S SITUATION AWARENESS

Sandra G. Hart
NASA-Ames Research Center
Moffett Field, CA

Sheryl L. Chappell
Informatics
Palo Alto, CA

ABSTRACT

Although it seems intuitively obvious that the addition of a cockpit display of traffic information (CDTI) should enhance a pilot's awareness of the current and projected situation of other aircraft, it has not been empirically determined that such is the case. Furthermore, there is some question about the utility of CDTI under conditions of relatively high pilot workload; when pilots become busy they may ignore the CDTI or take unilateral actions based on incompletely understood information. The current simulation was designed to determine how much information pilots could recall about eight aircraft simultaneously participating in a simulated approach task. A stop-action technique was used, so that each approach sequence was terminated at some point and the participants completed a written debriefing describing their recall of aircrafts' positions, situations, and intentions. The experimental variables included: (1) presence or absence of CDTI; (2) CDTI quality; and (3) level of concurrent workload. Four groups each consisting of three transport pilots, four instrument-rated general aviation pilots and one controller participated in the experiment. Concurrent workload but not CDTI quality or presence significantly affected the type and amount of information remembered. Rated workload and several types of communications were increased by the addition of CDTI and by the experimental manipulations intended to increase workload. The pilots reported feeling that CDTI afforded them a better understanding of the traffic situation, but this subjective impression was not supported by an improvement in the amount of information recalled.

INTRODUCTION

During the past ten years, considerable effort has been devoted to upgrading the air traffic control (ATC) system. This has occurred because of the need to replace aging equipment and to meet the demands of increased traffic density. The technology explosion during the same period has made it feasible to provide information to airborne crews that has been previously available to ground-based operators only. Not only is more information available in the cockpit, but transmission and presentation alternatives have broadened considerably. One of the systems that has been selected for implementation is the discrete beacon system (DABS). It will provide direct data link communications between individual aircraft and ATC to relieve radio frequency congestion.

One potential drawback of this system is that pilots would no longer overhear information about the status and intentions of adjacent aircraft as they now do on the radio frequency "party line". It has been suggested that traffic displays provided by digital data link could replace the lost information. In addition to assuring pilots about the continued safe conduct of the ATC system, the information displayed about the traffic situation might also enable pilots to assume airborne responsibility for spacing (ref. 1) or local separation, sequencing and traffic management (refs. 2,3,4). One possible benefit of such a modification in the division of responsibility between the air and ground might be to allow reduced separation under instrument flight rules (IFR) to approach that allowed under visual flight rules (VFR). This could improve system efficiency with no reduction in safety. Some of the potential liabilities of a CDTI include the possibility that pilot workload would be increased to unacceptable levels and that pilots might take inappropriate unilateral actions based on their perception of a situation observed on a CDTI.

The basic idea of presenting a graphic traffic display in the cockpits of aircraft is not new, in fact it was first proposed by RCA in 1946. Subsequent advances in computerized radar tracking systems and computer-generated displays have made the concept practically feasible. Increased traffic density and complexity have made it more compelling. Considerable research has been conducted to determine the impact of CDTI on the division of responsiblity between the air and ground (ref. 2), the perceptual and control problems associated with the use of CDTI to perform certain tasks such as intrail following (ref. 1), display content, format and symbology (ref. 5), and the display features that pilots actually look at (ref. 6). It is clear from these and other studies that pilots would like to have CDTIs and that they feel their performance is improved by them, even though traffic displays appear to increase their workload (ref. 7).

The completeness and correctness of a pilot's internal model of the environment through which he is passing depends on: the quality and format of the information available, its perceived relevance, and the amount of attention that the pilot can (or will) devote to situation awareness. A pictorial situation display could present information about navigation, terrain, and weather as well as the traffic situation. Much of the information that pilots now receive from a variety of sources could be combined in one place with a common format and scale. In addition, this display medium has the advantages of currency, immediacy, and availability. The disadvantages may include information overload and pilot complacency: pilots may become less aware of the surrounding situation. Because the information is available, they may not maintain a mental representation of their environment or remember information about the position or status of other aircraft. Finally, information via the audio channel occurs infrequently and only when a change has or will occur, thereby highlighting and drawing pilot's attention to potentially important events. This may be contrasted to a continuously changing visual display that depicts the ongoing situation rather than discrete changes or warnings, but must be consciously looked at to obtain information.

Display clutter may be a serious problem with a visually presented traffic display. So much information is available for display that intelligent selectivity must be exercised to optimally present just that which is relevant, and no more. However, it is still not clear just what additional

information a traffic situation display can or should provide to enhance pilot awareness and memory of the position, status, and intentions of other aircraft in the current ATC system. Placing the burden on the pilot by requiring him to select among a multitude of options is simply begging the question. A second issue that logically follows the first, is whether a CDTI could replace the "party line" source of information about other aircraft. Furthermore, it is likely that the amount of information that is extracted and remembered by a pilot will vary as a function of the duties that he is expected to perform with a CDTI and the content and format of the information presented. Thus, the environment into which CDTIs are to be introduced, the role that pilots will play, the display quality, and the impact of other attention-demanding concurrent tasks must be considered in answering the question, "What do pilots get from a CDTI?".

Melanson, Curry, Howell, and Connelly (ref. 8) addressed this issue using a stop-action quiz technique to determine pilots' knowledge about the position, altitude, speed, heading, rate of climb, identity, and landing sequence number of other aircraft. Their experiment involved one piloted simulator that conducted a series of instrument approaches in sequence with four other computer-generated aircraft. Communications were simulated by pre-recorded messages. The variables of interest were presence or absence of a traffic display, presence or absence of a "party-line" communications frequency, and level of conflict situations. It was found that the visual display was superior for detecting potential conflicts but that performing a concurrent spacing task limited pilot's attention to the aircraft being followed, thereby reducing their ability to detect problems with other aircraft. The addition of a traffic display did not significantly improve pilot's memory of identifications, altitudes, or speeds but did enhance their ability to recall positions of other aircraft. Workload, as measured by a light-cancellation side-task, was significantly impacted by display and communications conditions. This study provided the first information about how much pilots recall from a CDTI. However, the conclusions to be drawn were limited by the realism of the task; the "traffic" was computer-generated, their number limited, and the scenarios stylized.

The current study examined pilot's memory of the traffic situation under a variety of experimental conditions using a stop-action technique similar to that used earlier (ref. 8). The utility of visually presented information during the flight was also assessed with measures of aircraft control, traffic flow efficiency and safety, communications, pilot opinion, and information recall. The primary goal was to determine what additional information a traffic display provides about the status, position, and intentions of a pilot's own and other aircraft when it is superimposed on current ATC communications and procedures. This question is of considerable practical importance in specifying the critical and useful components of traffic situation displays (to reduce display clutter) and in determining which information might be best provided by radio and which by visual displays.

The experiment was conducted in a facility that provides realistic communications and decision-making behaviors by pilots and controllers by integrating seven piloted simulators flying instrument approaches under ATC control. It was anticipated that pilots would be motivated to remember

more information about other aircraft in this simulation than in the previous one because of the increased immediacy and uncertainty associated with piloted simulators as compared to computer-generated targets and because the traffic mix was realistically complex and varied. The quality of the traffic displays, the mix of CDTI-equiped and unequiped aircraft, aircraft types, and approach difficulty were manipulated to investigate the influence of different variables on pilot's memory of the traffic situation. Two different levels of workload were imposed by manipulating the simulation scenarios and by imposing additional tasks on the pilots. Scenario-related workload was varied by: introducing VFR flights that created different levels of conflict with IFR traffic, interrupting radio communications, disrupting the flow of traffic by temporarily closing the airport, and by pilot-declared emergencies. In addition, some of the pilots were required to perform a concurrent rating task on high workload trials. Although this task was not particularly demanding by itself, it did divert the pilots' attention from monitoring the traffic display. It was expected that traffic awareness would be reduced on the high workload trials, and that measures of control and communications activity, and concurrent and post-approach workload ratings would reflect the increased difficulty of the approaches.

METHOD

The experiment was conducted in the multi-cockpit facility at NASA-Ames Research Center, which has been described in detail elsewhere (refs. 7,9). Therefore, this section will focus on differences between past and current applications rather than provide a complete description. The facility provides an ATC station for two controllers, seven piloted simulators, and one computer-generated aircraft. The ATC station contains two communications setups and a 58 cm black and white monitor. The monitor depicts significant navigation features in a 84 km area surrounding the San Jose Airport and all approaching aircraft. (Figure 1) Three aircraft types are simulated for variety and realism (a medium jet, a light twin-engine aircraft, and a single-engine aircraft) with three levels of pilot control: discretely controlled autopiloted targets, manually controlled simulator cabs, and computer-generated targets. Four pseudo-pilot (PP) stations provide communications capabilities and discrete control over aircraft flying instrument approaches on autopilot. (The term "pseudo-pilot" reflects the limited quality of the controls and displays.) Heading, altitude, and speed changes are entered by discrete selections of digital command values effected by movement of a joystick and slide pot. The PPs share a single 58-cm horizontal traffic display depicting the same area as the ATC display (Figure 1) located 1.3 m from the pilots who are seated in a semi-circle around it. Three simple single-seat simulators with a stick, throttle, speed brake, flaps, gear, and radio are provided for the simulator pilots (SPs). Information necessary to control the simulated aircraft within a flight-envelope from cruise to landing is displayed on the top half of a 58-cm black and white monitor located 1 m in front of the pilot. (Figure 2) A 13-cm horizontal situation display located at the bottom of the monitor presents relevant navigation information (e.g. routes, navigation aids, and airports). When applicable, traffic information is superimposed on this display. Previously flown, regenerated VFR

flights typical of the San Jose area are included to increase traffic density and allow the programmed introduction of specific flights into the scenarios at critical times.

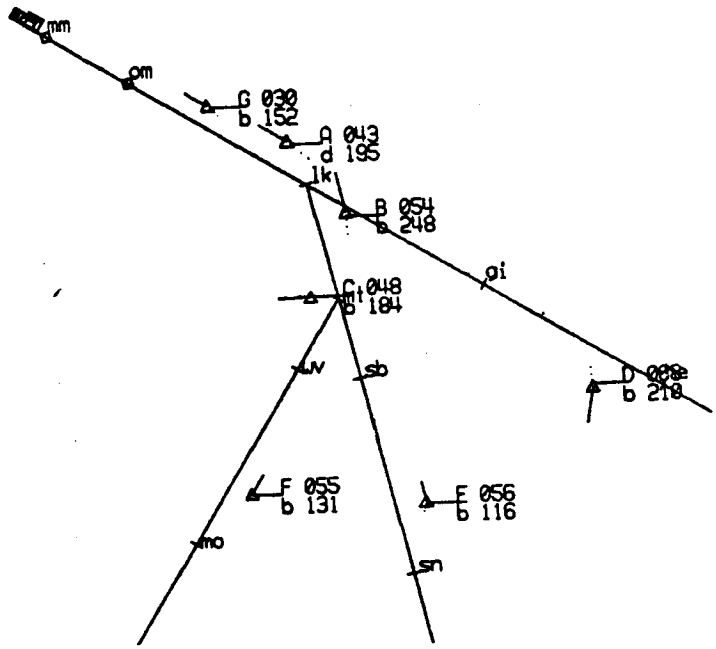


Figure 1: ATC and Pseudo Pilot traffic display.

Simulated Environment

Each experimental trial consisted of seven piloted approaches to the San Jose Municipal Airport from the south and west and one VFR "intruder" flying a variety of paths. Most approaches were initiated 45-65 km from the airport and were terminated when the majority of the aircraft were within 8-25 km of the field. Different wind conditions increased the realism. Current ATC procedures for IFR flight were followed throughout.

Scenarios

Sixteen scenarios similar to those used in a previous experiment (ref. 7) were used; four for training and twelve for the experimental trials. Aircraft under traffic control entered the terminal area with with adequate separation, simulating handoffs from an enroute controller. VFR aircraft entered, flew around, and left the area beneath and south of controlled airspace, flying patterns typical of general aviation traffic in the San Jose area. Each scenario was experienced once per group; however each time it was used, it was presented with a different combination of CDTI equipage (some or all aircraft) and workload (LO or HI).

Aircraft mix

Because one of the goals of the study was to determine how much information was remembered about the traffic situation, aircraft type (jet, single-engine, or twin aircraft) and call sign were varied from trial to trial for the PPs because the control requirements of their simulators were simple enough to allow such diversity. SP simulators required continuous

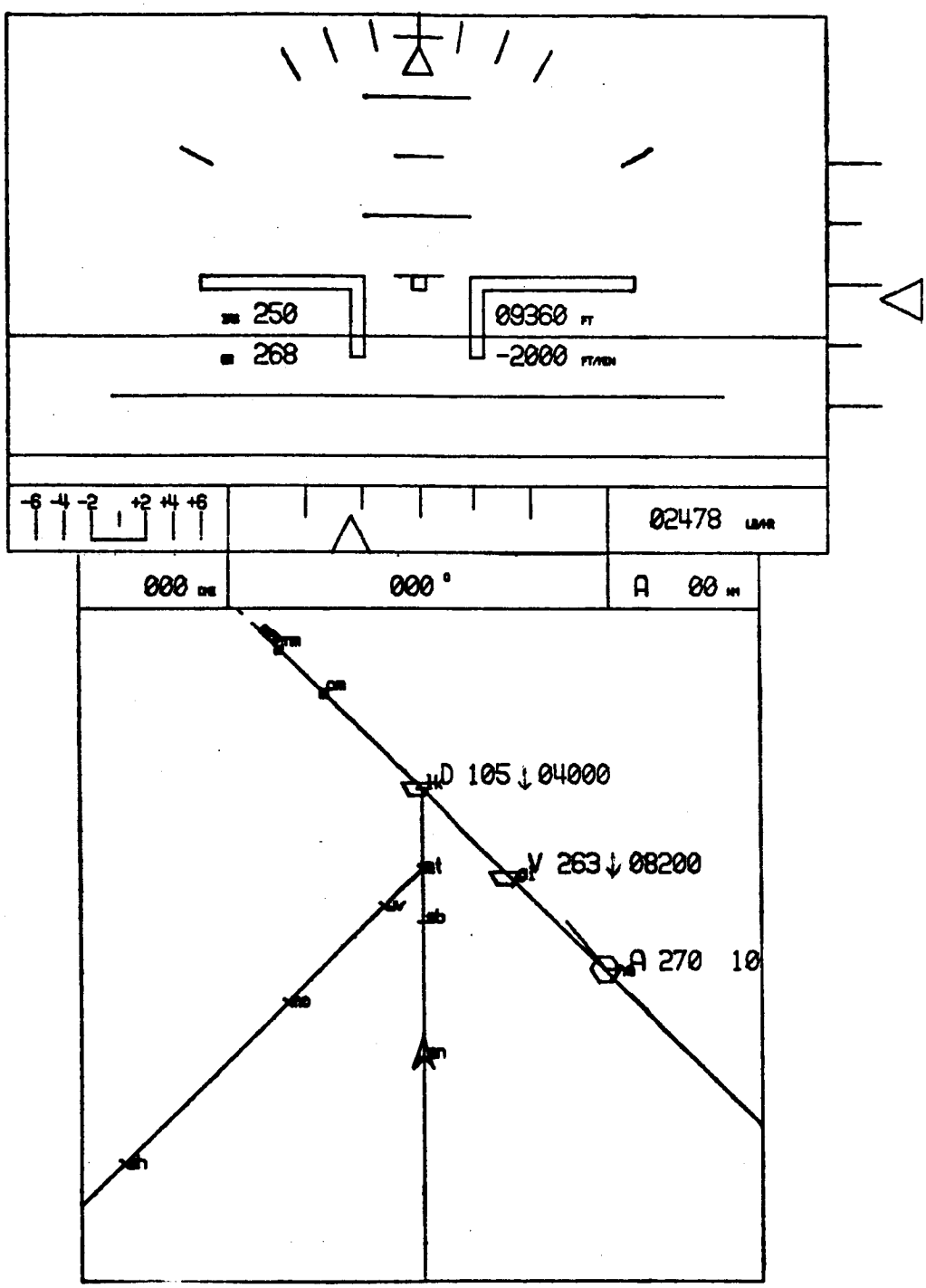


Figure 1: Simulator pilot display including traffic information

manual control, thus aircraft type was not varied for them. Two SPs flew a simulated medium jet and one a light twin aircraft. The first letter of the simulator's call sign (e.g. A, B, C,.....G) remained constant throughout, however the numeric portion was varied from trial to trial (e.g AA123 one trial might be AA35 on the next). Both general aviation and transport pilots were assigned to the SP and PP simulators. A VFR aircraft that was not in communication with ATC was included on every trial, although the level of intrusion was varied.

CDTI configuration

The CDTI symbology in the current study was the same as that used in a previous simulation (ref. 7). For SPs, the pilot's own aircraft was represented by a chevron. The other aircrafts' symbols reflected relative altitude; aircraft within +/- 150 m of ownship's altitude were represented by a full hexagon and those more than 150 m above or below were represented respectively by the upper or lower halves of a hexagon. Alphanumeric data tags with call sign, altitude, ground speed, and vertical direction of flight were available for all aircraft. A three-segment, 90-sec flight path predictor and 90 sec of flight path history were also provided. SP traffic displays were always presented track-up with a fixed map scale covering 52 km. All aircraft within the range of the map were displayed at all times, regardless of altitude. SP control over display alternatives (e.g. map range and traffic density displayed) was eliminated to allow experimental control over the information available.

The PP display presented a fixed-scale, north-up representation of the terminal control area south of San Jose Airport. (It was not ownship-centered, as it was shared by four pilots.) Aircraft were represented by equilateral triangles, oriented north-up regardless of the aircraft's direction of flight with no relative altitude encoding. This traffic display was, thus, considerably simpler and provided less information than that provided for the SPs. Aircraft were presented with 90-sec flight path predictors, flight-path histories, and digital data tags, as on the SP display.

Experimental variables

On half of the trials, all pilots were given CDTIs (allCDTI). On the remaining trials, only PPs were given a traffic display (someCDTI). Realistic radio communications were conducted between the pilots and an approach controller during all approaches. On half of the someCDTI and half of the allCDTI trials, the difficulty of the approach was increased by previously described variations in the scenarios and the addition of a continuous workload estimation task (HI). On the remaining trials, the workload was relatively low (LO). Three flights per group were made under each of the four combinations of CDTI equipage and level of workload. The experimental conditions and scenarios were presented in a different order to each group.

Workload estimates were obtained from SPs during HI workload approaches with a small box that contained ten response buttons patterned after the pilot and observer subjective workload assessment technique (POSWAT) (ref. 12). The left-most button represented a rating of extremely low workload; that on the far right, extremely high workload. A "beep" presented at 1-

min intervals signalled the pilots to make their next rating.

Written Materials

The participants were given a low-altitude chart, an approach plate, and information about the flying characteristics of the aircraft to be controlled. Each pilot was given a written clearance to simulate a handoff from center control and the approach controller was given "flight-data strips" for each IFR aircraft prior to its entry into the terminal area.

A written debriefing form was given to the pilots and controllers after each flight to assist them in recalling the location, status, and intentions of their own and other aircraft. They were asked to assign a number to each aircraft about which they remembered something (ownship was always first) to allow responses about aircraft whose call signs had been forgotten. They were asked to write whatever they could remember about the status of each aircraft (call sign, type, speed, altitude, vertical and horizontal direction of flight, and landing sequence number) and to locate each aircraft (including their own) on a navigation map. The degree of precision required for each parameter to be remembered for an item to be scored as correct (Table 1) was explained during the initial briefing.

Table 1
Definition of "correct" responses for recall task.

Dimension	Range	Units
Position	+/- 2	Nautical Miles
Call sign	Exact	Letters/numbers
Aircraft type	Exact	Manufacturer/type
Ground speed	+/- 10	Knots
Altitude	+/- 200	Feet
Landing Sequence Number	Exact	Assigned position
Direction: vertical	Exact	^ v -
Direction: Heading	Exact	One of 8, 45-deg compass sectors

After completing the traffic information questionnaire, the pilots and controllers were asked to evaluate their experience during the preceeding approach by placing a mark on each of 18 12.5-cm rating scales. The scales were presented in a bipolar adjective format with no numeric labels or interval boundaries:

TASK DIFFICULTY: VERY EASY_____VERY HARD

Fifteen of the scales had been used in previous research (refs. 10,11) to obtain information about the relationships among the subjective experience of workload and a number of related dimensions (e.g. task difficulty, stress, fatigue, effort, etc.). Three additional scales were added to obtain pilots' evaluations of the quality of their communications, aircraft control, and situation awareness.

A written debriefing was given at the end of the experiment to obtain the pilots' impressions about the simulation and the potential value (or liability) of CDTI. This same form was used in several other simulations conducted in this and other facilities (ref. 7)

Subjects

Four groups of pilots served as paid participants in this experiment. Each group consisted of three transport pilots and four instrument-rated general aviation pilots. A recently retired approach controller provided air traffic control for all four pilot groups. All of the subjects had participated in experiments conducted in this facility within one year.

Procedure

The experiment lasted one day per group. Each pilot received three hours of individual training with and without ATC. Because the pilots had flown the simulators before, they required less practice than the pilots had in the earlier simulations. Each set of approaches was relatively short (15 min, on the average) due to their premature termination. The trials began as before; aircraft entered the terminal area one at a time according to the scenario until all eight aircraft were active. Approximately 15 min after the beginning of each run, the display was blanked, the flight terminated, and the traffic situation questionnaire and rating scales completed. After four approaches, the participants were given a 15-min rest. A 1-hr lunch break was given between the practice and experimental trials.

RESULTS AND DISCUSSION

Measures of pilot effort and performance, system efficiency and safety, and communications were analyzed to determine the overall effects of CDTI equipage and workload level. The pilots' responses to the 18 bipolar rating scales were examined to determine whether two different levels of workload had been imposed, as intended. Finally, the information remembered by the pilots and controllers about the traffic situation was analyzed to evaluate the level of pilot's and controller's traffic awareness with and without CDTI. The analyses were performed, and in turn, reported with units of measurement typical of the aircraft environment (e.g. knots, feet, statute, and nautical miles).

The dependent measures were examined initially with one-way analyses of variance for independent groups to determine whether group membership was a significant variable. No significant differences were found for any measure, thus the remaining analyses were performed for individual pilots without regard for group membership.

System Performance

Three indices of system performance were analyzed: (1) intercrossing times, (2) separation violations, and (3) communications. For these and other measures, the effects of the experimental variables, including replications, were examined for the system as a whole, and for the two levels of simulator fidelity (SP and PP) independently.

Intercrossing times

The average interval between successive aircraft crossing specific points has been used as a measure of system efficiency. The more regular and optimally separated the flow of traffic, the better system performance has been considered to be. This measure was less appropriate for the current study, however, because the simulation was terminated before a significant number of aircraft were established on the ILS and had crossed the outer marker. The majority of piloted aircraft (63%) did cross the LICKE intersection, however. Their average separation was 49 sec. The intercrossing times were considerably less than have been obtained in previous simulations, and the variability was high (the overall standard deviation was 29 sec). It is possible that ATC, knowing the simulation would be terminated before the aircraft would have to land, allowed less separation between aircraft as they merged onto the final approach course.

The interval between successive aircraft lengthened from 46 to 52 sec when all simulators were CDTI-equiped, and the standard deviation increased from 26 to 36 sec. There was a smaller increase in intercrossing times and variability on high workload trials, probably because proportionally more aircraft were vectored into holding patterns during high workload trials and never crossed the LICKE intersection.

Separation Violations

Four levels of separation violation severity, reflecting different levels of system failure, were selected in advance of the simulation (ref. 7). The levels ranged from slightly less than standard separation to a collision. The frequencies of violations by experimental condition, simulator type, and severity category were computed, however no formal analyses were conducted because there were so few. No collisions occurred, nor did any near misses. There were eight separation violations in which two aircraft were closer than 3 mi and 1000 ft but maintained at least 1-mi separation. The durations of the separation violations ranged from from 4 to 48 sec. All of them occurred on HI workload approaches, and all but one involved a VFR aircraft not under ATC control. SPs were involved in only one violation, and this occurred when the pilot did not have a traffic display.

Communications

The communications portion of the simulation was conducted professionally by all of the 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 conditions on the ATC system. The communications between pilots and controllers were recorded and subjected to a frequency count and a detailed content analysis.

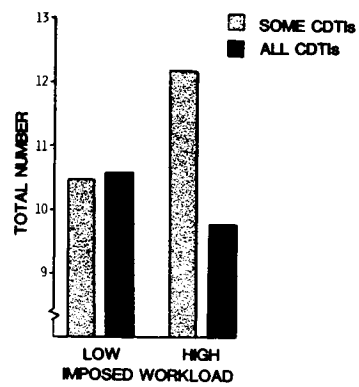


Figure 3: Average number of discrete transmissions to and from each pilot (n=28) per approach (n=12).

On the average, there were 10.7 discrete transmissions to and from each pilot per approach (Figure 3). For LO workload approaches, the number of transmissions was approximately the same with someCDTIs as with allCDTIs (10.49 vs 10.60). For HI workload approaches, however, there were 25% fewer communications when all aircraft were equiped with CDTIs than when only some were (9.77 vs 12.20). This difference was reflected in a significant workload level by CDTI-equipage interaction ($F(1,27) = 5.79$, $p < .05$). There were no significant differences due to replications, workload or CDTI-equipage alone. The reduction in the frequency of discrete transmissions with the addition of a traffic display that was found in this experiment for HI workload trials was similar to that obtained in an earlier simulation (ref. 7), both results suggesting one potential advantage of providing a visual traffic display in the cockpit.

A content analysis was performed on the communications received and initiated by pilots. Since many transmissions conveyed several different types of information, the number of messages was greater than the number of discrete transmissions. The total number of messages transmitted per approach was 117, on the average. This is considerably less than the average of 208 obtained in the previous simulation because the approaches in the current study were ended before the aircraft had been handed off to tower control and begun their final descent. The distribution of message types was similar between the two studies, however; approximately 50% were related to clearances and vectors, 30% were acknowledgements and readbacks, 15% were position reports, and 5% were traffic-related.

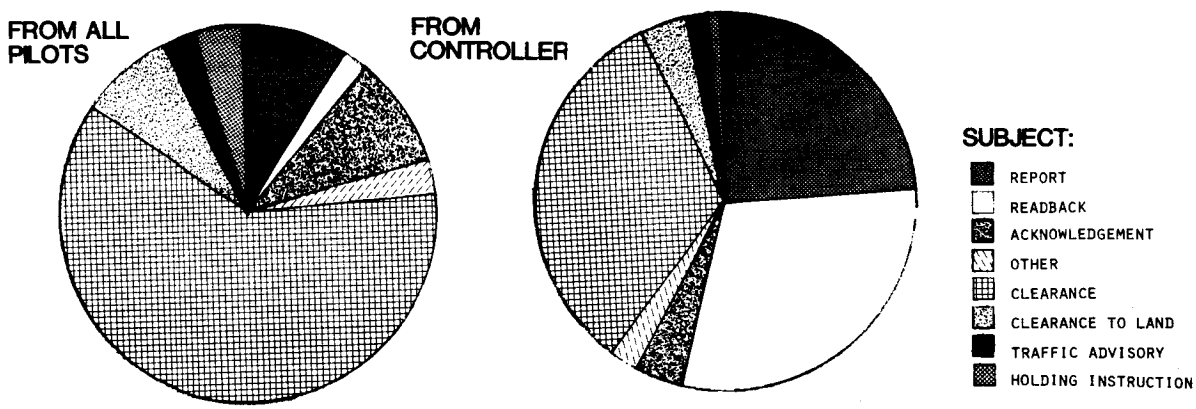


Figure 4: Relative frequencies of different types of communications initiated by ATC or pilots (n = 28; number of approaches = 12)

The messages were divided into eight different categories for analysis: reports, readbacks, acknowledgements, clearances in general, clearances to land, traffic advisories, holding, and other. As you can see in Figure 4, the types of messages transmitted by pilots are considerably different from those transmitted by controllers. Proportionally more of the controllers' transmissions involved readbacks and position reports and proportionally more of the pilots' communications were associated with clearances. Many of the communications were routine exchanges between the pilots and controllers and did not reflect differences in experimental conditions (e.g.

clearances or acknowledgements). For some types of communications, however, meaningful and predictable differences were found as a consequence of CDTI equipage and workload. Three-way analyses of variance for repeated measures were performed individually for each of the 8 types of messages for transmissions to pilots and from them. There were no significant differences found for acknowledgments, clearances, or "other". There was, however, a significant increase in transmissions related to holding instructions with HI workload from pilots ($F(1,27) = 24.17$, $p < .001$) as well as to them ($F(1,27) = 13.59$, $p < .001$). (Figure 5) This reflected ATC's response to the instability in the traffic flow introduced by VFR aircraft and other experimentally imposed problems (e.g. airport closure). There were significantly fewer clearances to land given under the high workload condition than under the low workload condition ($F(1,27) = 9.16$, $p < .01$), reflecting the relatively larger number of aircraft still executing holding patterns when the trial was terminated rather than executing approaches. Pilots made fewer reports in the LO workload, allCDTI and HI workload someCDTI conditions than in the remaining conditions (Figure 5). This was reflected in a significant workload level by CDTI-equipage interaction ($F(1,27) = 11.09$) $p < .01$). The frequency of complete readbacks (as opposed to simple acknowledgements) given by pilots was significantly greater with allCDTIs than with some ($F(1,27) = 12.79$, $p < .01$).

There was a significant interaction between workload level and CDTI-equipage in the number of responses given to traffic advisories by pilots ($F(1,27) = 4.23$, $p < .05$). More traffic-related responses were given under HI than LO workload conditions (because there were more traffic-related problems imposed on HI workload trials). Furthermore, there was an increase in traffic-related messages when all aircraft were equiped with CDTIs, as was found in previous simulations. ATC did use the pilots' abilities to see other aircraft on their CDTIs to involve them in maintaining "visual" separation from other aircraft.

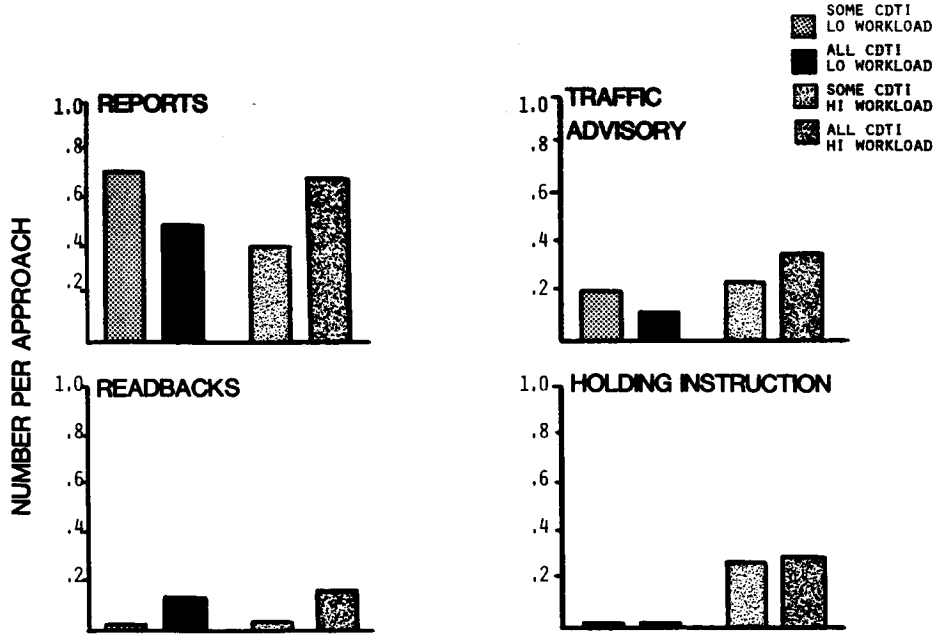


Figure 5: Significant variations in specific types of comunications from pilots as a function of experimental conditions.

No differences in communication types or frequencies were found that could be attributed to the fidelity of the simulator controlled by the pilots. This finding is different than that of an earlier simulation (ref. 9) in which communications differences were found between PPs and SPs. No communications differences were found that could be attributed to aircraft type, possibly because aircraft type was varied from trial to trial for more than half of the simulators while the pilots maintained a consistent communication style.

Pilot Effort

Several measures were used to assess the amount of effort exerted by the pilots under different experimental conditions: the number of SP throttle changes, root mean squared (RMS) aileron and elevator activity, and the discrete number of speed, heading, and altitude changes made by PPs. Each measure was subjected to a three-way analysis of variance for repeated measures. The analyses were conducted independently for PPs and SPs due to differences in the types of controls available to them. Other actions, such as radio frequency change, gear, or flap actuation occurred so infrequently that they contributed little to pilot workload and were not included in the analysis.

Pseudo Pilot Effort

Figure 6 depicts the number of speed, heading, and altitude changes per flight made by the PPs. These inputs occured partly as a function of vectors and clearances given by ATC and partly as a function of the pilot's own initiative. As has been found before, there were no significant differences in the number of heading or altitude inputs as a function of the experimental conditions. This is not surprising, as the PPs were affected by the display condition only indirectly and were not given the additional rating task on the HI workload trials as were the SPs. There was, however a significant interaction in the frequency of speed changes due to workload level and CDTI condition ($F(1,15) = 5.16$, $p<.05$). Under the LO workload conditions, there were approximately the same number of speed changes with both CDTI conditions. Under the HI workload condition, however, PPs made twice as many speed changes when SPs also had CDTIs as when they were the only pilots

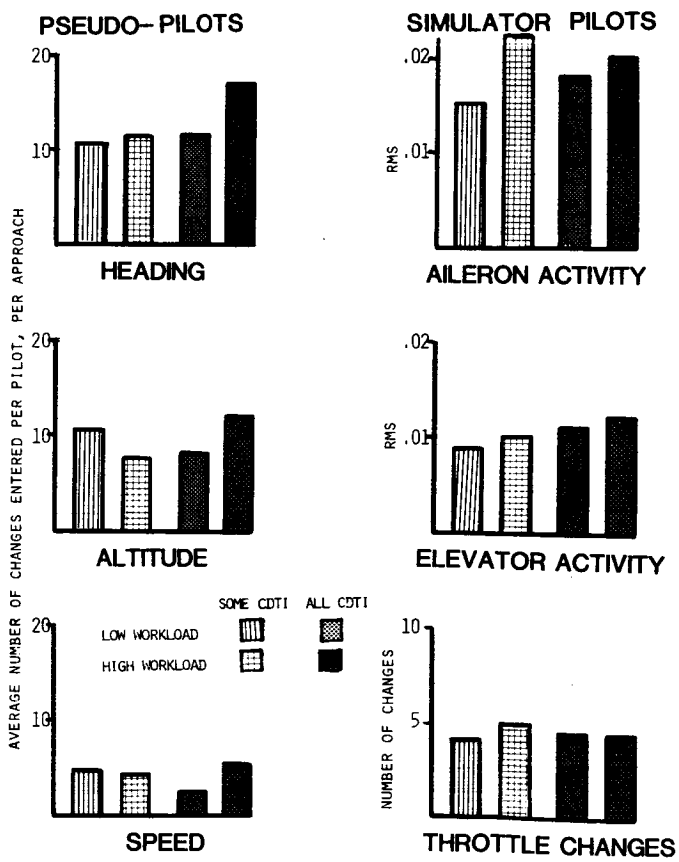


Figure 6: Measures of pilot effort.

Simulator Pilot Effort

There was a significant increase in SP elevator activity under the higher workload conditions ($F(1,11) = 5.48$, $p < .05$), but no significant change in throttle or aileron activity (Figure 6). There were no significant changes in any control activities as a function of CDTI-equipage.

Pilot opinion

Bipolar rating scales

At the conclusion of each approach, all of the pilots completed 18 rating scales. The scales represented dimensions of the task (e.g. difficulty, complexity, time pressure), operator-related variables (e.g. fatigue, stress and motivation), types of effort expended (e.g. physical, cognitive, and perceptual), and aspects of performance (e.g. total, communications, traffic awareness, and aircraft control). These dimensions have been found to be related to different individual's perception of the workload experienced. The pilots' ratings were quantified by assigning numerical values from 1-100 to the 12.5 cm length of the scale. To determine the impact of the experimental conditions on the pilots, individual analyses of variance were performed for each of the scales for SPs alone and for all pilots combined. A summary of the ratings may be seen in Table 2.

The ratings were remarkably similar for SPs and PPs, suggesting that the quality of the simulators was of less significance to the pilots than the essence of the tasks they were performing. Significant increases in ratings were found from the LO to the HI workload approaches for all of the

Table 2
Influence of experimental conditions on pilot ratings

Rating Scales	LO Workload				HI Workload			
	someCDTI		allCDTI		someCDTI		allCDTI	
	SP	PP	SP	PP	SP	PP	SP	PP
OVERALL WORKLOAD	40	40	47	44	48	49	62	58
TASK DIFFICULTY	37	36	40	38	42	42	53	50
TASK COMPLEXITY	33	34	36	38	37	41	48	48
TIME PRESSURE	33	33	33	36	35	37	50	45
PHYSICAL STATE	68	68	68	68	68	67	65	66
ENERGY LEVEL	38	44	40	44	45	47	53	52
EMOTIONAL STRESS	36	36	37	37	38	37	48	48
ATTENTION LEVEL	62	59	65	60	67	65	72	70
ACTIVITY LEVEL	39	39	48	44	46	47	61	55
PHYSICAL EFFORT	23	27	23	29	23	29	28	34
COGNITIVE EFFORT	51	47	55	50	55	52	63	59
SENSORY EFFORT	51	50	52	51	53	54	60	60
MOTIVATION	58	59	63	60	65	61	74	69
TRAINING	46	49	46	48	46	50	48	52
TOTAL PERFORMANCE	51	57	55	58	53	58	55	56
COMMUNICATIONS PERF	62	60	64	65	64	64	63	65
SITUATION AWARENESS	45	46	53	50	48	50	57	56
AIRCRAFT CONTROL	57	56	58	57	60	58	57	59

scales except sensory effort, total performance, aircraft control, and training. Significant increases were also found as a function of CDTI equipage. For all scales, except performance, training, and aircraft control, a significant increase was experienced by all of the pilots when traffic displays were provided. The pilots' ratings indicated that they felt that their own performance improved significantly across replications and that the adequacy of training improved significantly. No change attributable to replications was found for any other scale, however, and no interactions were found among the main effects for any of the scales.

The most important feature of the pilot ratings for the purposes of this experiment was the indication that two distinctly different levels of workload had been imposed, as intended. Furthermore, a significant increase in workload and related dimensions was found with the addition of graphic information about the traffic situation provided by a CDTI. This finding is similar to that reported in earlier studies, adding more weight to the suggestion that the addition of CDTI to a cockpit may have a significant impact on the pilot's workload. It must be remembered, however, that the pilots were required to manually fly the simulators and were operating in an unfamiliar single-pilot mode affording them little attention to spare for an additional display.

Concurrent ratings

During the HI workload trials, SPs were asked to provide discrete estimates of the workload experienced during successive 1-min intervals of flight. This task was added to test the acceptability of this method of acquiring workload ratings from pilots and to provide an additional distraction to the pilots during these approaches. The 10 discrete ratings were assigned numerical values of 10 to 100 in increments of 10 for analysis. As you can see in Table 3, rated workload was higher with CDTI than without, in agreement with the ratings obtained at the end of the approaches. There was no significant change in ratings across replications, nor was there any significant difference between groups. There was, however a significant increase in rated workload from the first to the last 3 min of each flight.

Table 3
Workload ratings obtained from SPs during HI workload approaches

	Time into flight		
	1-3 min	4-6 min	6-9 min
NO CDTI	37	40	50
CDTI	37	45	55

The concurrent ratings were somewhat lower than the post-approach overall workload ratings for the same flights (42 vs 48 with no CDTI and 46 vs 62 with CDTI, on the average), even though the dimension to be judged was defined by an identical phrase and the adjectives anchoring the ends of the scales were the same (e.g. "extremely low" and "extremely high"). This could be due to the way that the ratings were converted into numeric values, the fact that different techniques were used to obtain the ratings, or because the workload was perceived as having been higher in retrospect.

Traffic Situation Recall

As soon as the displays were blanked, the pilots and controllers were instructed to complete the traffic situation questionnaire. The positions of aircraft (including ownship) were indicated by writing the aircraft's identifier (the call sign, if remembered, or a number if not) on the map display at its remembered location. The status of each aircraft was indicated by writing the altitude, ground speed, direction of vertical flight, and heading, call sign, aircraft type, and assigned sequence number for landing . Figure 7 provides an example of the responses obtained from one pilot on a typical approach. The correct values are indicated in parentheses.

	CALL SIGN	A/C TYPE	SPEED (KT)	ALTITUDE (FT)	VERTICAL DIRECTION	LANDING NUMBER	HEADING (COMPASS SECTOR)
1 (OWNSHIP)	Am123	727		4200	^(v)-	2	3
	(A123)	(B-727)	(239)	(3600)		(3)	(7)
2	B456	727			^v-		7
	(B456)	(B-727)	(140)	(1500)		(1)	(7)
3			168		^V-		6
	(78C)	(C-310)	(167)	(5000)		(5)	(1)
4					^(V)-		7
	(99D)	(NAVAJO)	(169)	(3000)		(4)	(2)
5	34Echo			5000	^V-	3	2
	(34E)	(AEROSTAR)	(168)	(5500)		(6)	(8)
6	12Eox	Citabria	105	4000	^V(-)	4	8
	(12F)	(CITABRIA)	(105)	(4000)		(7)	(7)
7					^V-		7
	(56G)	(LEARJET)	(175)	(7000)		(2)	(7)
8							
	(VFR-UNKNOWN)						

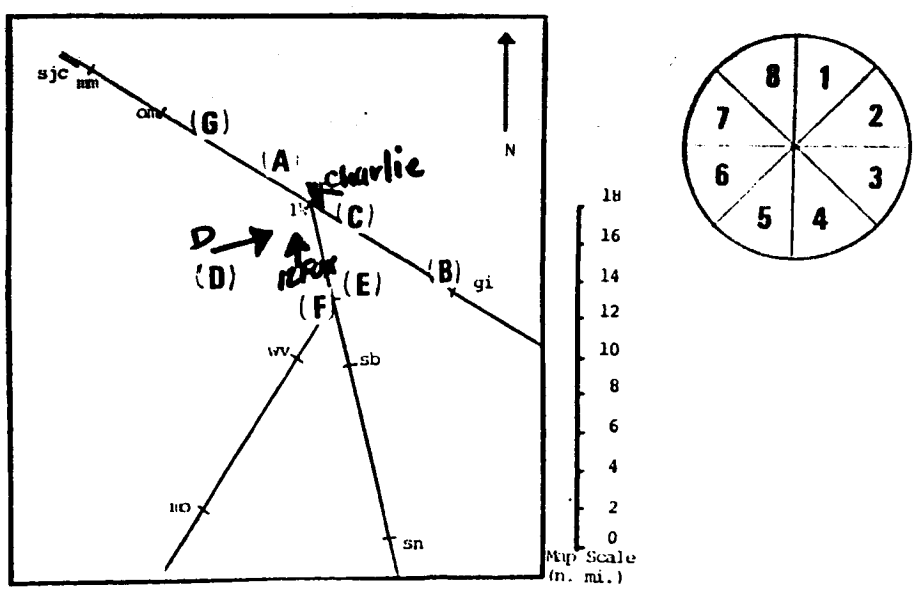


Figure 7: Sample Traffic Situation Questionnaire with responses made by one pilot and correct values for that approach (in parentheses)

As might be expected, more information was remembered correctly by the controller than by any of the pilots (Figure 8). The traffic pattern was his only concern in this simulation, and this task is one with which he was very familiar. The controller's responses were much more complete than were the pilots'. Errors were made, however because the controller responded with assigned speeds, headings, and altitudes, rather than with the current values. He was correct more that 90% of the time about aircraft type, call sign, and sequence number for landing, but less that 50% of the time for speeds, altitudes, and direction of vertical flight.

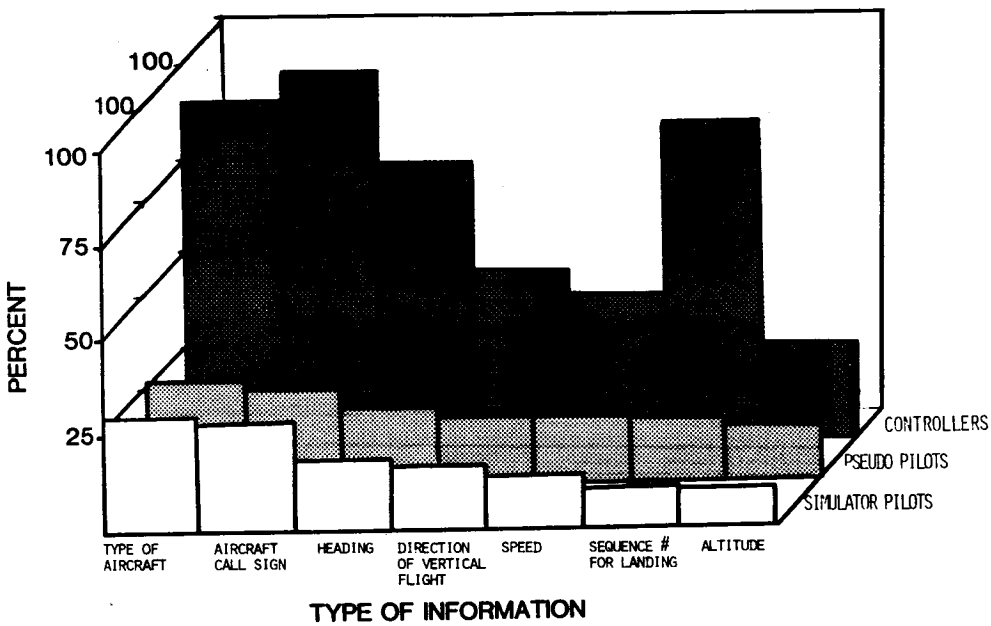


Figure 8: Percent of available information remembered correctly by pilots and controllers averaged across 12 approaches.

PP and SP responses were essentially equivalent, indicating that the quality and format of the CDTI was not particularly important in remembering information. The amount of information remembered correctly about their own and other aircraft was very low (Figure 8). The majority of the information requested on the debriefing was available to all pilots on all approaches, although current information about speeds, headings, and altitudes was not always available from the radio channel. The pilots remembered less than 17% of the total information requested about the status and intentions of their own and other aircraft. More than 25% of the types and call signs of the aircraft were remembered correctly, even though this information was presented at least once on all approaches. Less than 20% of the status information was remembered even though it had been continuously available to all of the pilots on 50% of the approaches and to more than half of them on the rest.

A four-way analysis of variance for repeated measures was performed on the percentages of information correctly remembered by all 28 pilots with replications, CDTI condition, workload, and type of information as factors. There was no significant change in information remembered across replications ($F(2,22) < 1$). There was a significant difference in types of information remembered ($F(6,162) = 17.62$, $p<.001$), but no significant

interaction between type of information remembered and workload or between type of information remembered and CDTI-equipage. There was a significant decrease in information remembered under the HI workload condition ($F(1,27) = 11.36$, $p < .01$); 20% of the items to be remembered were remembered correctly in the LO workload condition, whereas only 17% were remembered correctly in the HI workload condition. Individual comparisons performed for each type of information indicated that the overall differences found as a function of workload level were due to differences in information remembered about speed, vertical direction, and sequence number to land. There was a particularly large difference in remembered sequence numbers, because fewer aircraft flew standard approaches on HI workload trials than on LO workload trials.

The most surprising result was that there was absolutely no difference in the amount of information remembered as a function of CDTI-equipage ($F(1,27) < 1$) nor was there any significant interaction between CDTI-equipage and any other variable. As can be seen in Figure 9, the number of items remembered by PPs and SPs was exactly the same with CDTI as without. The fact that there was no difference for PPs is not surprising, as they were always equipped with a simple traffic display, but the lack of difference found for SPs when their responses were analyzed separately ($F(1,11) < 1$) was not anticipated. The percentage of items correctly remembered was 18.0 with no CDTI and 17.6 with. This finding indicates that the graphic traffic display did not provide any additional information remembered by pilots that was not already obtained from the voice channel.

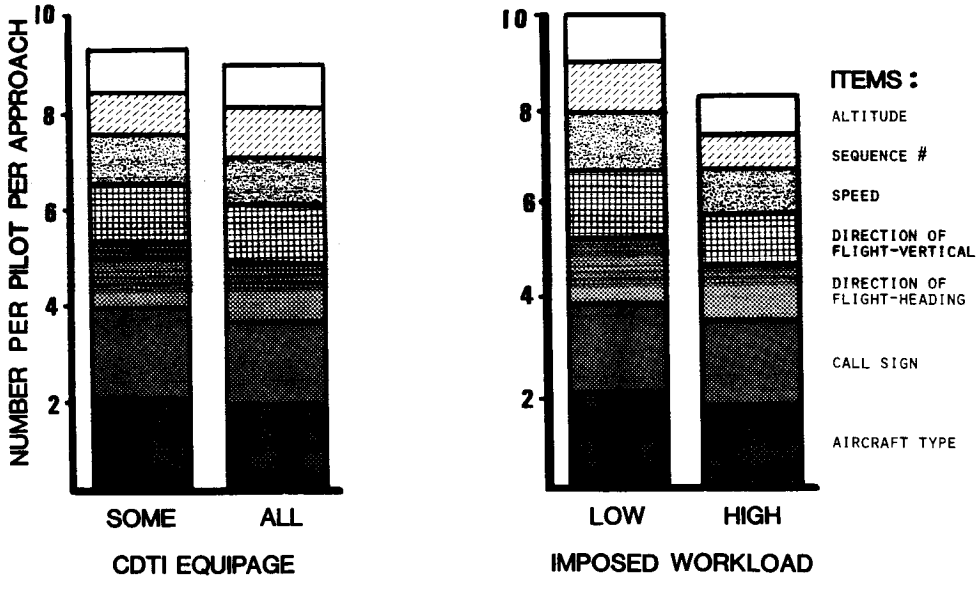


Figure 9: Correctly remembered information about own and other aircraft per pilot (n=28) per approach (n=12).

Because we were concerned that the low percentage of correctly recalled information might have been due to an overly conservative definition of "correct", the original data was reanalyzed to consider not only percent correct, but also the ratio of correct responses to total responses made

and the distributions of errors. The results of this analysis may be seen in Table 4. For call sign aircraft types, ground speeds, and vertical direction of flight, all of the responses made were made correctly. For the remaining items, more than 50% of the information provided by the pilots was in error. For some items, more responses were made with CDTI than without, and for others, the reverse was true. Thus, it did not appgear that pilots were given false confidence (as indicated by more items listed) that the traffic situation was more completely understood with CDTI than without. These data indicate that the pilots were willing to offer information about some parameters if, and only if, they were sure they were correct. For other types of information (e.g. altitude, sequence number or heading) the pilots were either more willing to guess or they were more confident of their traffic awareness than they should have been. These results could have implications for the quality of decisions that pilots might make based on imperfectly understood information derived from a traffic display.

Table 4
Total responses and (percent correct) by experimental condition

	CDTI-equipage				Workload			
	some		all		LO		HI	
Call sign	14	(100)	20	(100)	21	(100)	12	(100)
Type	14	(100)	13	(100)	14	(100)	13	(100)
Speed	9	(100)	7	(100)	8	(100)	8	(100)
Vertical direction	10	(100)	7	(100)	9	(100)	9	(100)
Altitude	13	(48)	13	(53)	13	(49)	13	(52)
Sequence number	14	(57)	10	(48)	13	(43)	6	(62)
Heading	16	(66)	13	(68)	15	(68)	14	(67)

There was no apparent pattern in the distributions of errors for the three categories for which many errors were made. For sequence number, the pilots were typically off by one, but equally often in both directions. For heading, errors were made as often in one direction as the other, although errors of more than 45 deg were found rarely. For altitude, errors as great as 1000 ft were found, however the majority of errors were within 300 to 400 ft of the actual altitude. The errors were evenly distributed between over- and under-estimates (Figure 10), and did not differ as a function of any experimental condition.

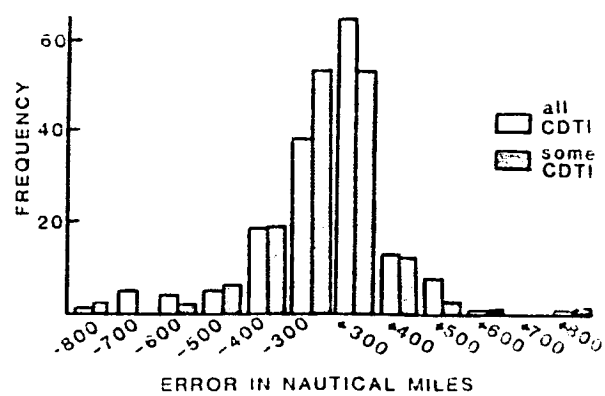


Figure 10: Altitude estimation errors by 28 pilots for 12 approaches.

The pilots were asked to position their own and other aircraft on the map provided. Surprisingly, the pilots were not particularly accurate in their memory of their own positions, even though the position of their own aircraft was continuously throughout the simulation. Virtually all of the pilots attempted to indicate the position of their own aircraft for every trial, but they missed by more than 2 n. mi. on all LO workload trials, and by 4 n. mi., on the average, on HI workload trials. Their error was greater as well on allCDTI trials than on someCDTI trials, even though their own position was identically displayed under both conditions.

The correctly recalled relative positions of other aircraft with respect to the pilot's own position are graphically presented in Figure 11. Approximately 10% of the positions of other aircraft were remembered correctly by the pilots, even when a relatively generous definition of "correct" was used in scoring (within 2 n. mi.). It is interesting that the same percentage of aircraft positions were remembered in LO and HI workload conditions (10%), yet 3% more were remembered with allCDTIs than someCDTIs, a reversal in the pattern of results found for the status information. Although this difference was not statistically significant, it does indicate that there was some improvement in pilots' memories of graphically-presented information about other aircraft (e.g. position), even though there was no improvement in their memory of other aspects of aircraft status. A final interesting feature of this analysis is that the number of aircraft that were positioned correctly is evenly divided between those in front of and those behind ownship. This finding is different than that of the earlier stop-action study (ref. 8), and is not what one would expect from the traffic-display literature. It had been expected that aircraft immediately in front of or closing on the pilot's ownship would be the only ones remembered correctly and few, if any, other aircraft would be noticed.

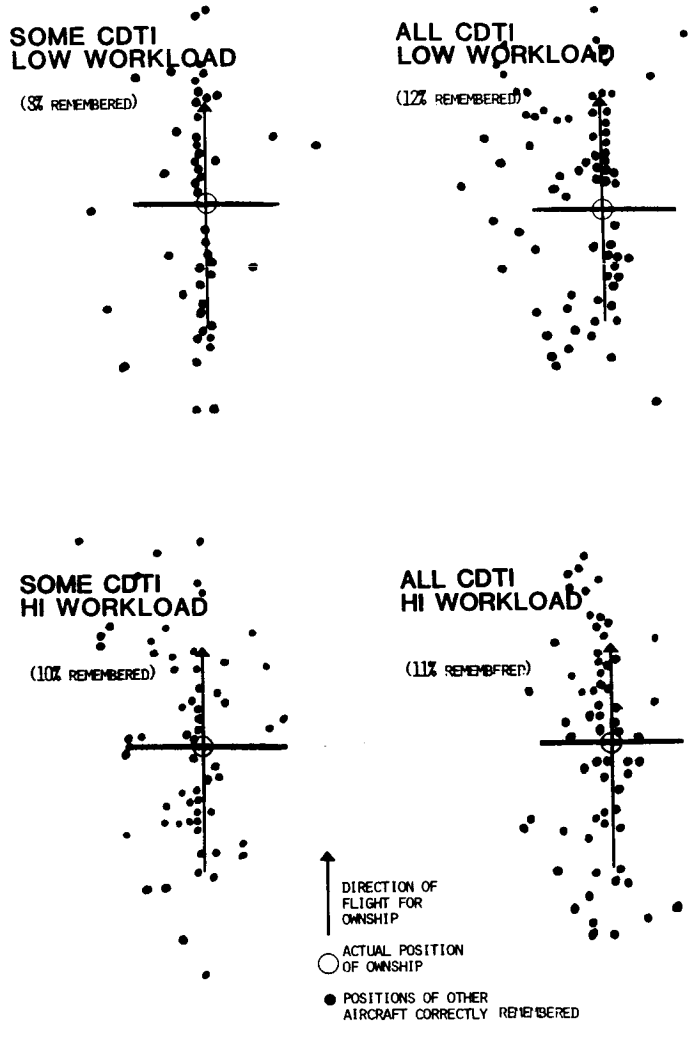


Figure 11: Correctly remembered positions of other aircraft relative to the position and direction of flight of own aircraft.

Final Debriefing

At the conclusion of the simulation, all pilots were asked to complete a written debriefing about the quality of the simulation, their experiences during the experiment, and the potential benefits and liabilities of CDTI. The SPs responded that the simulation was adequately realistic, whereas the PPs felt that it was not, reflecting the difference in their simulation facilities. There was considerable disagreement among the pilots whether the workload imposed on them during the simulation was greater or less than that which they would normally experience in flight. However, the majority agreed that the tasks that they were required to do with the CDTIs would be equivalently difficult to perform in the air as in the simulation. More than 70% of PPs and SPs predicted that there would be increase in their workload over present levels if a CDTI was provided in their cockpit, but that the increase would be acceptable. All of the pilots indicated that the addition of a CDTI would substantially enhance their ability to plan ahead and remain aware of the distance between their own and other aircraft. The pilots agreed that they would be willing occasionally to accept responsibility for spacing between their own and another aircraft with a CDTI, indicating a reluctance to assume more responsibility in the cockpit even with the addition of a traffic display.

The potential benefits cited by both general aviation and transport pilots included improved situation awareness, "electronic" visual separation, and better information about closing rates. The disadvantages included concern over cockpit clutter, increased pilot workload, increased time with "head in the cockpit" in the terminal area, and the length of time required to interpret data. The primary disadvantage cited by the general aviation pilots was the difficulty of monitoring a visually presented traffic display while hand-flying an aircraft in single pilot operations. The general aviation pilots also expressed concern about the cost and weight of such a system. The responses obtained in the debriefing are similar to those of earlier simulations performed in the same facility (ref. 7).

CONCLUSIONS

This simulation was conducted with the express purpose of determining what information pilots remember about the status, positions, and intentions of other aircraft when provided with a graphic traffic situation display. Because the simulation was conducted under instrument flight rules, and pilots did not take unilateral action, it was not possible to determine what advantages a traffic display might provide to pilots required to use such displays for local traffic management or pilot-determined sequencing. The only deviation from standard IFR procedures allowed was that ATC could give clearances to land behind or to follow aircraft in sight on the traffic display. The measures of performance that were taken during the approaches indicated that the addition of traffic displays had relatively little impact on aircraft control, separation violations, or traffic flow regularity. Some differences were found in the number and types of communications as a function of CDTI-equipage, as have been found in earlier studies, with increases in some and reductions of others as a function of presence or absence of a traffic display. Pilot opinion measures of workload indicated, as has been found in several previous studies, that the addition of a CDTI increased workload, even though its benefits warranted this increase from the point of view of the pilots. The most important

finding was that relatively little information was remembered by any of the pilots regardless of CDTI-equipage and with "party-line" communications throughout. Furthermore, as overall pilot workload was increased, the information remembered was reduced even further, and was not improved by the addition of a traffic display. Approximately 10% of the positions and 20% of the status information were remembered correctly by the pilots even though they were aware of the purpose of the study and had to recall information about the traffic situation repeatedly. Our fear that the pilots would remember an artificially high percentage of information due to the design of the simulation appeared to be unfounded, unless pilots remember even less in flight than the limited amount that they did in this experiment.

The results of this study suggest that a very careful analysis of the information transfer from a traffic display to the pilot should be made under realistic conditions in which the pilot is performing tasks that require the use of the display. In addition, adequate practice is needed to insure familiarity with the meaning and format of the display and procedure. Remembered information should be but one element in the analysis. Information usage and strategies selected during the simulation should be studied as well, to obtain a complete analysis of the information extracted from a traffic display as a function of the tasks a pilot is performing. The results of this analysis and that of the current simulation could then be used to specify the optimal display content, method of presentation (graphic or alphanumeric), and medium (visual or audio) for traffic information under a variety of ATC environments. The role that a pilot will play with a CDTI remains the key issue in such an analysis; if the CDTI is to provide no more than pilot assurance, then this goal has been satisfied in this and other simulations. If the goal is to increase the information obtained, used, and remembered by the pilots, then it is possible that a graphic traffic display, at least as currently formulated, may not provide enough additional information or alter performance sufficiently to warrant the cost entailed by its addition into the national airspace system.

REFERENCES

1. Chappell, S. L. & Palmer, E. A. In-trail following during profile descents with a cockpit display of traffic information. (Proceedings of the First Symposium on Aviation Psychology.) Columbus, Ohio, 1982.

2. Kreifeldt, J. G. Simulation Fidelity and Numerosity Effects in CDTI Experimentation NASA CR-166459, (March 1983).

3. Hart, S. G. & Kreifeldt, J. G. Multiple Curved Descending Approaches and the Air Traffic Control Problem NASA TM-78,430 (August 1977).

4. Chappell, S. L. & Kreifeldt, J. G. Air traffic control of simulated aircraft with ond without cockpit traffic disaplays. (Proceedings of the 18th Annual Conference on Manual Control). Wright Patterson Air Force Base, OH, 1982 (in press)

5. Hart, S. G. & Loomis L. L. Evaluation of the potential format and content of a cockpit display of traffic information. Human Factors, 1980, (1980), 591-604.

6. Ellis, S. R. & Stark, L. Pilot scanning patterns while viewing cockpit displays of traffic information. (Proceedings of the 17th Annual Conference on Manual Control) Los Angeles: UCLA, June 1981.

7. Hart, S. G. The effect of VFR aircraft on approach traffic with and without cockpit displays of traffic information. (Proceedings of the 18th Annual Conference on Manual Control). Wright Patterson Air Force Base, OH, 1982. (Proceedings in press)

8. Melanson, D., Curry, R. E., Howell, J. D., & Connelly, M. E. The effect of communications and traffic situation displays on pilots awareness of traffic in the terminal area. (Proceedings of the 1973 Annual Conference on Manual Control). Cambridge: Massachusetts Institute of Technology, 1973, 25-39.

9. Chappell, S. L. & Kreifeldt, J. G. Facility requirements for cockpit traffic display research. (Proceedings of the 1982 Meeting of the Human Factors Society). Seattle, 1982.

10. Childress, M. E., Hart, S. G. & Bortolussi, M. R. The reliability and validity of flight task workload ratings. (Proceedings of the 26th Annual Meeting of the Human Factors Society). Seattle, 1982

11. Hauser, J. R., Childress, M. E., and Hart, S. G. Rating consistency and component salience in subjective workload estimation. Paper presented at the 18th Annual Conference on Manual Control. Dayton, OH, May 1982. (Proceedings in press)

12. Stein, E. S., Fabry, J. & Rosenberg, B. The elusive goal of measuring pilot workload in general aviation. (Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics) Edwards AFB: AFFTC-TR-82-5, 1982, 275-280.

"Investigation of the Steady State Visually Evoked Brain Response"

Andrew M. Junker
Air Force Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Dayton, Ohio 45433

Karen J. Peio
Systems Research Laboratories, Inc.
Dayton, Ohio 45440

ABTRACT

When light stimulus is presented to a subject under the right intensity and modulation levels, a corresponding response can be detected in the subject's EEG. If the stimulus is sine wave modulated light with a frequency above 1 hz the result is called a steady state response. By sine wave modulating the light intensity with a sum of sine waves, evoked responses can be measured at a number of frequencies simultaneously. From this data, visual/cortical describing function measures can be obtained. These measures reflect dynamic properties of the system within the brain being stimulated. This paper will present results obtained from using sums of 11 and 14 sine waves (from 5 to 50 hz) as the visual evoking input. Responses were obtained from subjects while they attended to the evoking stimulus only and while they performed a supervisory control task concurrently with the evoking stimulus. The effects of the control task on the evoked response are given. The results will be presented in terms of input/output describing functions and background EEG power.

1983 Annual Manual Informal Presentation

STI WP-1167-8

A DIGITALLY MECHANIZED CROSS-COUPLED INSTABILITY TASK FOR CONTROL WORKLOAD RESEARCH

Henry R. Jex
Principal Research Engineer

Richard A. Peters
Principal Specialist

Systems Technology, Inc.
Hawthorne, CA 90250

The Cross-Coupled Instability Task (CCIT) is a system for measuring the workload margin of a set of primary control tasks such as flying or driving. It uses the STI developed principle of controlling a slightly unstable, auxiliary task whose degree of instability, λ_x, is adaptively adjusted to use up whatever control-workload margin remains after the primary tasks are attended to sufficiently well to maintain the average performance "close to the initial unloaded" level. When normalized by the critical-instability, λ_c, (obtained while performing the auxiliary control task alone), a non-dimensional measure of "excess-control" capacity is obtained ($XC \equiv \lambda_x/\lambda_c$), which has been found to correlate well with subjective measures of task workload and handling qualities ratings.

Portable digital mechanization of the CCIT has been made for use by the Crew Performance Branch of the U.S. Naval Air Test Center at Patuxent River, MD, in the Display Evaluation by Flight Test (DEFT) program. The 0.3 ft^3 CCIT computer package fits into a slot in the Calspan NT-33A Variable Stability Aircraft which is being used to evaluate various head-up display concepts. The CCIT interfaces with various aircraft motion sensors, controls, and DEFT display channels. The check-pilot controls the CCIT operation via a small keyboard terminal in the rear cockpit, while the front seat pilot flies the adaptively loaded tasks via the NT-33A's fly-by-wire control system. A simple model-follower-loop option is available to create aircraft motions in consonance with the CCIT supplied head-up display motions. All parameters of the several functional modules (which monitor performance, produce the dynamics, and perform sequencing and scoring of the tasks) can be accessed and optimized by the experimenter to suit individual situations with the help of a portable Ground Support Unit. The sequencing of various CCIT, Critical Instability Tests (CIT), and Self-Coupled Tasks (SCT) is automatically controlled by the device, which computes and logs all of the settings and scores for later playback to a hard-copy recorder. The CCIT computer can also be used in the simulation lab to prepare or validate flight tests, or in a stand-alone mode.

The CCIT has been developed, delivered and successfully checked out in both lab and flight and is ready for formal experiments in the DEFT Program.

This paper briefly describes the CCIT features, functions and their mechanization, gives examples of its use, and shows typical results.

Acknowledgements

This work was sponsored by the U.S. Navy Flight Test Center's Crew Performance Branch, at Patuxent, MD. The technical Monitor was Dr. Sam Schiflett, who has encouraged the application of various auxiliary task workload measuring approaches for application to the Navy's Display Evaluation by Flight Test (DEFT) program, so that various Head-Up Display (HUD) formats could be compared by more sensitive methods than either overall performance or subjective evaluations alone. He has acted while others studied.

The authors are grateful for the occasionally heroic contributions of others to this project:

Thad Dubzik of Instrumentation Technology Systems, Inc. for computer hardware design.
Henry Hernandez of Executive Systems Inc. for help in the digital hardware specialization and debugging.
Lou Knotts of Arvin/Calspan, for his efficient flight testing of the CCIT on a very limited budget.

Background

In 1972 the Cross-Coupled Instability Task grew out of experience in using the STI Critical-Instability Task Tester (CTT) in central-display research involving display scanning and divided-attention allocation.[1-4] It has been used successfully in various analog-computer mechanizations on a number of research projects, such as display format effects,[3] pilot-handling qualities evaluations,[5] integrated-display format simulations[6] and control manipulators[7] digital displays.[8]

The present all digital version grew out of a need to provide a sensitive test for the often subtle head-up display variations planned for the Navy's Display Evaluation by Flight Test project, in which different HUD formats (e.g., for terrain following or landing) would be compared in the CALSPAN NT-33A Variability Stability research aircraft. Only about 0.3 ft^3 of space was available, and a number of interfaces with the existing NT-33A and HUD analog signals, subject-pilot controls and check-pilot CCIT operation, mode-switching, and controls had to be accomodated. These complex sequencing and numerous test options dictated a digital mechanization for the DEFT CCIT, with MIL-grade flight hardware.[9]

In addition, it was felt desirable to have the NT-33A aircraft physical motions follow the displayed motions to reduce motion-cue-conflicts. Almost as an afterthought, the requirement to mechanize an aircraft-like vehicle model and a tight model follower loop was added to the design specifications. It turned out that this was the first model-follower loop ever employed on the WT-33A, and its careful, versatile and flight-safe design turned out to be the-tail-that-wagged-the-dog.

Despite these severe functional and environmental requirements, all were met and demonstrated in a pair of lab and flight "shakedown" sessions that took only a few days each.

The following compendium of viewgraphs illustrates the above material.

REFERENCES

1. Jex, H. R., J. D. McDonnell, and A. V. Phatak, "A 'Critical' Tracking Task for Manual Control Research," IEEE Trans. on Human Factors in Electronics, Vol. HFE-7, No. 4, Dec. 1966, pp. 138-145.

2. Allen, R. W., W. F. Clement, and H. R. Jex, Research on Display Scanning, Sampling, and Reconstruction Using Separate Main and Secondary Tracking Tasks, NASA CR-1569, July 1970.

3. Jex, H. R., W. F. Jewell, and R. W. Allen, "Development of the Dual-Axis and Cross-Coupled Critical Tasks," Proc. of the Eighth Annual Conference on Manual Control, University of Michigan, Ann Arbor, Michigan, AFFDL-TR-72-92, May 1972, pp. 527-552.

4. Jex, H. R. and W. F. Clement, "Defining and Measuring Perceptual-Motor Workload in Manual Control Tasks," Mental Workload, Its Theory and Measurement, Neville Moray (Ed.), Plenum Press, New York and London, 1979, pp. 125-177.

5. McDonnell, J. C., Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities, AFFDL-TR-68-76, Dec. 1968.

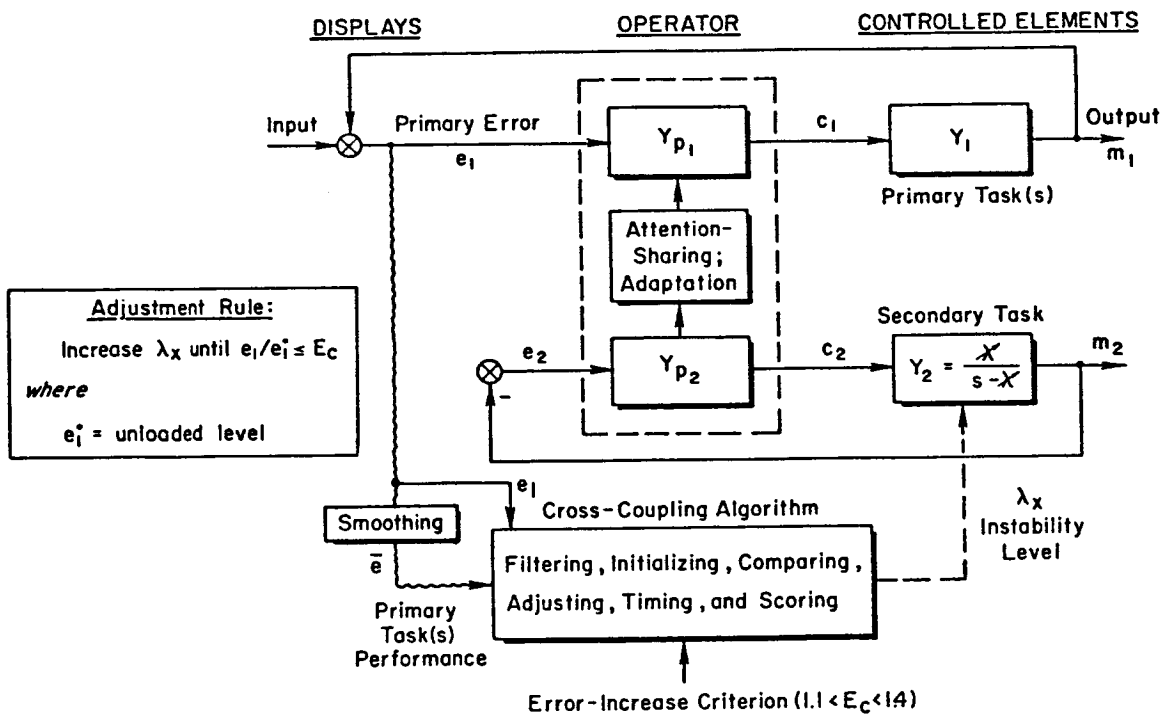
6. Clement, Warren F., Investigation of the Use of an Electronic Multi-Function Display and an Electromechanical Horizontal Situation Indicator for Guidance and Control of Powered-Lift Short-Haul Aircraft, NASA CR-137922, Aug. 1976.

7. Merhav, S. J. and Orna Ben Ya'Acov, "Control Augmentation and Work Load Reduction by Kinesthetic Information from the Manipulator," IEEE Transactions on Systems, Man, & Cybernetics, Vol SMC-6, No. 12, Dec. 1976, pp. 825-835.

8. Hess, Ronald A. and Walter M. Teichgraber, "Error Quantization Effects in Compensatory Tracking Tasks," IEEE Trans., Vol. SMC-4, No. 5, July 1974, pp. 343-349.

9. Jex, Henry R. and Richard A. Peters, Excess Control Capacity System Description and Specification for the DEFT Application in the Calspan NT-33A, Systems Technology, Inc., TR-1167-1, Mar. 1981.

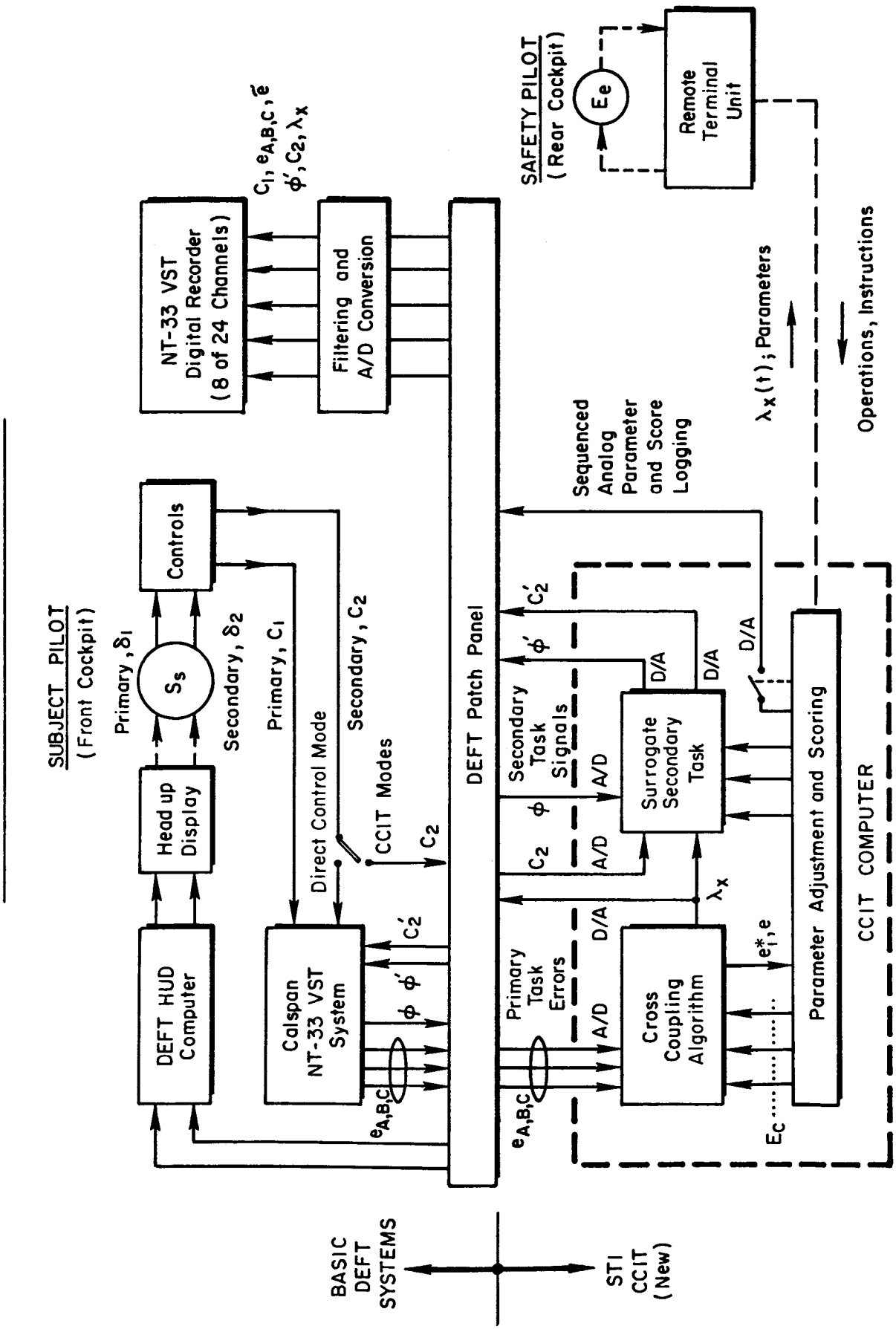
PRINCIPLE OF CROSS-COUPLED-INSTABILITY TASK



DEFT/CCIT REQUIREMENTS

- HUD Research; Display Evaluation by Flight Test
- A/C: Calspan NT-33A; air, ground, lab
- Primary Tasks (typical): γ_c, $\theta \rightarrow \delta_e$; $u \rightarrow \delta_T$ } Typical
- Auxillary Task: $\phi \rightarrow \delta_a$ } Typical
- Motion Follower Loop: $|\phi_{motion}/\phi_{model}| \approx 1.0$ up to 2 Hz
- Subject Pilot (front)
  - Direct Control
  - Fly-by-Wire
  - Model Follow
  - Abort
- Check Pilot (rear)
  - Control Tasks, Modes
  - Monitor Task; Scores
  - In-Flight Adjustments
- Smooth Mode Switching; A/C Safety
- Versatile Options Many Adjustments
- Space: 0.3 ft^3, Severe Environment, 28V
- Automatic Run & Score Logging For Replay

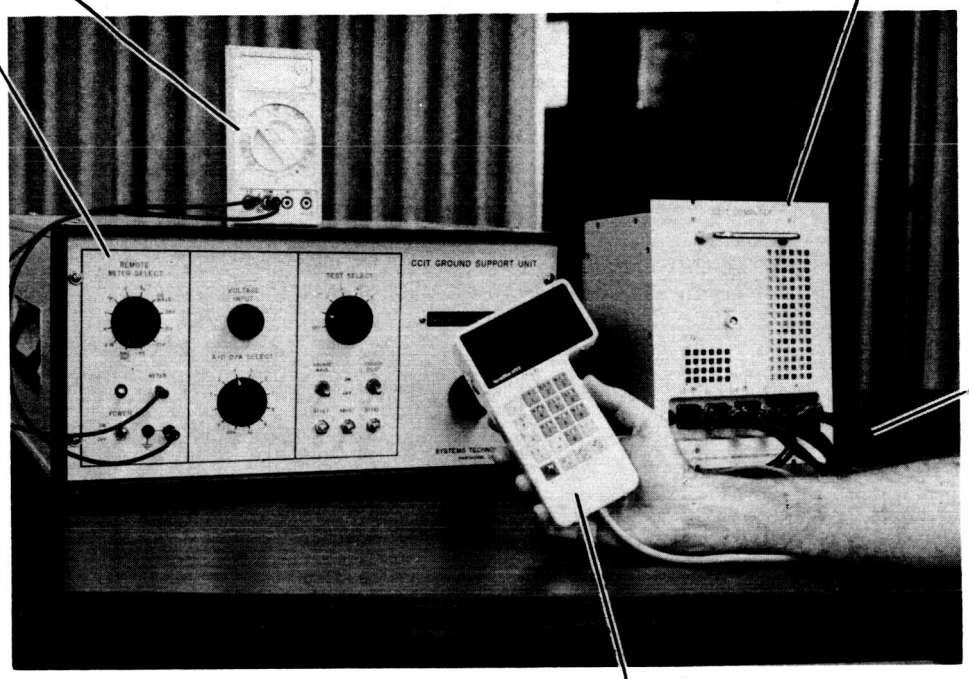
CCIT/NT-33A INTERFACES AND SIGNALS



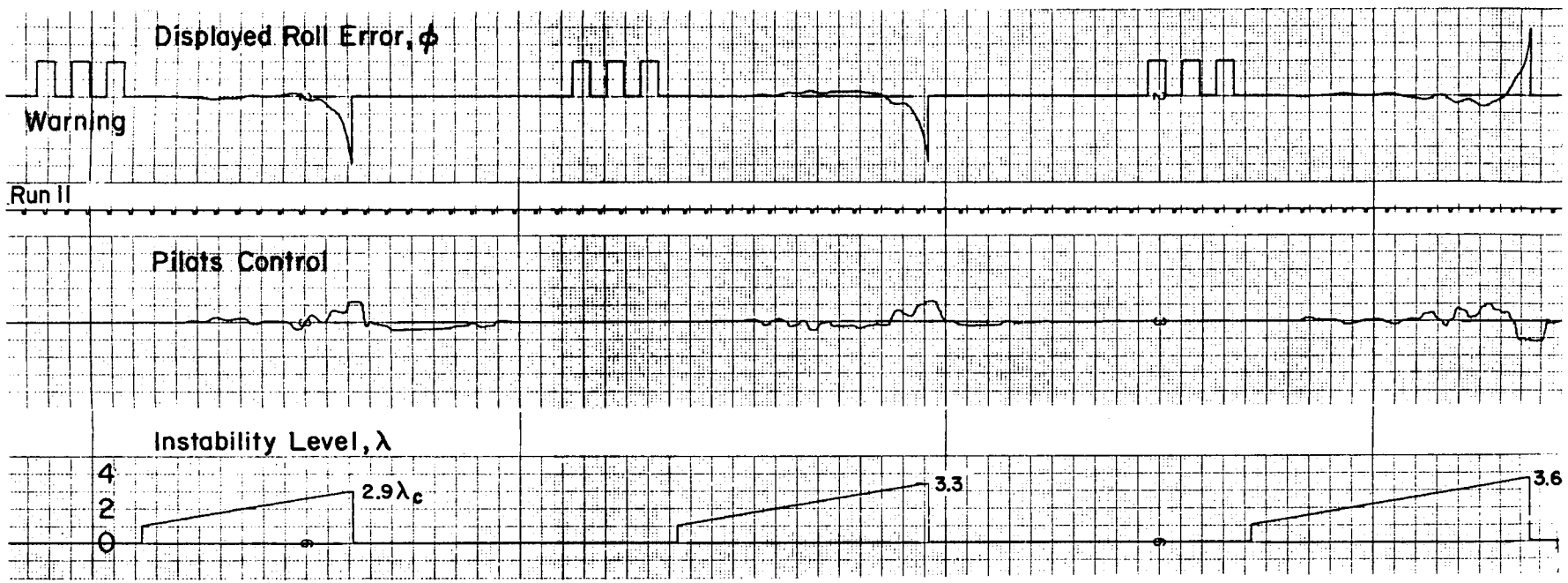
CCIT TESTER OPTIONS AND SCORES

MODE	TEST	SCORES
GROUND TEST (GT)	--	
DIRECT CONTROL (DC)	--	
FLY-BY-WIRE (FBW)	--	
MODEL FOLLOWER (MF)	--	$e_{rms} \rightarrow e_{par}$
GT, DC	CRITICAL INSTABILITY (CTT)	$\hat{\lambda}_c$, $\bar{\lambda}_c \pm \sigma$ (1-9 trials) $(\tilde{\lambda}_c \rightarrow \lambda_{par})$
GT, DC, MF	SELF-COUPLED INSTABILITY (SCT)	e_1^*, λ_s @ T_{run} $(\tilde{\lambda}_s \rightarrow \tilde{\lambda}_{par})$
GT, MF	CROSS-COUPLED INSTABILITY (CCIT)	e_1^* all primary tasks, unloaded $\lambda_x(t)$, λ_F @ T_{run} $E \equiv e_1^*/e_{par}$ Normalized Error $XC = \lambda_F/\lambda_{par}$ (Excess Capacity) $P = K_E E + XC^{-1}$ Performance Index

MAIN CCIT COMPONENTS



TYPICAL CRITICAL TASK RUN (3 TRIALS)

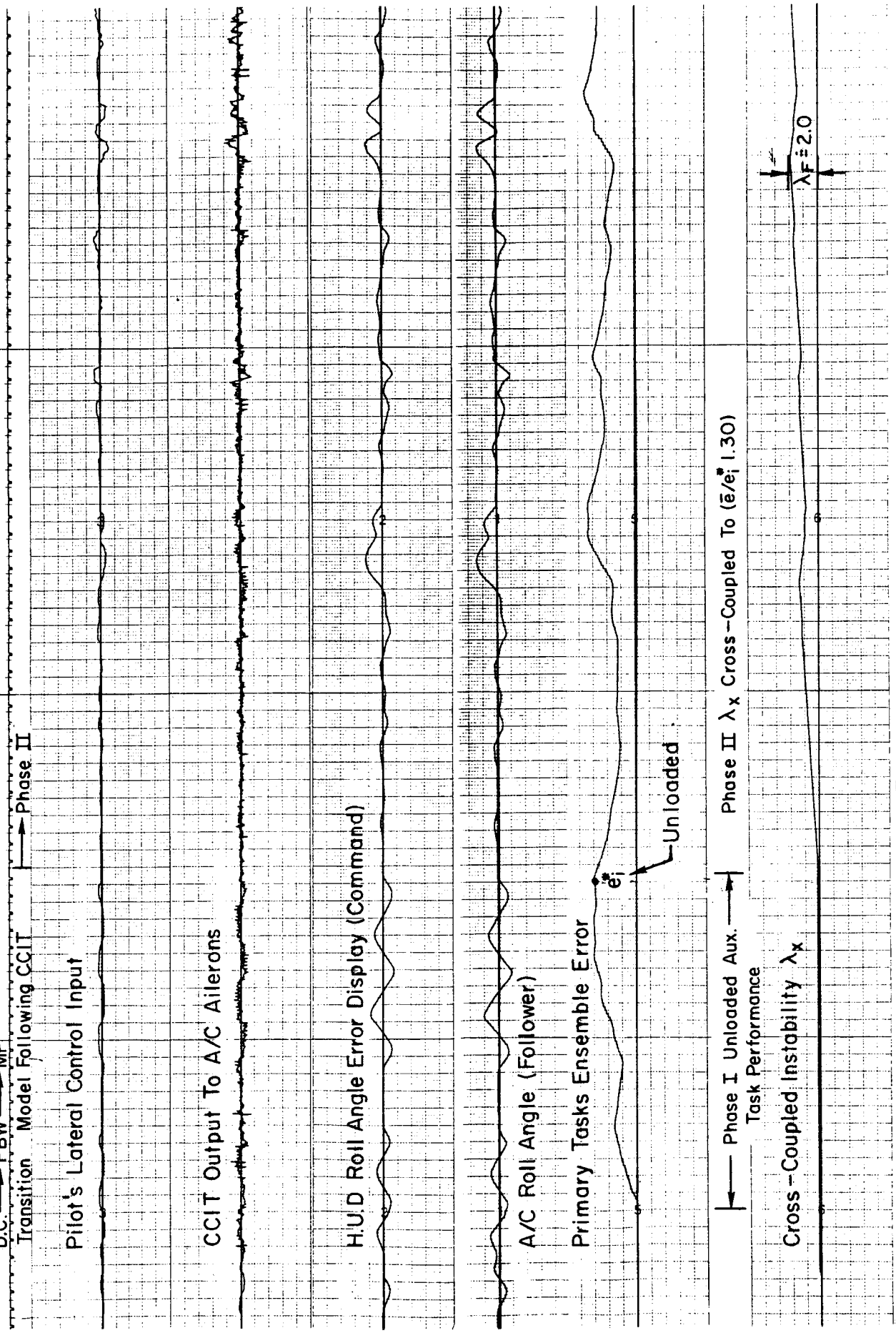


AUTOPACED TESTS

RUN NO	LAMBDAC BY TRIAL NO 1	2	3	4	5	6	7	8	9	LBDA MED	LBDA AVE	STD DEV
11	2.94	3.34	3.62							$\hat{3.34}$	$\overline{3.30}$	$\pm$ 0.34
12	2.91	3.38	3.18							3.18	3.16	0.24

TYPICAL CCIT RUN TRACES

ON NATC CREW PERFORMANCE BRANCH, F-18-LIKE SIMULATOR: TERRAIN FOLLOWING HUD



Run 10, $\bar{e}_i^* = 6.0$ deg ϕ, $\bar{\lambda}_x = 2.0$, XC = 2.0/3.3 = 6.1

SHAKEDOWN TESTS

LAB

- BUILT-IN SOFTWARE TESTS
- HRJ λ SCORES CK MK 8A CTT

SIMULATOR

- AUG. 1982 NATC - CREW PERF. BR. (PATUXENT RIVER, MD)
- F-18-LIKE, FIXED BASE HUD SIMULATOR (TERRAIN FOLLOWING TASK)
- INSTALLED AND CHECKED OUT (SAME DAY)
- TYPICAL CTT AND CCIT RUNS
  (SIMULATOR DELAYS, $\tilde{\lambda}_c \approx 3.2$)
  STABLE CROSS-COUPLING DEMONSTRATED
  LONG TERRAIN FOLLOWING HUD ($XC = \lambda_x / \lambda_c = .61$)

FLIGHT

- NT-33A PROJECT PILOT: LOU KNOTTS
- INSTALLED AND FUNCTIONALLY PROVEN IN 2 DAYS
- INADVERTENT MODEL FOLLOWER LAGS INDUCED LATERAL PIO (NO HUD)

STATUS

- READY FOR PARAMETER OPTIMIZATION R&D AND APPLICATIONS

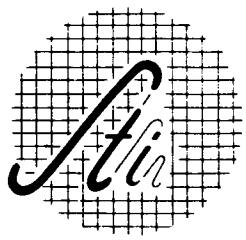
WORKING PAPER 1167-8

A DIGITALLY MECHANIZED CROSS-COUPLED-INSTABILITY TASK (CCIT) FOR CONTROL WORKLOAD RESEARCH

HENRY R. JEX

RICHARD A. PETERS

SYSTEMS TECHNOLOGY, INC.
HAWTHORNE, CA 90250



MENTAL WORKLOAD IN SUPERVISORY CONTROL OF AUTOMATED AIRCRAFT

by

Keiji Tanaka, Ahmet Buharali, and Thomas B. Sheridan
Department of Mechanical Engineering,
Massachusetts Institute of Technology

SUMMARY

The purpose of this study is to investigate the nature of pilot mental workload in highly automated aircraft. On the basis of Rasmussen's model where human behavior is divided as *skill*, *rule* and *knowledge-based*, we hypothesize that mental workload is multi-dimensional, and that different aspects of workload are associated with each level of human behavior. In order to examine these hypotheses, a laboratory flight simulator was developed, functions of which included dynamics of a general aviation aircraft, autopilots, and navigational aids, as well as artificial air traffic controllers. Terminal-area approachs were simulated based on several scenarios where pilot tasks included aircraft guidance, navigation, aircraft configuration changes, and communication with air traffic control. In each case workload was measured by employing subjective rating scales and the number of pilot actions. It was shown that the level of automation available affects only the workload of *skill-based* behavior, whereas the abnormality of the situation resulted in an increase in workload for *rule-* and *knowledge-based* behavior.

INTRODUCTION

The move to automate the cockpit environment is highlighted by the successive introduction of improved devices such as autopilots and CRT displays, made possible by the recent development of computer technology. The emphasis of cockpit automation so far has been on alleviating human control workload. This trend inevitably shifts most of the pilot's tasks from conventional manual control to supervisory and monitoring control. Yet, it is

possible to say that in this cockpit environment the effects of cockpit automation on pilot mental workload are still unclear, and that assessing human supervisory and monitoring behavior has just begun. In order to establish the future automated cockpit, it becomes crucial to elucidate the mental workload associated with the supervisory and monitoring tasks which pilots should accomplish in the present cockpit. Several successful advances in analyzing mental workload have recently been reported[1-7]. It seems clear that in automated aircraft, workload is still one of the most important factors for evaluating the pilot-aircraft system, even though the nature of workload has greatly changed from that of manually controlled flight, which now also forms a basis for handling quality criteria.

It seems to be an emerging consensus that there are multiple factors in mental workload. Based on Rasmussen's model of the behavior of process operators[8,9], we hypothsize that there are also three behavioral levels in pilot-aircraft systems: *'skill'-*, *'rule'-*, and *'knowledge-based'*. The *skill-based* behavior corresponds to conventional manual control tasks such as manipulating controls to maintain aircraft attitude. This forms the inner-loop feedback in the pilot-aircraft system. In most modern transports, however, these actions have been replaced by the stability and control augmentation system and the hold-type autopilot, such as pitch hold, speed hold, or altitude hold. Rule-based behavior corresponds to procedural activities such as configuration changes or ATC communications, which are specified by operational flight manuals or regulations. *Rule-based* behavior has been partly replaced by command-type autopilot functions such as VOR coupled navigation or automatic landing. *Knowledge-based* behavior corresponds to judgements or decisions about which rules apply in case of an emergency situation. This loop forms the most outer feedback from the standpoint of aircraft operations, and is considered to be the slowest information exchange between the pilot and the aircraft. That is, in this loop the results of the pilot's actions as revealed by aircraft responses cannot be recognized as quickly as, for example, the pitch response to a control column deflection. This concept of multi-level human behavior allows us to assume that different kinds of information processes should cause different kinds of mental workload.

Based on this hypothetical hierarchy of the pilot's behavior, in the present study we simulated the operation of a current automated aircraft in order to reproduce supervisory control situations. By introducing changes in ATC commands, we measured three kinds of subjective workload which correspond to the respective levels of human information processes. It is our present aim to examine whether there is a situation where a pilot experiences a

relatively high workload in an automated cockpit. Based on subjective rating scores, we analyzed the influence of cockpit automation and emergency situations on three aspects of mental workload.

EXPERIMENT

A laboratory flight simulator was developed for this experiment. Fig.1 shows a block diagram of the simulator, which consists of a mini-computer, a monochrome graphic display, and a control box. The functions of the simulator included real-time computation of the following items:

1. Flight dynamics of a general aviation aircraft, the four-engine Jetstar[10], with fixed aerodynamic coefficients
2. Dynamics of an automatic flight control system
3. Calculation of navigational information
4. Calculation of visual perspective information
5. Artificial air traffic control commands
6. On-line data recording

The data-update rate was set at 0.1 [sec]. The display format is shown in Fig.2. The functions of the automatic flight control system were emphasized in the software of this simulator. The autopilot functions are summarized in Table 1. Switching autopilot modes as well as radio channels and gear and flap movements are made possible through switches attached to the control box shown in Fig.3. Availability of this autopilot was one of the experimental variables.

Another specific function of this simulator was to simulate air traffic control. Pilot requests were submitted from a keyboard by using abbreviations such as RA for 'request clearance to waypoint A', or RL for 'request landing clearance'. The requests to the ATC controller and corresponding ATC commands were given on the Megatek display. Two control towers were prepared; the pilot had to communicate with the proper tower at the proper time.

The simulation flight proceeded based on one of the prepared scenarios and occasional pilot requests. Table 2 is a summary of the procedures for a simulated approach and landing in a terminal area [Fig. 4]. As shown in this

table, fuel management was not simulated, and no written checklists were provided. Here, for the convenience of rating the mental workload by the pilots, the entire flight was divided into three phases; 'to WPT(waypoint) A' which is the period from starting the flight to completing the descent to 1500[ft] and before making the heading change to intercept the final course or just before requesting radar vectoring; 'to final' which is the period succeeding 'to WPT A' until completing the landing configuration and the final turn; and 'on final' which is the period during tracking the glideslope and localizer, until making the flare and landing.

The experiment had pilots conduct a flight in a terminal area following the above procedure. As another experimental variable, the scenario was varied from flight to flight according to the degree of abnormality we intended to impose on the pilot. While flying along the nominal procedure, ATC sometimes requested an unexpected change in approach pattern as well as landing runway, or even a change of the landing airport. The following variables could be varied in scenarios: heading and position of the alternate runway; the ATC responses to the requests, meteorological variables such as the ceiling, wind magnitude, and its direction.

In order to simplify the experimental design, however, only the following three scenarios were employed in this research: a normal flight (Normal), a radar vectored flight to another runway on the same airport (Vectoring), and a flight with an unexpected situation where changing weather closes the airport and the aircraft must land at another airport not familiar to the pilot (Closed). These scenarios are summarized in Table 3. For the 'Vectoring' case, the landing runway was either No.1 or No.3. For the 'Closed' case, information such as the position of an alternate airport and its runway heading was not given to the pilot until the closing of the original airport. After he obtained the information about the new airport by asking for runway information, a map indicating the airport position and distance to the final fix from a certain VOR/DME station was given to him. Airport 2 was not equipped with an ILS. In this way, the change in the scenario was designed to affect only the workload experienced in the phase of 'to final'. The initial flight phase,'to WPT A' remained unchanged for all scenarios. The ratings for the phase 'to WPT A' were utilized to evaluate the consistency of rating behavior of each subject and the biases of the ratings with those of the other subjects possibly due to differences in their flight experience. Throughout the flight, the ceiling was set at 1000 [ft]. The magnitude of the wind was 10 [kt], and the wind direction was designed so the pilot could anticipate that it was changing direction.

Four instrument rated pilots (subjects 1 - 4) participated in the experiment. Their flight experiences ranged from 600 to 2400 [hours]. The order of the experiments was arranged so that the type of scenario was randomized, and the subject might perform 'Manual' flight and flight utilizing the 'Autopilot' alternately. Each pilot performed more than 20 flights, requiring a total of about 6 hours. The time required to accomplish one flight varied from 6 to 20 [min] according to the scenario. To obtain the entire data for each subject required three sessions, each consisting of several flights, and performed on different days.

After each landing, the subjects were asked to evaluate their workload for each flight phase as well as provide a subjective rating about their performance in the flight phase. The rating scales used in this study are shown in Table 4. It should be noted that the rating scores compared the workload experienced in the simulation with one usually experienced in each subject's daily flight; this enabled us to obtain a basis for generalizing the results to the actual flight environment. In addition, we measured the number of changes in switch settings and control stick positions, which were recorded in real-time.

RESULTS AND DISCUSSION

Excluding several initial normal flights for familiarization, scores of 16 flights for each pilot were obtained. Figures 5 - 8 show the subjective rating scores of flight performance and of each tentatively classified mental workload aspect for each subject. Results indicate that none of the subjects used the worst rating 5. The tendency of changes in ratings according to a change in scenarios which is shown in these figures, is consistent except for subject 3. For reasons discussed below we took the average for only the subjects 1, 2, and 4 after normalizing for the mean biases of the ratings of each subject. These biases were obtained by comparing the average ratings of their workload in the flight phase, 'to WPT A', for either 'Autopilot' mode or 'Manual' mode, because this phase was the same throughout all the flights. The following data manipulation was made by using the data of 1, 2, and 4. The raw ratings of each flight were shifted so that average ratings in the flight phase 'to WPT A' for either the 'Autopilot' or the 'Manual' mode for each subject became the same values as the overall averages of all the subjects for this flight phase and mode. These normalized workload ratings are shown in Figure 9. This figure gives the following results:

1. In the initial phase 'to WPT A', all the subjects rated their

workload to be very low for all three levels of behavior. Although the workload of the 'Manual' mode became higher than that of the 'Autopilot' mode, it appears that there was no significant effect of the autopilot on the workload in this phase.

2. In the phase 'to final', significant shifts of subjective workload were obtained by adding the autopilot and by changing the scenarios.

   a. Emergencies in the scenarios affected mainly the *rule-based* and *knowledge-based* workload in both the 'Autopilot' and 'Manual' modes. As the scenario became increasingly abnormal, i.e. in order of 'Normal', 'Vectoring', and 'Closed', both ratings of *rule-* and *knowledge-based* workload increased significantly. It should be noticed here that these increases are independent of availability of the autopilot. However, the averaged ratings of *rule-* and *knowledge-based* workload show the same tendency. This means that either the present variation of the scenarios or the present rating scales might not be sensitive enough to make the subjects distinguish *rule-based* workload from that which is *knowledge-based*. On the other hand, the *skill-based* workload was almost independent of any scenario changes.

   b. The effcts of autopilot were most evident in the workload of *skill-based* behavior. Although the ratings of the 'Normal' scenario showed the largest influence of the autopilot, it can be said that, regardless to the scenario, the autopilot consistently decreased the *skill-based* workload.

3. In the phase 'on final', the autopilot is effective in decreasing *skill-* and *knowledge-based* workload. On the other hand, workload was generally higher in the 'Manual' mode and 'without ILS' mode, which occurred in 'Closed' cases. One of the causes for this increase in workload came from having less information on visual perspective in the display than in the real world.

4. The subjective flight performance had a strong proportional relationship with *skill-based* workload. However, it appears that *rule-* or *knowledge-based* workload is independent of subjective flight performance or *skill-based* workload.

The ratings of the phase 'to final' given by subject 3 show considerable saturation due to large biases compared with the other three. It is difficult in this flight phase to see the influence of the autopilot. One of the reasons for these biases came from the differences of the flight experience of 3 from the others; the aircraft class of subject 3 was large transports, while the class of the others was smaller airplanes. This suggested to us that using a unified rating scale among pilots who have different flight experiences would not be quite valid. One possible method to avoid these saturations might be to improve the rating scales by specifying what actual tasks are within the three rating scales so that pilots can more easily compare the workload of each level with that of their actual experience.

Figure 10 shows the number of changes in switch positions and control deflections. This figure shows that even though the basic procedure listed in Table 2 reproduced a lower number of tasks than those of actual flight operations, it enabled us to control the density of the pilot tasks in each flight. The numbers of pilot actions per unit time, however, indicate no significant relationship with any of the workload ratings. This implies that even rather condensed tasks can be readily performed when the tasks are well arranged and predictable.

CONCLUDING REMARKS

In the present study, a laboratory type flight simulator has been introduced. It seems evident that this flight simulator, although simple, could successfully reproduce high workload situations even in an automated environment.

Pilot behavior in an automated cockpit was tentatively classified into three levels; *skill-*, *rule-*, and *knowledge-based*. The scores of the workload ratings corresponding to these three levels were compared among different experimental conditions. Experimental variables were availability of the autopilot and the degree of scenario abnormality.

The results, especially in the 'to final' phase, indicate that the autopilot was consistently effective in reducing only the *skill-based* workload, whereas the scenario change caused a significant increase in the *rule-* and the *knowledge-based* workload. This independence of causality between the autopilot effect on *skill-based* workload and the effect of scenario change on the *rule-* and *knowledge-based* workloads leads us to a key hypothesis about the nature of human mental workload. In addition, the results indicated the present ratings did not successfully distinguish the difference between subjectively judged

workload of *rule*-based and *knowledge-based* behaviors, and that the difference in the type of flight experience among the pilots might result in large biases of their workload ratings. Also suggested was the need to improve our multi-dimensional rating scales. It is our intention to continue investigation of qualitive descriptions of *knowledge-based* pilot behavior during an emergency flight.

REFERENCES

1. Anon.: *A300FF Minimum Flight Crew Determination*, Airbus Industrie, AI/VF-AI/TE4-023/82, 1982.

2. Sheridan,T.B.: *Mental Workload in Decision and Control*, Proc. of 1979 IEEE Decision and Control Conf., Miami, Fla., 1979, pp.977-982.

3. Bird,K.L.: *Subjective Rating Scale as a Workload Assesment Technique*, Proc. of 17th Annual Conf. on Manual Control, 1981, pp.33-39.

4. Hart,S.G.; Childress,M.E.; and Bortolussi,M.R.: *Defining the Subjective Experience of Workload*, Proc. of the Human Factor Society 25th Annual Meeting, 1981, pp.527-531.

5. Hauser,J.R.; Childress,M.E.; and Hart,S.G.: *Rating Consistency and Component Salience in Subjective Workload Estimation*, Proc. of 18th Annual Conf. on Manual Control, 1982.

6. Hart,S.G.; Childress,M.E.; and Hauser,J.R.: *Individual Definitions of the term 'workload'*, Proc. of 1982 Psychology in the DOD Symposium, 1982.

7. Childress,M.E.; Hart,S.G.; and Bortolussi,M.R.: *The Reliability and Validity of Flight Task Workload Ratings*, Proc. of the Human Factor Society 26th Annual Meeting, 1982, pp.319-323.

8. Rasmussen,J.: *The Role of Cognitive Models of Operators in the Design, Operation and Licensing of Nuclear Power Plants*, Proc. of Workshop on Cognitive Modeling of Nuclear Plant Control Room Operators, NUREG/CR-3114, ORNL/TM-8614, 1982, pp.13-35.

9. Goodstein,L.P.: *Descriptive Display Support for Process Operators*, Human Detection and Diagnosis of System Failures, Edited by Jens Rasmussen and William B. Rouse, Plenum, 1981, pp.433-449.

10. Heffley,R.K.; and Jewell,W.F.: *Aircraft Handling Qualities Data*, NASA CR-2144, 1972.

11. Anon. : *Air Traffic Control*, U.S. Air Traffic Service, Federal HIghway Administration, 7110.65c, 1982.

TABLE 1 AUTOPILOT FUNCTIONS

	Longitudinal Autopilot
1	Attitude Hold
2	Speed Hold
3	Altitude and Speed Hold
4	Automatic Landing
	Lateral Autopilot
1	Manual Heading
2	VOR Coupled Automatic Navigation
3	Localizer Coupled Automatic Approach
	Basic Stability and Control Augmentation System
1	Pitch CWS
2	Roll CWS
3	Yaw Damper
4	Dynamic Pressure Compensation

TABLE 2 NOMINAL APPROACH PROCEDURES

flight phase	position	procedures
start	5[nm] south of VOR 1	200[kt], 3000[ft] level flight to VOR 1
to WPT A	to VOR 1	call to Tower 1 request clear to WPT A (request runway information)
	over VOR 1	call to Tower 2 turn to WPT A
	to WPT A	request clear to landing descend to 1500[ft]
to final	over WPT A	turn to the outer marker
	to OM	decrease speed to 150[kt] tune the ILS lower flaps and gear
	over OM	final turn
on final	to decision	follow the glideslope and localizer
	at decision	flare and landing

TABLE 3 SCENARIOS

name	route
Normal	start VOR1 WPTA OM4 runway 4
Radar 1	start VOR1 Radar to WPTB OM3 runway 3
Radar 2	start VOR1 Radar to WPTC OM1 runway 1
Close 1	start VOR1 WPTA Radar self navigation to final fix (14[nm] north of VOR1) runway 2 (6[nm] north of OM3)
Close 2	start VOR1 WPTA Radar self navigation to final fix (4[nm] west of VOR2) runway 2 (3[nm]south of VOR2)
Close 3	start VOR1 WPTA Radar self navigation to final fix (2[nm] east of VOR2) runway 2 (4[nm] north of VOR2)

* Runway 2 is not equipped with ILS

TABLE 4 SUBJECTIVE RATINGS OF PERFORMANCE AND WORKLOAD

performance	How would you rate your overall performance compared with your normal performance? EXCELLENT. COULDN'T HAVE DONE ANY BETTER. 1 ABOVE AVERAGE PERFORMANCE. SOME ROOM FOR IMPROVEMENT 2 AVERAGE 3 BELOW AVERAGE. . . 4 POOR. I USUALLY DO MUCH BETTER. 5
workload caused by skill based behavior	How great was the workload for basic flying skills compared with the workload you normally experience? (climbs, descents, turns to headings, etc.) EASY. VERY LITTLE WORK TO DO. 1 SOME WORK, BUT LESS THAN AVERAGE. 2 ABOUT AVERAGE. 3 HEAVIER THAN AVERAGE 4 VERY HEAVY WORKLOAD. ALMOST GOT TO BE TOO MUCH. 5
workload caused by rule based behavior	How heavy was the procedural workload compared with the workload you normally experience? (procedures specify aircraft configuration, FAA and ATC regulations, etc.) The scale is the same as the above.
workload caused by knowlegde based behavior	How heavy was the judgement and/or decision workload compared with the workload you normally experience? The scale is the same as the above

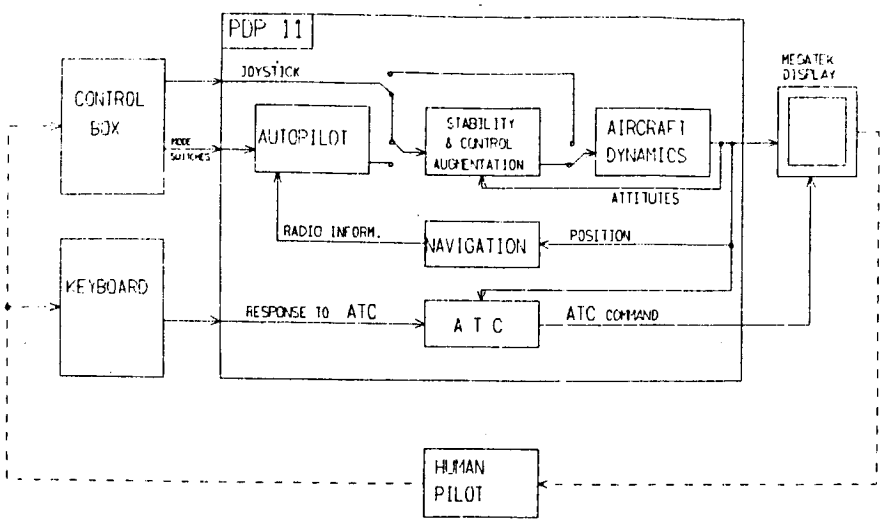


FIGURE 1. BLOCKDIAGRAM OF THE FLIGHT SIMULATOR

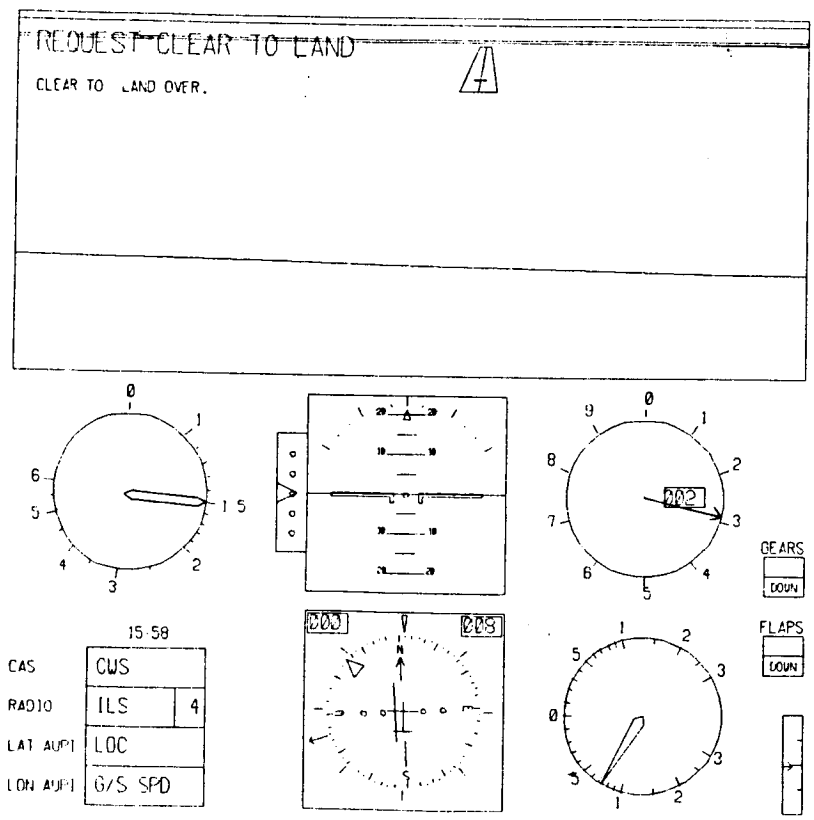


FIGURE 2. MEGATEK DISPLAY FORMAT

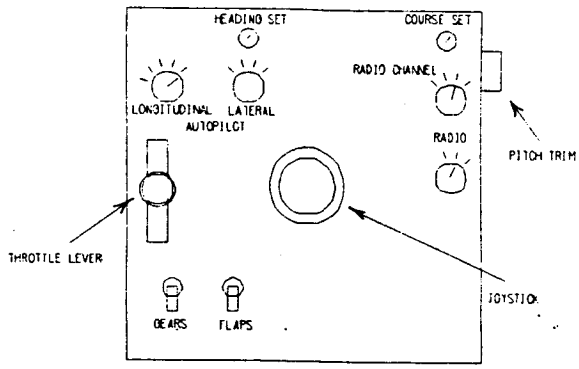


FIGURE 3. CONTROL BOX LAYOUT

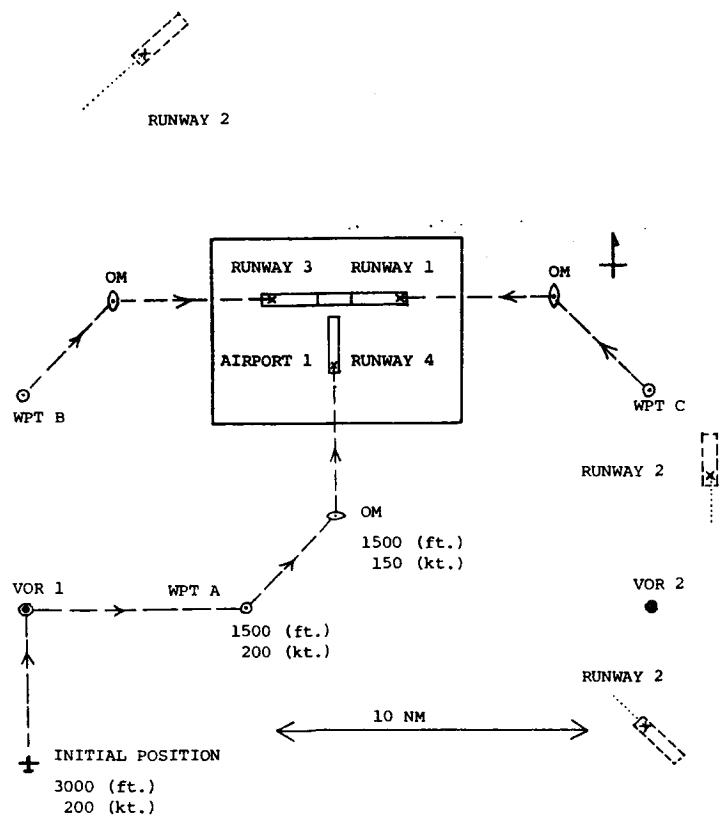


FIGURE 4. TYPICAL FLIGHT PATTERNS

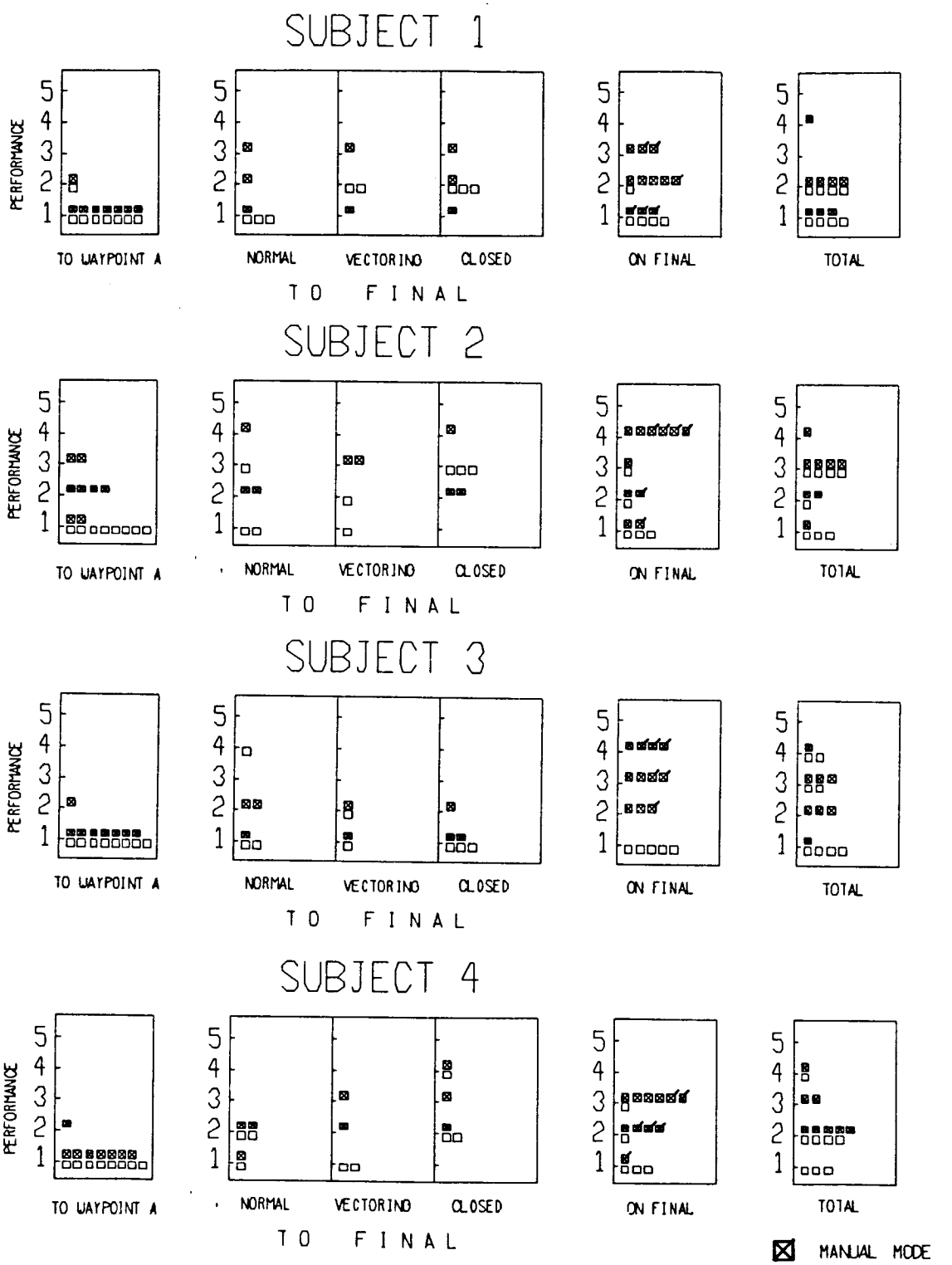


FIGURE 5. PERFORMANCE RATINGS OF EACH SUBJECT FOR EACH FLIGHT PHASE

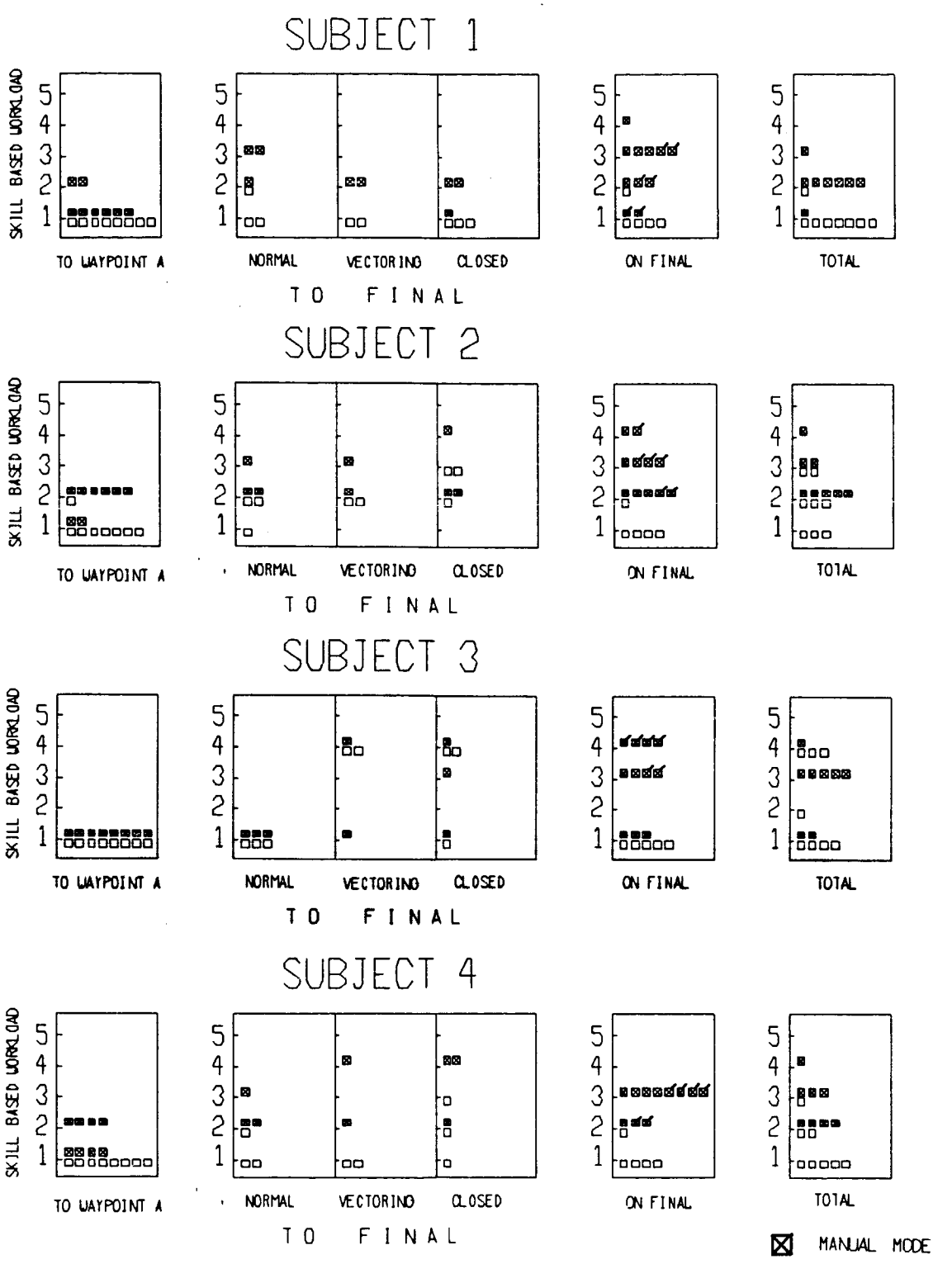


FIGURE 6. SKILL-BASED WORKLOAD OF EACH SUBJECT FOR EACH FLIGHT PHASE

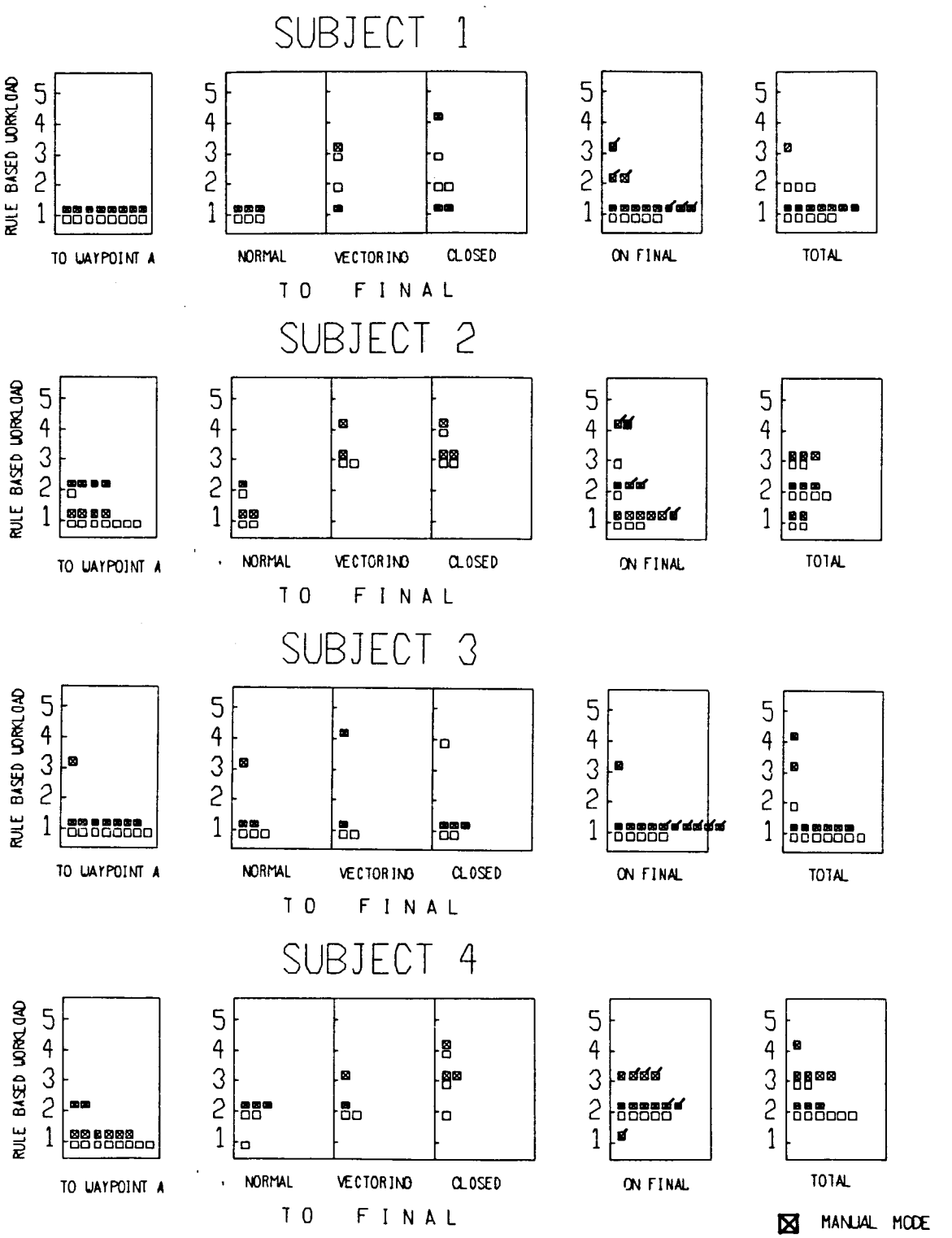


FIGURE 7. RULE-BASED WORKLOAD OF EACH SUBJECT FOR EACH FLIGHT PHASE

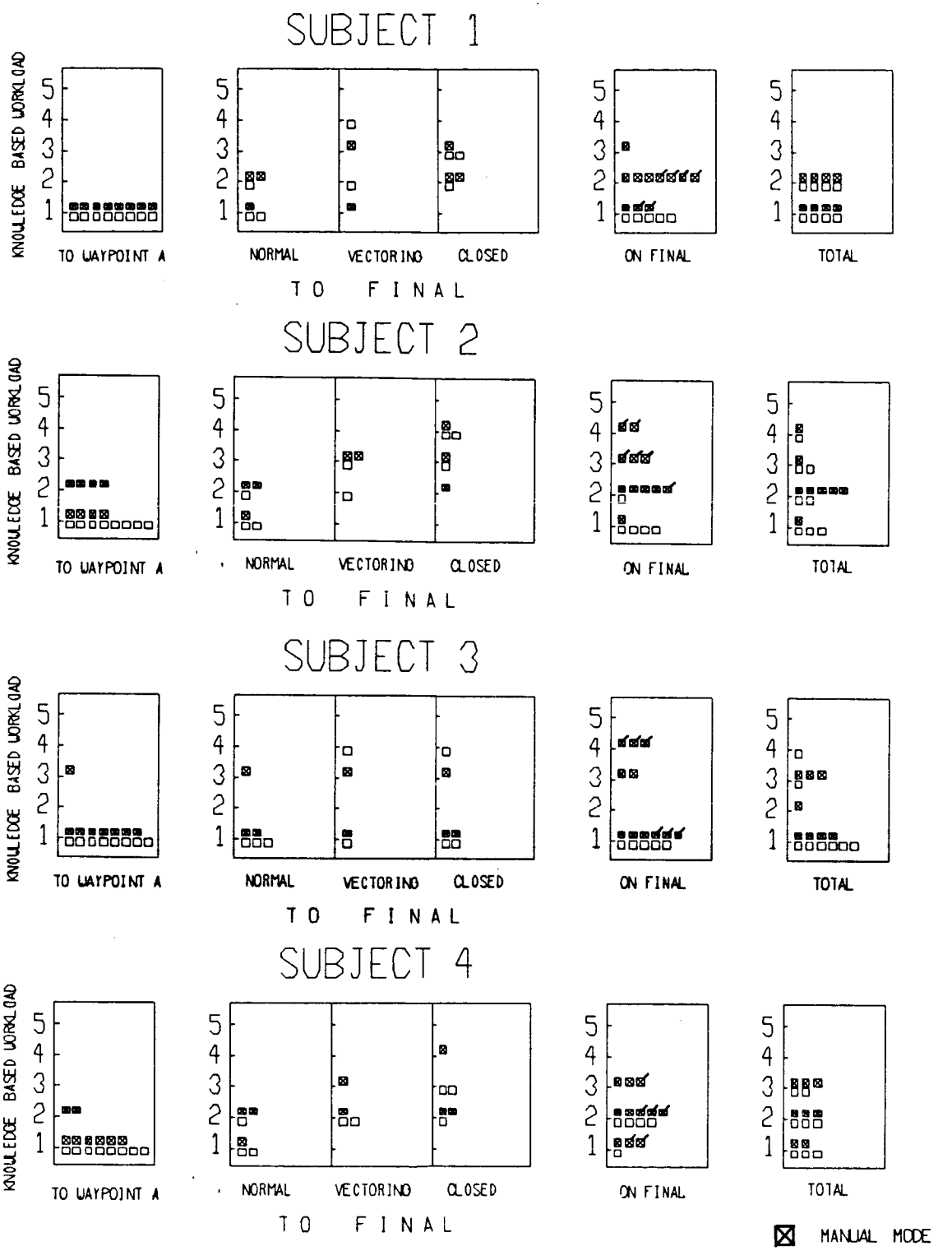


FIGURE 8. KNOWLEDGE-BASED WORKLOAD OF EACH SUBJECT FOR EACH FLIGHT PHASE

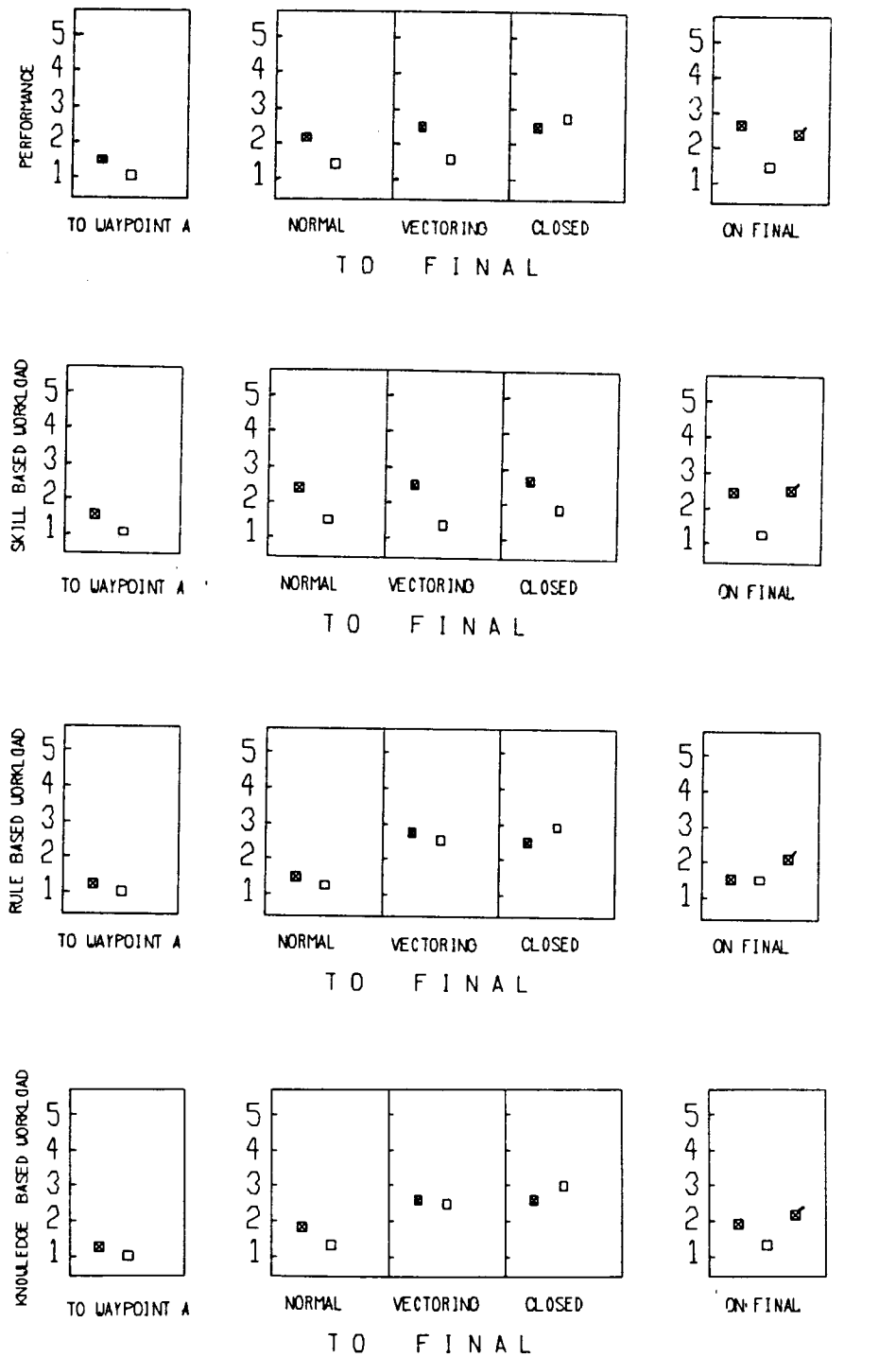


FIGURE 9. NORMALIZED WORKLOAD AND PERFORMANCE RATINGS OF EACH FLIGHT PHASE

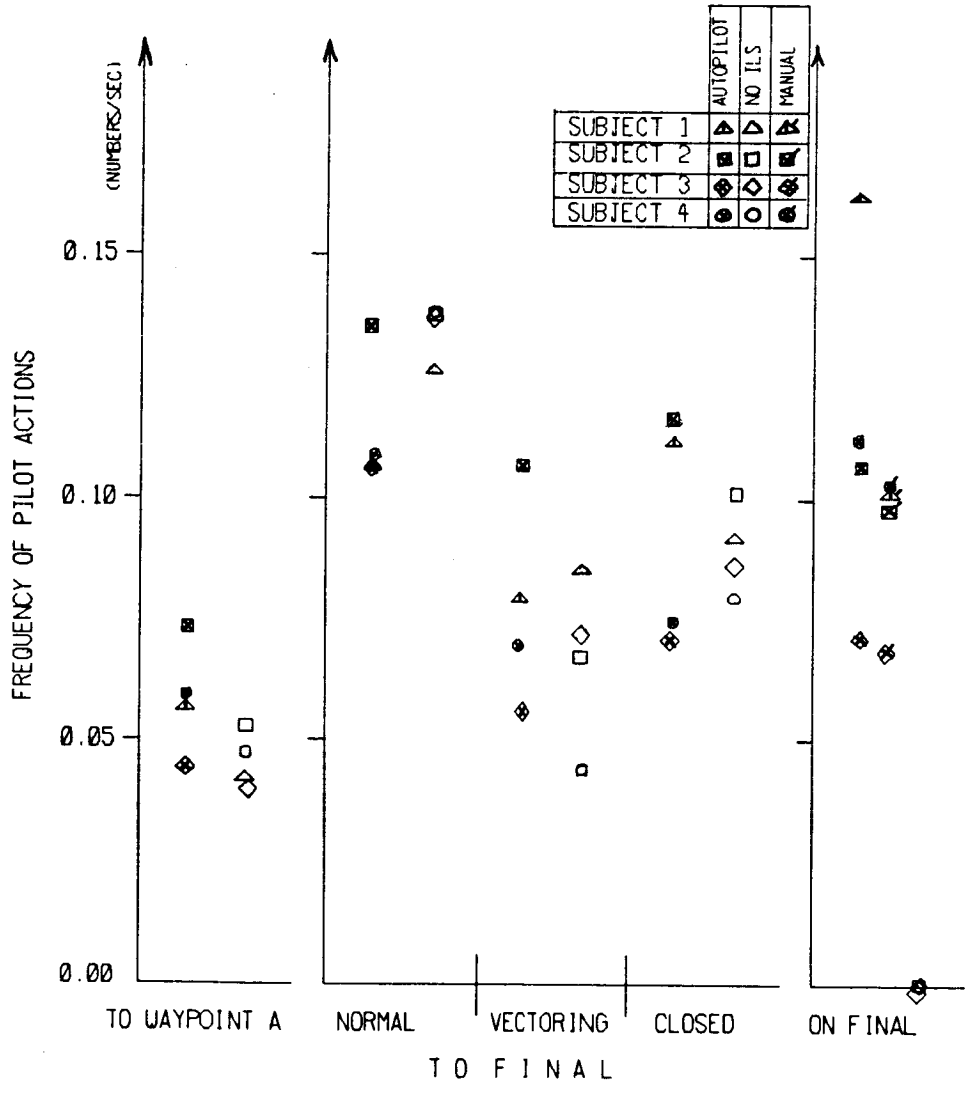


FIGURE 10. AVERAGE NUMBER OF ACTIONS FOR EACH FLIGHT PHASE

WORKLOAD ASSESSMENT IN HUMAN DECISIONMAKING

by

Eric P. Soulsby and David L. Kleinman
Department of Electrical Engineering
and Computer Science
University of Connecticut
Storrs, Connecticut 06268

ABSTRACT

It has long been accepted that in order to describe the overall behavior of a man-machine system, whether it be a manual control or a supervisory decision-making system, one needs to model the human's behavior in addition to describing the response of the non-human system components. In addition to measuring system performance, it has also been recognized that there exists a need to determine the level of "workload" imposed on the human operator in achieving design objectives. A central component of the human's "workload" is the amount of effort required in overcoming the uncertainty involved in making decisions. Several sources of uncertainty may exist, including external disturbance signals, human observation noise, a lack of a good internal model of the external system, etc. The relation between uncertainty and workload is explored in relation to a multi-task decision problem.

I. HUMAN OPERATOR WORKLOAD

In layman's terms, the phenomenon referred to as human "workload" is easily associated with how much the human must do to perform a specified operation. However, when attempts are made to quantify the "workload" imposed on a human operator by a particular design or operational procedure, the need quickly arises for a more precise specification of "workload", and for concomitant assessment criteria and techniques. As indicated in [4,6], specification of operator workload can be as diverse as the many disciplines involved in the man-machine research community. For example, the systems engineer may emphasize operational definitions based on the time available to perform a task, whereas psychologists have the tendency to emphasize information processing aspects of workload, defining it in terms of measures related to channel capacity. Physiologists take an alternate approach by emphasizing considerations of operator stress and arousal.

Intuitively, "workload" is related to the extent one is "mentally occupied" and to the effects of this occupation on the human organism. A good conceptualization of human operator workload is one which encompasses the three functionally relatable attributes: input load, operator effort, and performance or work result [4]. Sources of input load can be separated into three classes: environmental, design-induced or situational, and procedural. Examples of environmental variables are noise, vibration, temperature etc. Design or situational variables are concerned with characteristics of displays and control devices, vehicle dynamics, etc. Procedural variables involve instructions, task sequencing, and mission/task duration.

The performance level attained in alternatively designed systems can not alone be indicative of the amount of workload imposed on the operator. This follows from the notion that the same level of performance can be achieved under different loading situations. Therefore, the crucial ingredient in investigating human operator workload is the need to correlate operator effort exerted with the performance achieved. Operator effort may be envisaged as a function of the input load, the operator state, and an internal performance criterion. The operator state depends upon such factors as motivation, attentiveness, general background and personality. The internal performance criteria are maintained by the human operator and influence his tolerated error level [4,6].

A central component of the human operator's activity is in dealing with various sources of uncertainty in the work situation. As indicated in Fig. 1, these sources include

(a) external disturbance signal

(b) varying parameters of the system structure external to the human operator

(c) human-produced noise in observing the task stimuli

(d) lack of a good internal model of the external system

(e) human-produced distortions in interpreting the externally stipulated criterion of performance

(f) human-produced motor noise

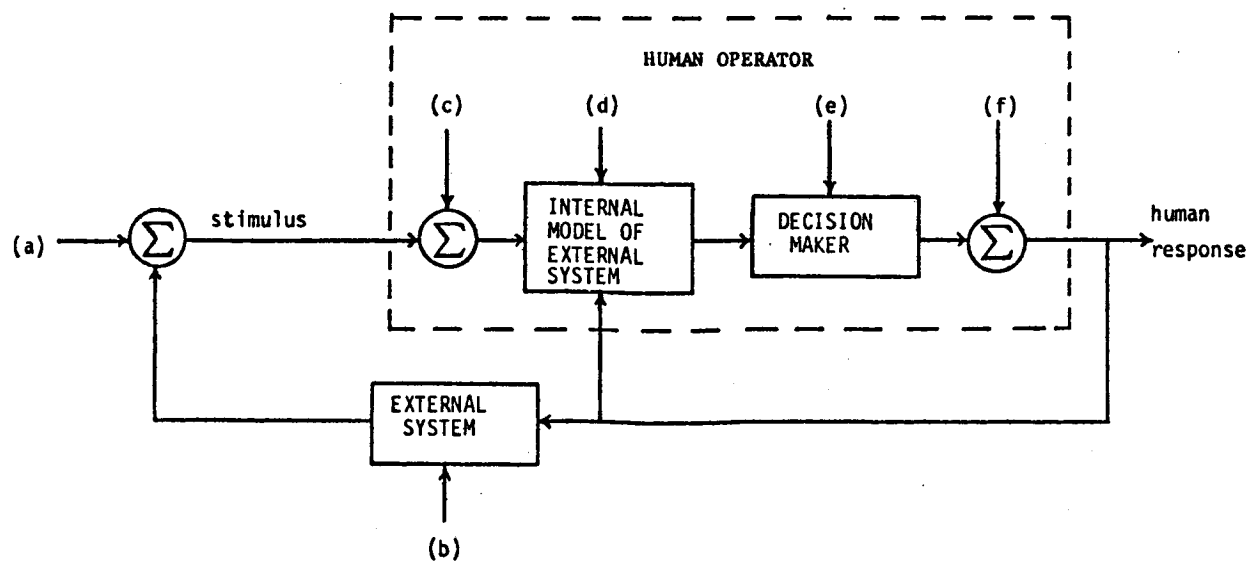


Fig. 1

It is a commonplace that humans use "images" or "internal models" of the world around them in organizing and executing their everyday activities. From the description of the man-machine environment presented in Fig. 1, one can conceive the notion that human operator workload can be related to the cognitive effort required to update or improve the operator's internal model and estimate of other signals in the system. The hypothesis here is that the better one's knowledge about the system (i.e. the better one's internal model), the less effort need be spent in achieving the task objective. Theoretically, if the internal model and statistical understanding of the various disturbance signals approach perfect matches to the actual, then the operator's behavior would become automatic and hence his workload would diminish to zero. An improved internal model is hypothesized as being an integral part of the mechanism by which the decisionmaker chooses a desired course of action. In other words, decision-making would become easier as the mental model of the decisionmaker approached a true "picture" of reality.

II. A MODEL FOR HUMAN DECISIONMAKING

Previous research at the University of Connecticut CYBERLAB was aimed at modeling single-human decision-making processes in multiple-task environments. The experimental effort focused on developing a canonical decision paradigm through which the decision maker's performance was studied as task parameters were changed [1]. This paradigm was motivated, in large part, by the C^3 problem of target selection. In this situation, targets of various type move across the display scopes of the human operator, vying for his attention. Since each target has different velocity and different distance to travel, the operator has a variable length of time available to process a target before it disappears. Each target has different threat value and processing time requirements. The human, therefore, is faced with the problem of sequencing tasks dynamically, so as to maximize the performance of the system.

Fig. 2 shows the fundamental decision loop that is considered in our

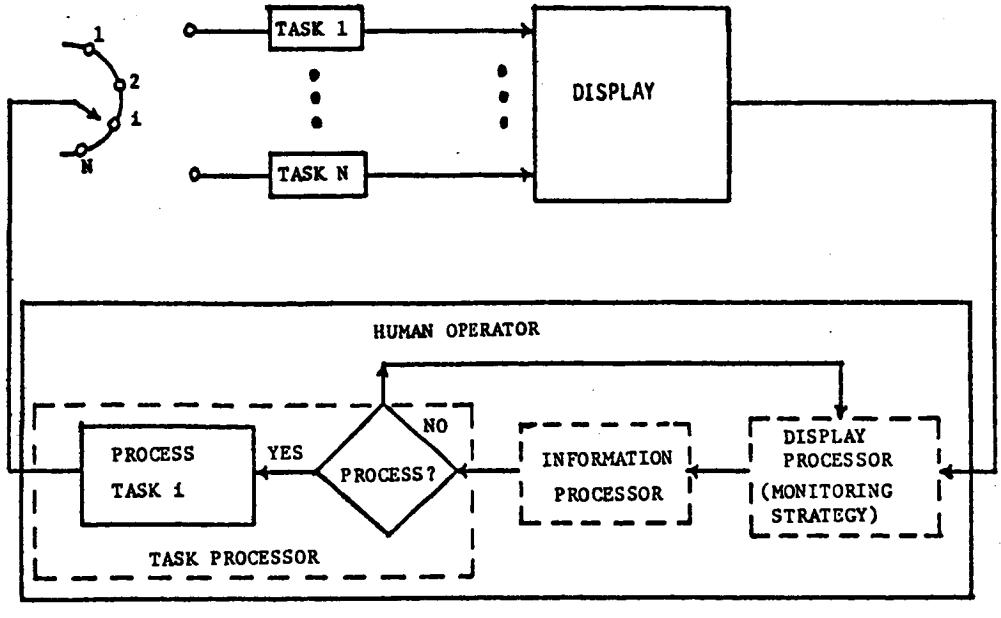


Fig. 2 MULTI-TASK DYNAMIC MONITORING/DECISION LOOP

approach. The human decision process involves 1) whether to process a task or gather more information (i.e. monitor); and 2) which of N tasks to act upon (N is time varying), in order to maximize the system performance (e.g. maximize reward, minimize regret, etc.). The decision loop is dynamic in nature. As time evolves, tasks of different value, duration (processing time), and opportunity window demand the human's attention, while others depart. The opportunity windows shrink with time as the tasks approach their dead-lines.

A simple, yet realistic, computer controlled experimental set-up was considered as indicated in Fig. 3. In the experiments, the subjects observe a CRT screen on which multiple, concomitant tasks are represented by moving rectangular bars. The bars appear at the left edge of the screen and move at different velocities to the right, disappearing upon reaching the right edge. Thus, the screen width represents an "opportunity window".

The height of each bar corresponds to the reward (value) of the task. The amount of time required to process a task (in seconds) is represented by the number of dots displayed on a bar. A task is processed by the subject when he pushes the appropriate push-button as shown in Fig. 3. By processing a task

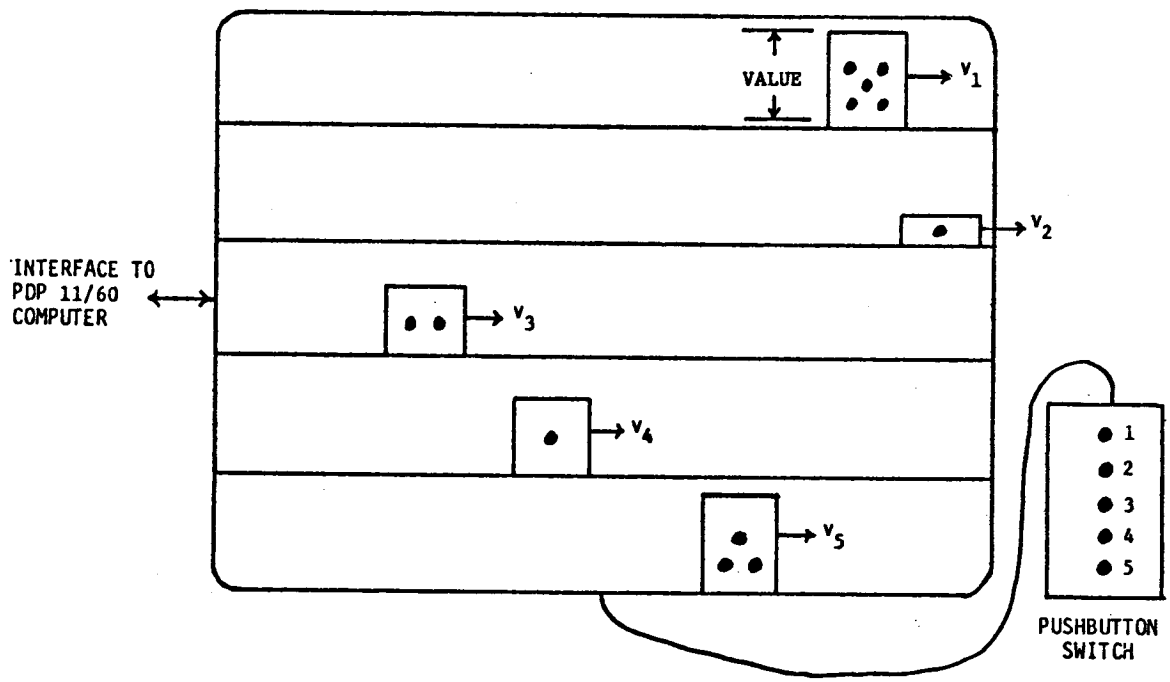


Fig. 3 EXPERIMENTAL PARADIGM FOR TASK SELECTION

successfully, the subject is credited with the corresponding reward (r_i), and the completed task is eliminated from the screen. An attempt by the decision-maker to act on a task that can not possibly be completed (i.e. the time required is greater than the time available) constitutes an error. No partial credit is given.

Retaining the essential features of the multi-task decision problem in a manageable, yet manipulative, context, this experimental framework was used in studying the effects of key task variables on human decision-processes [1]-[2].

Review of the Dynamic Decision Model [1]-[2]

The approach used in modeling human decision behavior paralleled that of the optimal control model of human response in spirit, but not in form. In the OCM, the control and information-processing strategies are separable; once an estimate of the system state is available, the linear feedback control law uses this estimate as if it were the true state.

In the present dynamic decision model (DDM), this type of separation has been found to be plausible. For any task i in the opportunity window, it is convenient to distinguish between the *task state*, $\underline{x}_{Ti}$, and the *decision state*, $\underline{x}_{di}$. The *task state* describes the dynamical variables internal to each task i; it consists of the instantaneous position and velocity of the bar, and the time required to process the task. The *decision state* provides the complete running summary of past actions (decisions), and is a memoryless functional transformation of the task state. The decision state variables consist of the time required to complete task i starting at time t, $T_{Ri}(t)$, and the time available/remaining to work on task i at time t, $T_{Ai}(t)$. The statistics of decision states, along with the task values, $r_i(t)$, and a performance metric, are used to compute the decision strategy. By analogy to the control-theoretic OCM, the values $r_i(t)$ play the role of cost functional weights, while the decision state variables correspond to system state variables.

A block diagram of the DDM structure is shown in Fig. 4. Each of the N tasks in the opportunity window is represented by a dynamic subsystem acted on by disturbances to account for (perceived) non-stationarities in the task characteristics. The perceived outputs $\{\underline{y}_{pi}\}$ are delayed, noisy versions of the task states $\{\underline{x}_{Ti}\}$, and are contingent upon the monitoring process. The perceived outputs are processed to produce the best linear unbiased estimates of the task states $\{\hat{\underline{x}}_{Ti}\}$, and their associated covariances $\{E_i\}$ via a Kalman filter-predictor submodel. The statistics of the task states $\{\hat{\underline{x}}_{Ti}, E_i\}$ are, in turn, used to determine the first and second order statistics of the decision states $\{\hat{T}_{Ri}, \sigma_{Ri}\}$ and $\{\hat{T}_{Ai}, \sigma_{Ai}\}$. The statistics of the decision states, along with the task values, $r_i(t)$, are combined to determine the attractiveness measure, $M_i(t)$, of each task in the opportunity window. Subsequently, the measures are used in a stochastic choice model based on Luce's choice axiom to generate the probability $P_{di}(t)$ of acting on each of the N tasks and the probability $P_{do}(t)$ of not acting on any task. The validation of the DDM results is presented in [2] and [5].

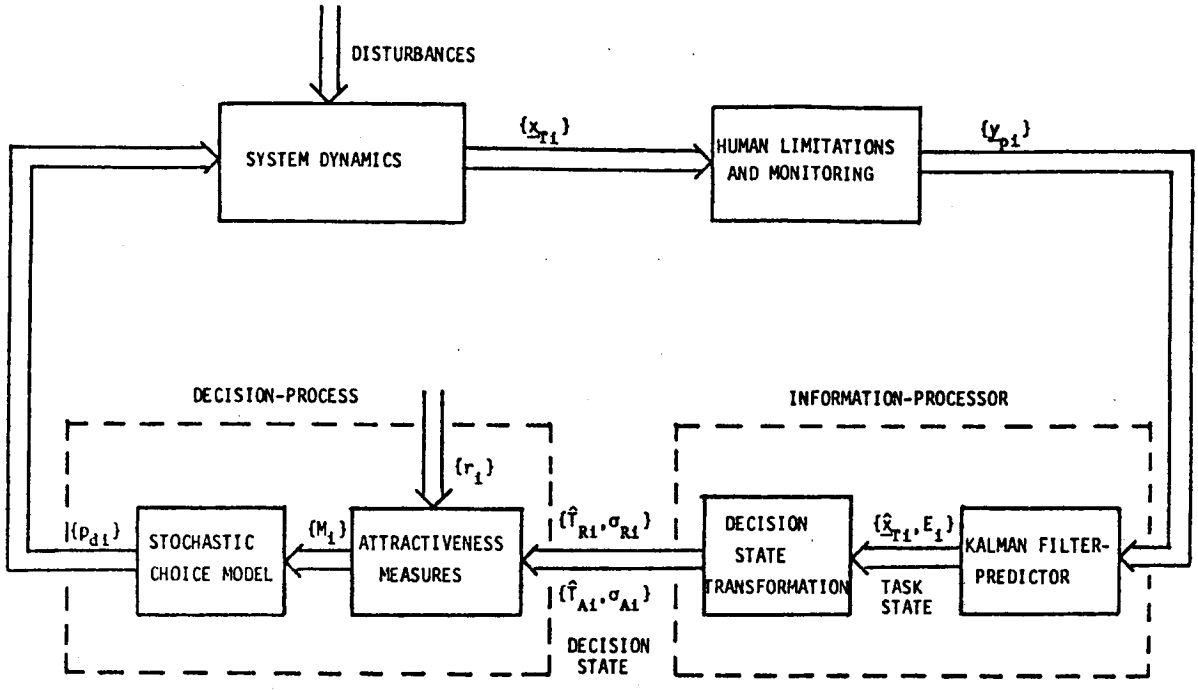


Fig. 4 DYNAMIC DECISION MODEL OF HUMAN TASK SELECTION PERFORMANCE

III. ASPECTS OF THE DDM RELATING TO WORKLOAD

In a supervisory, decisionmaking situation the human operator must process information presented en route to choosing an appropriate course of action. With regard to the multi-task selection paradigm, several sources of uncertainty must be dealt with. These include human produced distortions in observing variables presented on a display, uncertainty in determining the status or state of the system, various hypotheses as to possible courses of action, and the difficulty in envisioning consequences of actions subject to the overall task objective. Clearly all of these sources of uncertainty are influential in determining the amount of workload imposed on the human.

In the present experimental paradigm, from which the DDM was developed, several factors influence the decisionmaker's workload. The human must process information involving the amount of time required (processing time dots), the value of a particular task (height), and the position and velocity of the task relative to the task's opportunity window deadline. All of these factors must be considered for each of the N task lines (channels) that are simultaneously vying for the decisionmaker's attention. By processing this information (i.e. reducing his uncertainty about the variables presented), the decisionmaker then uses his knowledge of the "state" of the system to develop an appropriate choice of which task to act upon.

Not only must the human process several sources of input information en route to his decisionmaking processor, but he must do so at some rate corresponding to the rate of input information. The rate at which information must

be processed thus becomes a major factor in the workload imposed on the operator. Another viewpoint concerning this rate of information processing required by the human is the rate at which the decisionmaker must reduce his uncertainty about the system. Hence, the rate of change of uncertainty should reflect the effort exerted by the human in an attempt to "keep up" with increased information processing demands.

In the DDM the information processor portion of the model consists of a Kalman filter-predictor submodel. This submodel produces the best linear unbiased estimates of the task states and their associated covariances. The uncertainty with which the human "knows" the system is therefore exhibited via these covariance terms, and consequently the rate of change of this uncertainty should yield an indication of the human's information processing load. Similarly, the rate at which the human can reduce the uncertainty in the decision states (time available, time required) is reflective of the workload involved in making decisions. Intuitively this is appealing, for if the human's knowledge of the decision states was very poor one would expect that the course of action chosen would be arbitrary, reflecting the decisionmaker's inability to foresee the consequence of his action. On the other hand, if the human has sufficiently reduced the uncertainty surrounding these decision states then his immediate and perhaps future course of action is easily determined. Therefore, by investigating the rate of change of the covariances associated with the decision states, an indication of the amount of effort exerted by the decisionmaker (i.e. the amount of workload imposed) may be obtained.

IV. FUTURE DIRECTIONS

An aspect of human workload that has not yet been mentioned is the operator's response to increasing information. When subjected to information near or in excess of the operator's capacity, several mechanisms reflecting "strategy changes" are available to the decisionmaker. These include omission (nonprocessing of information), error (processing incorrect information), queuing (delaying processing of a task presuming that the human will "catch up" during a lull), filtering (neglecting to process certain categories of information while processing others), and employing multiple channels (processing information in parallel) [4].

Due to the adaptive nature of the human it seems that incorporation of these mechanisms into the DDM is warranted to compensate for workload situations involving strategy changes by the human. Some sort of thresholding capability which discriminates favorable tasks from unfavorable tasks in comprising the decisionmaker's action set would appear to reflect the adaptive nature of the decisionmaker.

In summary, we have demonstrated that a need for examining the workload imposed on the human in a supervisory or decisionmaking capacity is a vital ingredient in system design. In particular, the effect of input information load and its interaction with the effort required by the human in overcoming uncertainty is a crucial topic of investigation. By examining the quality of the human's internal model of the world about him and the rate at which the human improves this model, an understanding of the workload imposed may be achieved.

REFERENCES

[1] Pattipati, K. R.; Ephrath, A. R. and Kleinman, D. L., "Analysis of human decision-making in multi-task environments", Univ. of Conn. Tech. Report EECS-TR-79-15, Nov. 1979.

[2] Pattipati, K. R., "Dynamic decision-making in multi-task environments: theory and experimental results", Univ. of Conn. Tech. Report EECS-TR-81-9, August 1980.

[3] Kleinman, D. L.,; Soulsby, E. P., and Pattipati, K. R., "Decision Aiding - An Analytic and Experimental Study in a Multi-task Selection Paradigm", Proc. of the Fourth MIT/ONR Workshop on Distributed Information and Decision Systems Motivated by Command-Control-Communications (C^3) Problems, MIT, LIDS-R-1159, Oct. 1981.

[4] Soulsby, E. P., "Human Operator Workload: A Survey of Concepts and Assessment Techniques", Univ. of Conn. Tech. Report EECS-TR-82-4, Nov. 1982.

[5] Pattipati, K. R., Kleinman, D. L.,; Ephrath, A. R., "A Dynamic Decision Model of Human Task Selection Performance" IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-13, No. 2, March/April 1983.

[6] Moray, N., Mental Workload: Its Theory and Measurement, 1979, Plenum Press, N.Y. (NATO Symposium on Theory and Measurement of Mental Workload, Mation, Greece, 1977).

A WORKLOAD-ORIENTED MATH MODEL OF THE NAVY CARRIER LANDING
(Informal Paper)

Robert K. Heffley
Manudyne Systems, Inc.
349 First Street
Los Altos, CA 94022

INTRODUCTION

Math models of the task, pilot control strategy and controlled element can be instrumental in the analysis of such diverse topics as pilot workload, aircraft flying qualities, and even pilot skill development. However, the math modeling of the pilot-vehicle-task system must go beyond that of the conventional long-term continuous tracking task and address the time-bounded, deterministic, and discrete-control nature of many actual flight operations. In so doing it is also possible to appreciate more fully the role of the pursuit and precognitive-level pilot behavior which contributes to successful task execution.

This paper illustrates how task modeling can be used to examine a particularly crucial Navy mission flight phase, the carrier landing. The ultimate objective is to determine a means for drawing an explicit quantitative connection* between pilot workload and aircraft flying qualities requirements. The task model structure is given in same detail here although work is still ongoing to quantify model components.

The approach used to define the piloting task is based on the manual control theory point of view represented in Reference 1 but is augmented by recognition that the task itself is a major component in the overall system description. The closed-loop view of pilot performance is a major key to quantifying the task and pilot control strategy. The purpose of this paper is to illustrate how the overall carrier landing task can be credibly cast in such terms and especially how this permits effective analysis. The full daytime VMC carrier landing task depicted in Figure 1 can be stated in terms of a chronological series of perceptualmotor pilot-vehicle-task loop structures. Each component of the series is connected by cognitive procedural or decisional events. The result is a control-law program which provides an effective basis for exploring the sensitivity to any of the pilot-vehicle-task system

*Sponsored by the Naval Air Development Center, Warminister, PA, under contract N62269-82-R-0712.

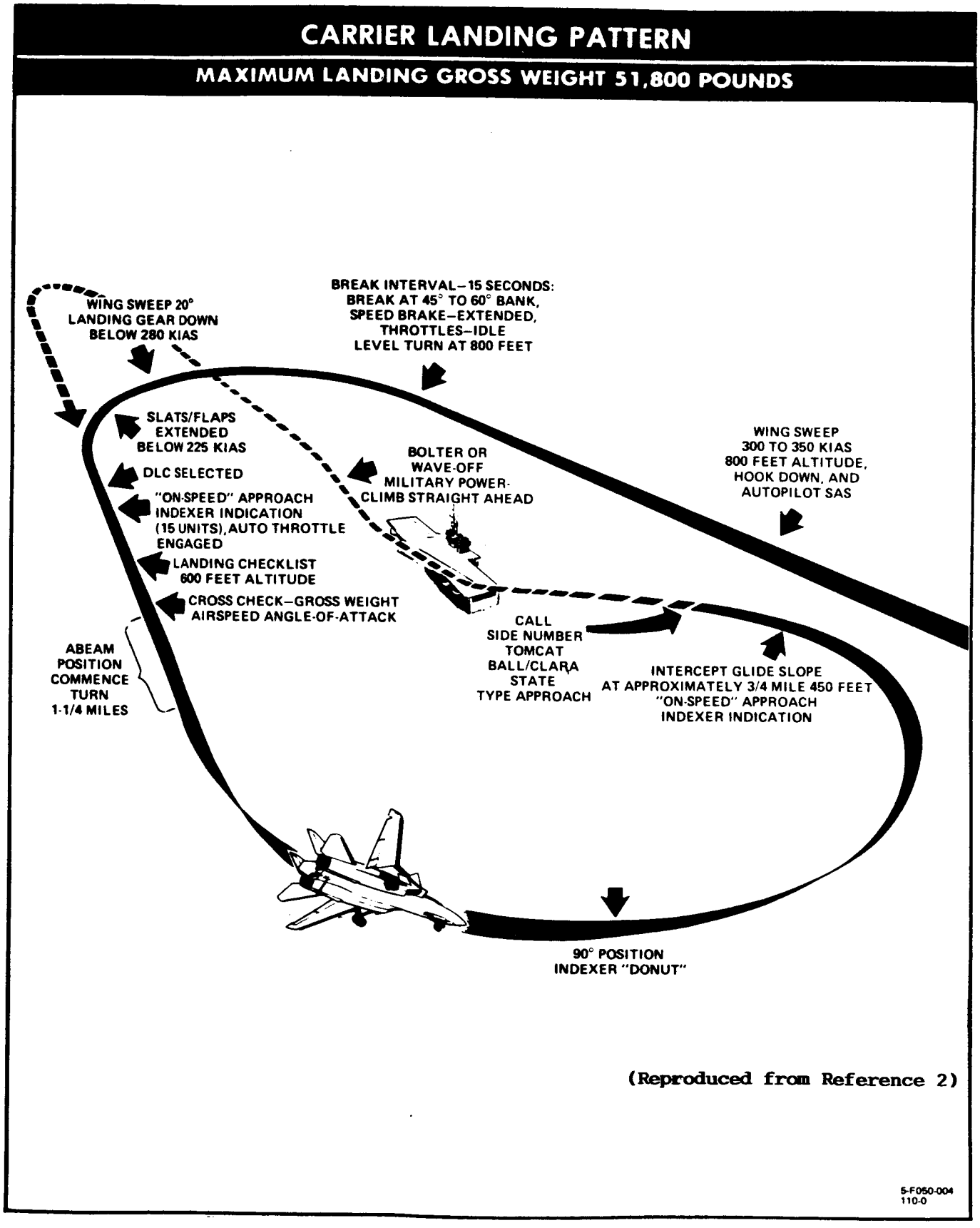


Figure 1. Carrier Landing Pattern

parameters. This provides a reasonably correct and complete operating context in which to examine the pilot workload as a function of aircraft flying qualities.

FEATURES OF THE TASK MODEL

The model presented here is intended to address three aspects of pilot workload suggested in Reference 3: (1) mental effort load, (2) time load, and (3) stress load. While "stress" is inherently difficult to express analytically, the mental effort and time loadings can be approached quantitatively. This can be done, in part, by using the relationship between excess control capacity and controlled-element form which is reflected in data from Reference 4, then augmenting it with the kind of discrete-maneuver pilot control strategy model described in Reference 5. The general idea is that time loading can be estimated by assessing the time available and time required for a limited-duration task or subtask. Mental effort can be estimated by representing some key feature of the uncompensated controlled element such as amplitude rolloff or phase angle at the effective operating point.

It is useful to break both the pilot control strategy and controlled element into units according to the control axis and support-loop roles. This at least allows some estimation of the the degree of difficulty (mental effort) of each loop taken individually. Figure 2 shows the components which can be used to define basic pilot control strategy for a given control axis and loop. When combined with the respective controlled element, this forms the pilot-vehicle-task system. The main system features which allow an analysis of time and mental effort workload include:

- **Task Duration**
- **Outer Loop Bandwidth**
- Outer Loop Controlled Element
- Support Loop Interval
- Support Loop Bandwidth

Time and mental effort aspects can be quantified for each control loop by evaluating the controlled element characteristics for the operating point corresponding to the time required to complete a finite-duration task or sub-task. In fact, a kind of closed-form time/mental effort tradeoff can be constructed if mental effort is expressed as, say, the uncompensated phase margin. Simply stated, the shorter the execution time for a discrete maneuver, the higher the bandwidth requirement and the lower the phase margin in both the outer and supporting loops.

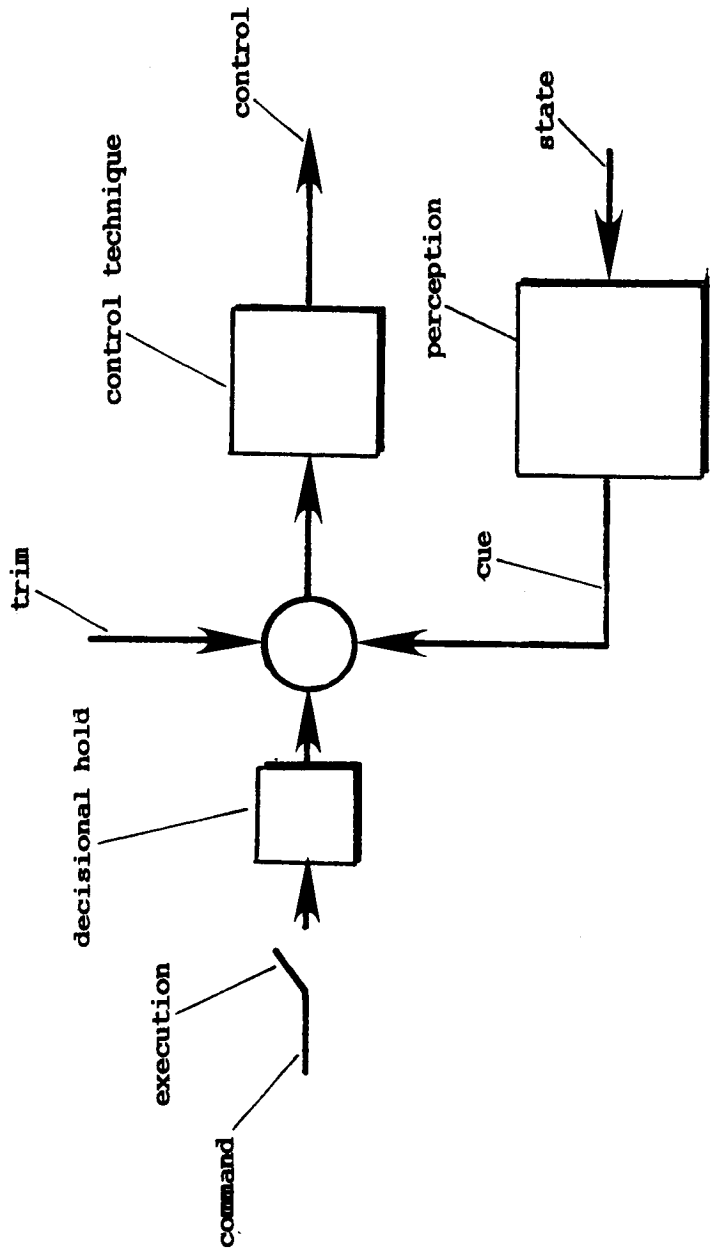


Figure 2. Pilot Control Strategy Topology

Figure 3 shows generically how the components of the task model can be divided into phase margin debits. Hence, we can express an explicit tradeoff between the phase margin, ϕ_{M_1}, (reflective of outer loop mental effort) and the time required to accomplish the task, T_1 (reflective of time load when compared to the time available for the task).

Note that the ϕ_{M_1} and T_1 tradeoff can be adjusted by the inner loop duty cycle, T_2. But, at the same time, T_2 is limited by the inner loop bandwidth available, ω_{c_2}. Such tradeoffs can be illustrated and explored in the analysis of the carrier landing.

DESCRIPTION OF TASK SEGMENTS

Based mainly on Navy F-14 fighter pilot interviews conducted at fighter squadron VF-111, NAS Miramar, detailed multiloop block diagrams of the pilot-vehicle-task system have been constructed for each segment of the carrier landing. These have been refined using F-14 flight data from the Naval Air Test Center, Patuxent River, MD (Reference 6). Training manual descriptions (Reference 7) were also consulted. The four major segments of the daytime racetrack pattern include:

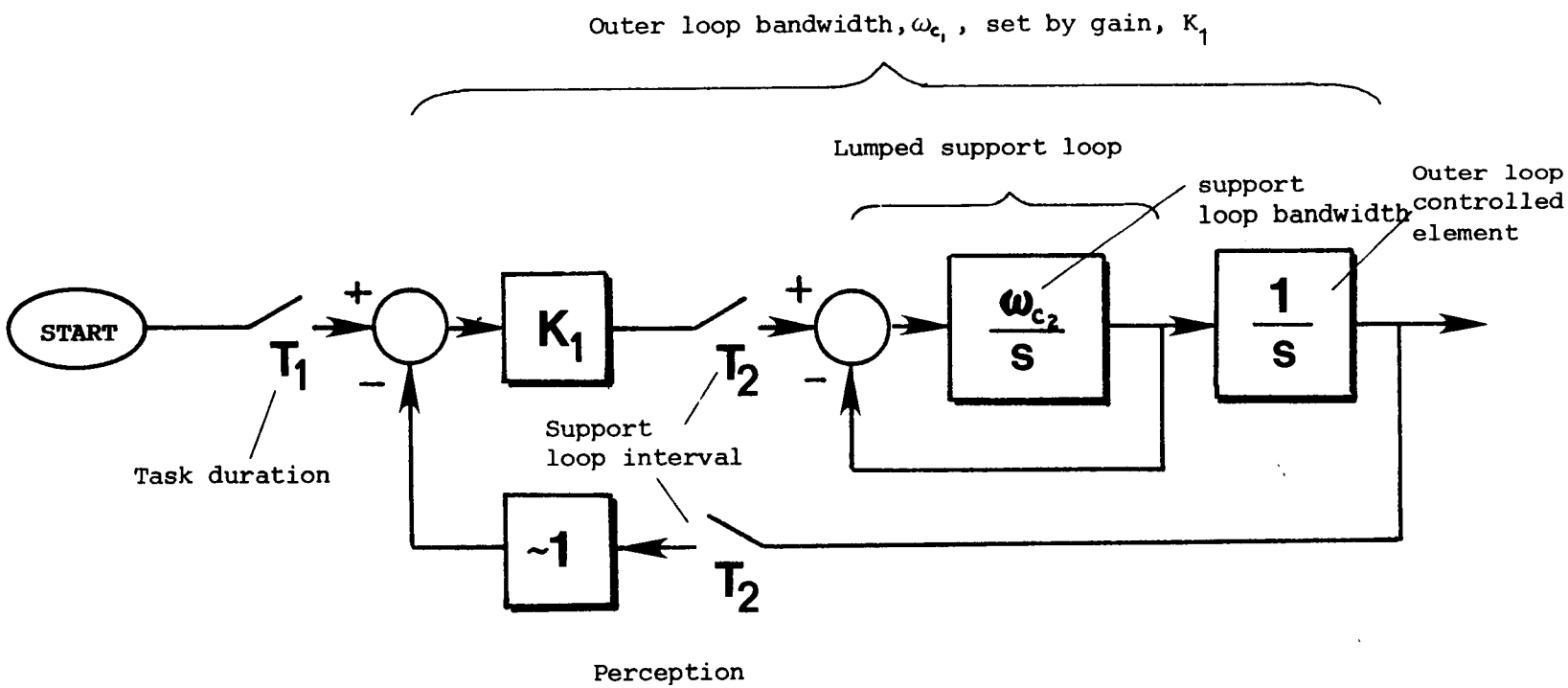
- Initial approach from astern
- Break (turn to downwind leg)
- Turn from downwind to final leg
- Final approach leg (using optical guidance)

Each of these segments is characterized by a fundamental shift in pilot control strategy and is described in detail below.

Initial Leg

The purpose of the initial leg is to arrive overhead the carrier on a standard course, heading, and altitude in preparation for executing the racetrack pattern. As shown in Figure 4a, this leg formally begins three miles astern the ship at 1200 ft and ends above or slightly beyond the bow. For the lead aircraft the main flight tasks during the initial leg are to arrive over the bow, on the base recovery course (BRC), and at 800 ft. altitude. (For aircraft flying formation on the lead aircraft, their task is limited only to maintaining formation and not to navigation.) Airspeed is set at the prerogative of the lead aircraft between 300 and 400 kt with the F-14's wings fully swept.

The pilot control strategy of Figure 4b, based on pilot descriptions, involves a compensatory management of course and altitude with



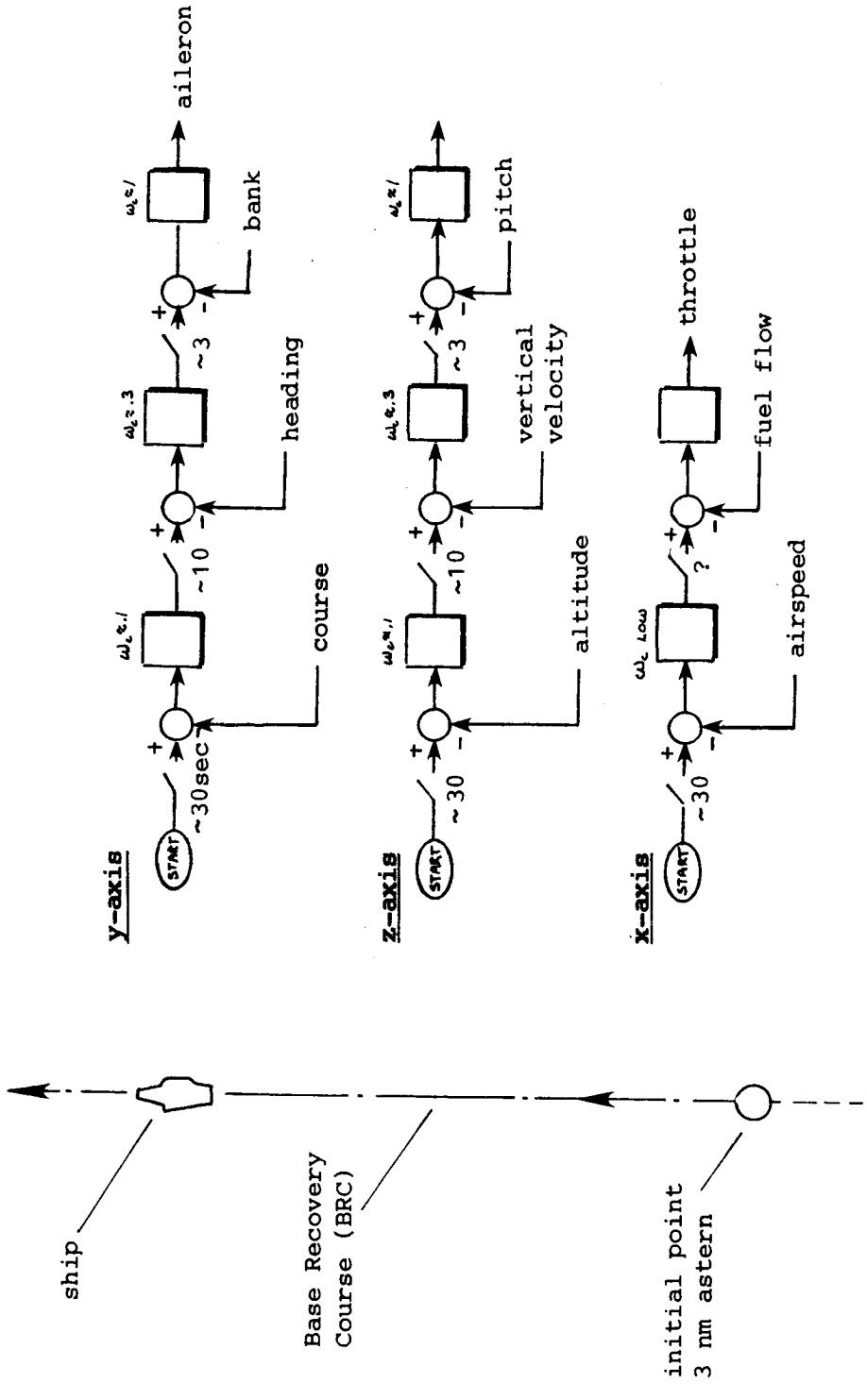
$$\text{Outer Loop Phase Margin, } \phi_{M_1} \approx \pi - \frac{\pi}{2} - \omega_{c_1}\frac{T_2}{2} - \tan^{-1}\frac{\omega_{c_1}}{\omega_{c_2}}$$

controlled element debit

support loop interval debit

support loop bandwidth debit

Figure 3. Discrete Maneuver Factors for a K/s Outer Loop



a. Nominal Trajectory

b. Likely Pilot Control Strategy

Figure 4. Features of the Initial Leg Segment

timing and loop gains set mainly by the 30 second duration of the leg. Both axes are similarly characterized by supporting middle and inner loops. That is, course is supported by a middle heading loop which, in turn, is supported by an inner bank angle loop. Altitude is supported by vertical velocity, and it in turn by pitch attitude. The third axis, airspeed, appears not to involve any substantial active regulation. Throttle or fuel flow rate is set at a nominal position and left there.

The controlled-element dynamics during the initial leg are benign. The 300 to 400 kt speed range ensures minimal effective lag in pitch, roll, and vertical flight path. Thus the controlled element is essentially either "k" or "k/s" for each of the three series loops in the two main control axes. The resulting mental effort required for these perceptualmotor tasks is therefore low. However, the large excess control capacity can be used up by decisional tasks connected with deck spotting and planning for a minimum-interval approach.

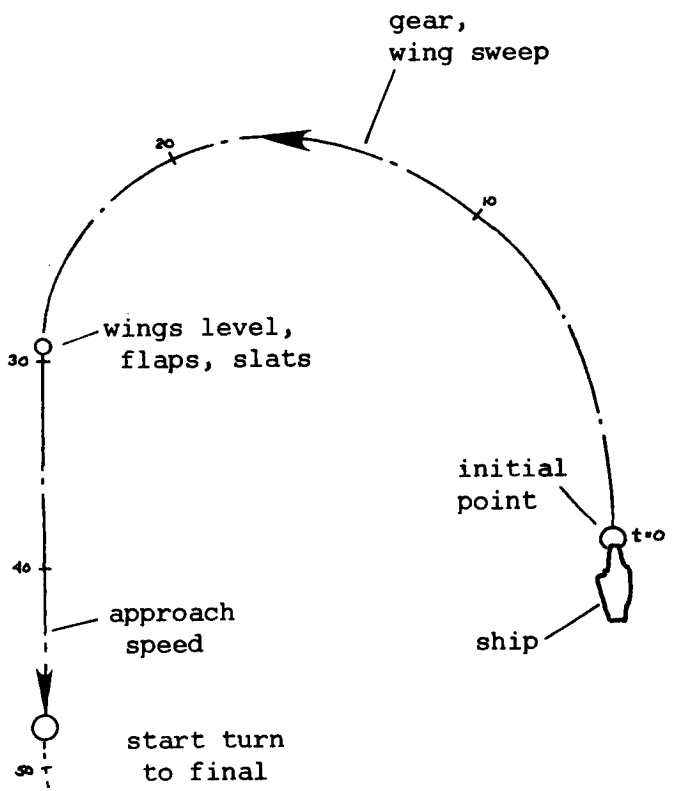
Break Maneuver

As shown in Figure 5a, this phase of the approach starts the 360 degree racetrack course and includes crucial deceleration and reconfiguration events. In addition a new course and altitude must be attained toward the end of the break on the downwind leg.

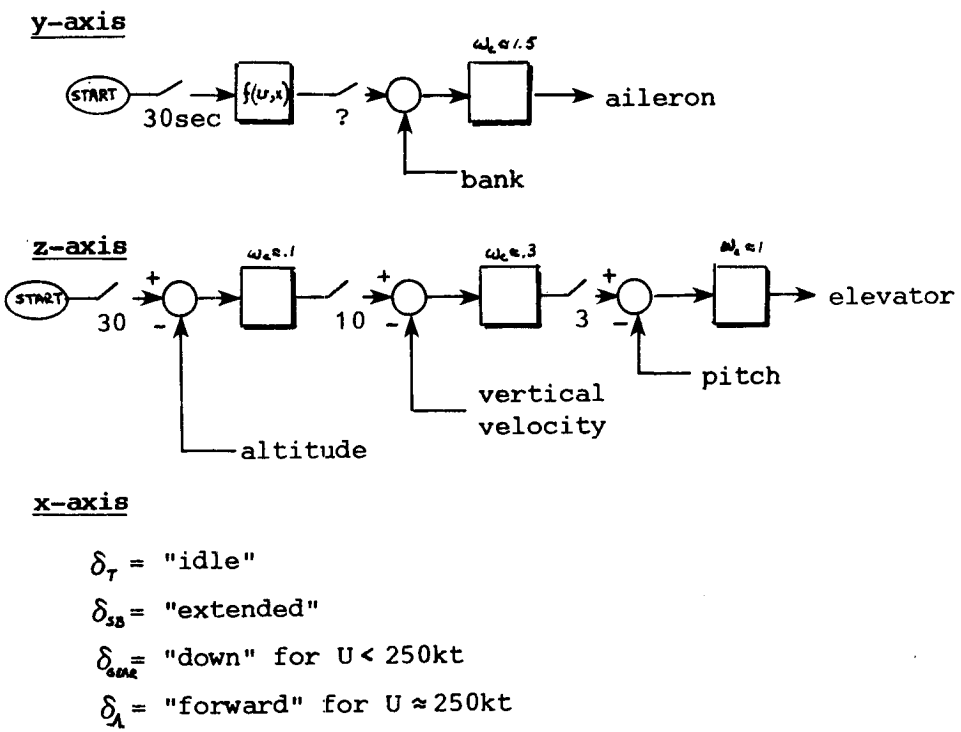
A dramatic change in pilot control strategy accompanies the break. Figure 5b shows that trajectory control is essentially precognitive as is airspeed. Only altitude retains the same compensatory character seen in the initial leg.

The desired horizontal-plane trajectory in the break a downwind course about 1.1 miles abeam the ship. This is achieved by an open-loop bank angle command at the start of the break. This bank can range from 45 to 70 deg depending upon initial airspeed and the pilot's judgement of the resulting nonuniform turn radius. No visual position cues relative to the ship are really available until well around the 180 degree turn. At this point a minor heading change might be used to adjust the distance from the ship.

Airspeed is a procedural matter determined by the reconfiguration sequence. Simultaneous with the break the speed brakes are deployed. A few seconds later the wings are unswept but not so early as to compromise the benefit of high induced drag. Then, as quickly as airframe limits allow, the gear is lowered and flaps extended to help the deceleration. Timely execution of each discrete step in the break can be crucial to the pilot arriving at the subsequent flight phase, prepared for the next set of tasks. It should be noted that the break maneuver involves several discrete actions which depend on airspeed and is therefore closed-loop in nature.



a. **Nominal Trajectory**



b. **Likely Pilot Control Strategy**

Figure 5. Features of the Break Maneuver Segment

Altitude control strategy in the break is similar to that of the initial leg although a pursuit technique involving pitch/throttle coordination is beneficial as the approach speed is reached.

Controlled-element features during the break are highly dynamic owing to the changing airspeed and high normal acceleration, but the pilot control strategy is fundamentally tolerant to this change. For example, an intermediate sink rate loop minimizes the effects of varying heave damping on flight path.

Turn-to-Final

This portion of the approach sets up the final leg (Figure 6a). Precise execution is essential for success since, as in the break, another open-loop lateral trajectory is involved. The controlled-element dynamics have now reached a relatively sluggish level compared to earlier phases, but they remain steady because speed is constant.

Shown in Figure 6b, lateral axis pilot control strategy is precognitive just as in the early part of the break. When precisely "abeam the LSO platform," the pilot executes a 27 deg banked turn toward the ship which is again necessary because of the absence of explicit lateral guidance cues. In effect, this segment is performed "on instruments".

Altitude continues to involve about the same control strategy as previous segments. Starting downwind at 600 ft the next target is 450 ft at the 90 deg point in the turn. A middle vertical velocity loop supports altitude. However, because of the low airspeed, a coordinated use of thrust and pitch must be used to support vertical velocity.

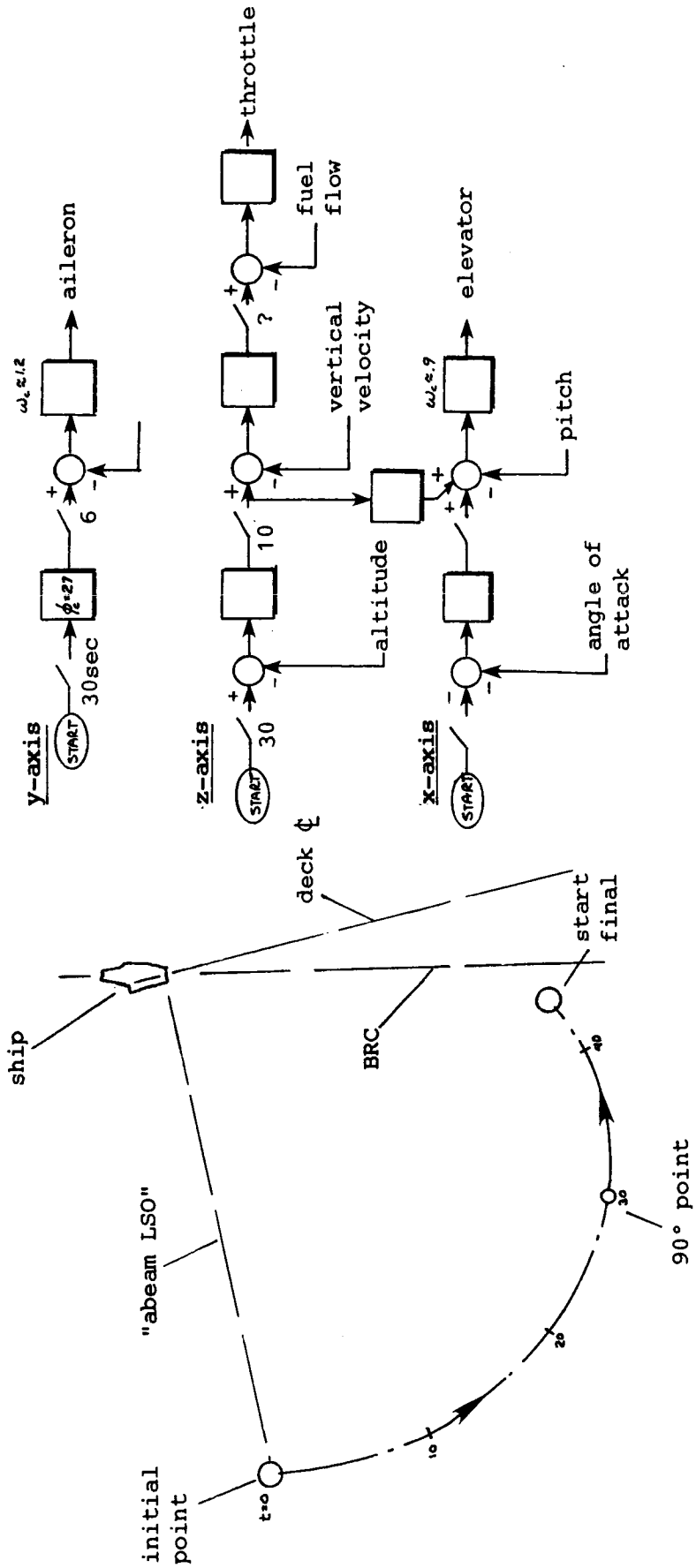
The x-axis involves a loose compensatory regulation of angle of attack by varying pitch attitude. Upsets to this axis are minimized by effective thrust/pitch coordination in flight path.

The controlled-element characteristics are typically "low-speed." Heave damping is low, speed damping high, adverse yaw a factor, and loss of lift due to lateral spoilers a problem. The last feature induces pilots to use lateral control sparingly in order to avoid upsetting sink rate, especially on the final leg.

Final Approach Leg

The final leg really begins while the aircraft is still in the turn to final (Figure 7a). This corresponds to the acquisition of final approach visual guidance -- the carrier Fresnel Lens Optical Landing System (FLOLS). The objective of this leg is to land precisely within the narrow confines of the deck arresting gear.

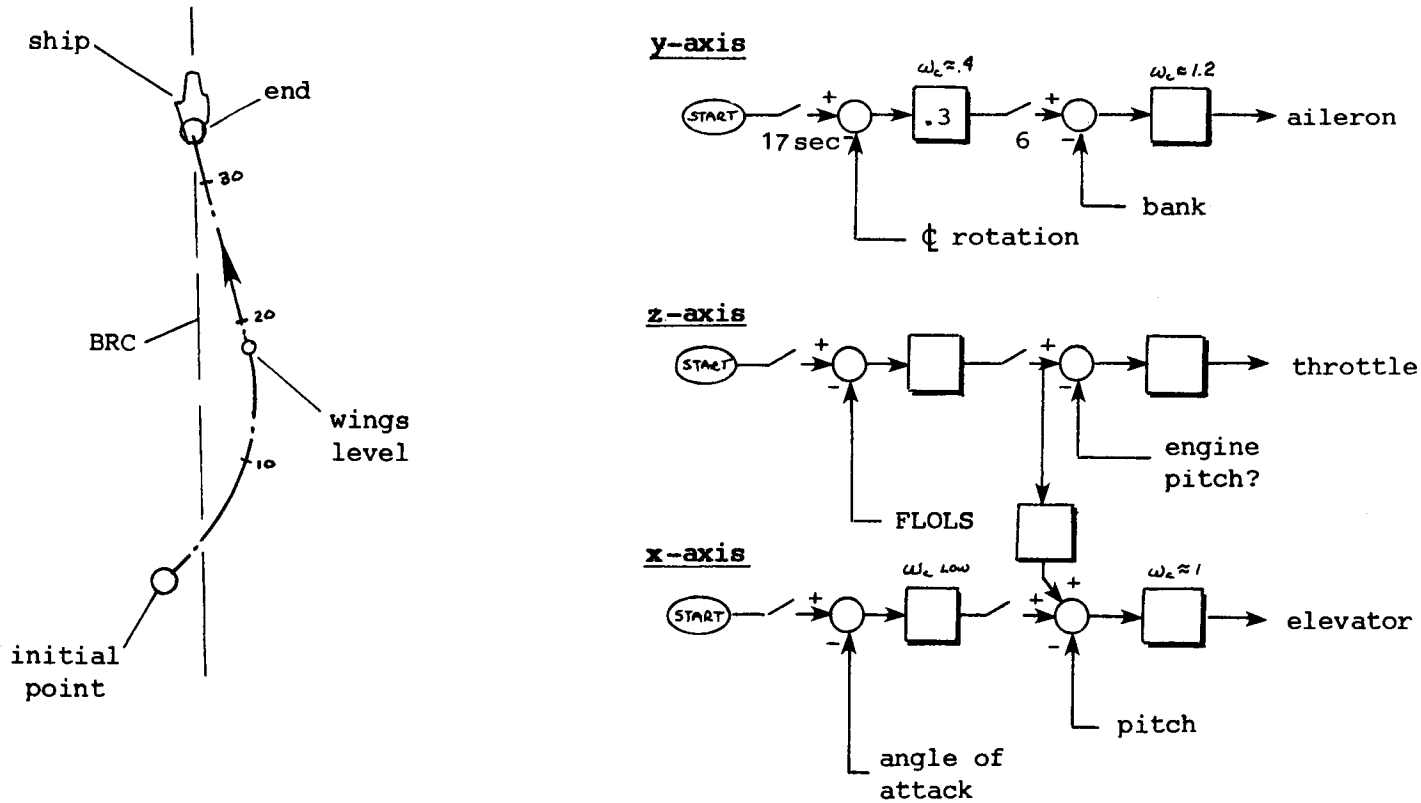
Pilot control strategy in the outer loops now adapts to a pursuit



a. Nominal Trajectory

b. Likely Pilot Control Strategy

Figure 6. Features of the Turn-to-Final Segment



a. Nominal Trajectory

b. Likely Pilot Control Strategy

Figure 7. Features of the Final Approach Segment

level. The vertical axis strategy is to null and stabilize the FLOLS, and for the lateral axis, to line up with the deck center line. Inner loop strategy must operate at a pursuit level in order to maximize outer loop bandwidth and minimize perceptualmotor mental effort. This is dictated by the short time-to-go (15 to 25 sec) and slow pitch, roll, and heave response. A pursuit crossfeed of pitch and thrust is needed to maximize path response and minimize angle of attack upset.

One additional aspect of the final leg is the pilot's interaction with the Landing Signal Officer (LSO). This is another source of flight path, position, and angle of attack information. The LSO assures the pilot of a clear deck or the need to wave off via light signals.

DISCUSSION OF WORKLOAD FACTORS

Analysis of the task segment trajectories and pilot control strategy diagrams given above provide a basis for estimating mental effort and time loadings during the carrier landing. Also the crucial cognitive events can be itemized. The following is a brief recap of some of the workload factors.

One important step in the perceptualmotor workload analysis is to examine the controlled element dynamics in the context of pilot control strategy. The effective controlled element response, say, for flight path, can vary significantly depending upon how the pilot chooses to manage it. As shown earlier in Figures 4 through 7, strategy is varied depending upon the demands of each task segment.

For the initial leg the controlled element lags are all minimal because of the high speed and the ability to partition the y-and z-axes into three loop structures. The x-axis requires little or no active regulation. A substantial excess control capacity in the initial segment permits deck spotting and planning for executing the racetrack pattern.

In the break the pilot's mental effort shifts to the x- and z-axes with the y-axis being mainly a precognitive banked turn. Here procedural tasks must be performed as quickly as airspeed reduction permits. This loading is not a fuction of time but rather of flight condition and will vary depending upon where and how fast the break was initiated. The closer to the ship and the higher the airspeed at the break, the more will the reconfiguration tasks pile up toward the end of the break maneuver. If not completed before the turn to final, they will begin to intrude on execution of the next task segment.

The turn to final marks the beginning of higher perceptualmotor loading and less cognitive. The pilot must hold a steady turn toward the ship, increase sink rate, and stabilize angle of attack in order to arrive on final in a steady, well-managed condition. Substantial pre-

cognitive behavior is evident such as holding a steady roll attitude, making a pre-determined fuel flow adjustment to set sink rate, and altering the nominal angle of attack to compensate for the effects of the turn. In this segment it appears that the pilot operates at high levels of control organization in order to maximize performance, while keeping mental effort and time loading manageable. Subjective assessment of workload is high at this point according to pilot commentary.

The final approach leg begins with various indications of lateral position relative to deck centerline. These include crossing the ship's wake and acquiring the FLOLS beam visually. Analyis of the rollout onto final has revealed an economical two-loop lateral axis structure involving the rotation of the centerline perspective as the outer loop and bank angle as the inner loop. This strategy permits quick lateral adjustments (about 17 seconds) and moderate bank angle bandwidth (about 1.2 rad/sec). It appears that the pilot has time for no more than two lateral corrections and about the same for the vertical. As discussed in Reference 8, a pursuit strategy is essential in the vertical axis in order to excute path corrections with acceptable mental effort in such a short period. In addition, an experienced pilot will apply subtle precognitive vertical path corrections just prior to landing in order to counter peculiarities of the carrier's air wake.

Carrier pilots emphasize that effective management of workload depends upon performing tasks on schedule and upon the degree of anticipation applied to making corrections. The adequacy of aircraft flying qualities, therefore, needs to be judged according to how well they support these rather deterministic demands as well as in countering random disturbances.

CONTINUING TOPICS OF STUDY

The model described herein is continuing to be refined and better quantified using all available flight data. The culmination will be a workload model (distinct from the task model) containing the factors of time, mental effort, and stress--structured to permit analysis of aircraft flying qualities. The carrier landing task model forms the operational context for applying the workload model.

Many of the numerical values given here are estimates, of course. Still needed are a comprehensive set of in-flight measurements to fill out the math model quantification. However, the form of the task models accomodates simple and direct parameter identification techniques such as suggested in Reference 5.

Construction and analysis of the carrier landing model has helped to identify some of the more crucial factors missing in the pilot workload data base. These include an understanding of how perceptualmotor elements build pilot workload in terms of multiple axes of control and

multiple support loops. Also, as the model demands, these aspects need to be evaluated in a time-bounded and realistic flight task context.

REFERENCES

1. McRuer, D. T., and E. S. Krendel, Mathematical Models of Human Pilot Behavior, AGARD-AG-188, January 1974.

2. Anon., NATOPS Flight Manual, Navy Model F-14A Aircraft, NAVAIR 01-F14AAA-1, October 1, 1980.

3. Reid, Gary B. and Clark A. Shingledecker, "Application of Conjoint Measurement to Workload Scale Development," Proceedings of the Human Factors Society--25th Annual Meeting, October 1981, pp. 522-526.

4. McDonnell, J. D., Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities, AFFDL-TR-68-76, December 1978.

5. Heffley, Robert K., Pilot Models for Discrete Maneuvers, AIAA 82-1519CP, August 1982.

6. Anon., Phase I NPE, F-14, Naval Air Test Center Report FT-03R-72, May 8, 1972.

7. Anon., Flight Training Instruction TA-4J, Part X CQ Stage Advanced Strike, Naval Air Training Command CNAT P-1545, 1978.

8. Heffley, Robert K., Warren F. Clement, and Samuel J. Craig, "Training Aircraft Design Considerations Based on the Successive Organization of Perception in Manual Control," Sixteenth Annual Conference on Manual Control, Massachusetts Institute of Technology, Cambridge, MA, May 1980, pp. 119-127.

HELICOPTER INTEGRATED CONTROLLER RESEARCH

Sherry Dunbar
San Jose State University
Psychology Department
San Jose, CA 95192

E. James Hartzell
US Army Aeromechanics Lab.
NASA Ames Research Center
Moffett Field, CA 94035

P. Madison
Informatics General, Corp.
Palo Alto, CA 94303

R. Remple
Informatics General, Corp.
Palo Alto, CA 94303

ABSTRACT

In an effort to improve the efficiency of helicopter flight and reduce pilot workload, the present research investigated the use of an integrated side-arm controller for the primary flight control of helicopters. Two types of integrated controllers were studied, with pitch, roll, and yaw as the axes of interest: (1) an isotonic controller where output is proportional to displacement; and (2) an isometric controller where output is proportional to force. The two types of integrated controllers were compared with the conventional configuration and with each other via performance on a three axis compensatory tracking task. Display error in the tracking task was generated by a combination of sum of sine waves and filtered white noise in such a way as to allow study of cross-coupling among the axes. A formal ANOVA was performed on the tracking data with RMS as the dependent measure. It was expected that both integrated controllers would result in better performance than the conventional configuration. Results of the analysis indicated that the isometric controller was better than the conventional controls, but the isotonic controller was worse. Several issues are discussed which may have contributed to this outcome. A subjective rating scale was administered; the results of which paralleled the objective data.

INTRODUCTION

Helicopters are being employed in an increasingly wide range of civil and military applications due to their unique flight capabilities. Such capabilities include nap-of-the-earth (NOE) flying, hovering, and precision landings in small volumes of airspace in crowded or otherwise inaccessible areas. These functions constitute the most rigorous of helicopter operations placing great demands on the skill of the pilot. The result of these demands is an increase in the pilot's workload. One of the primary effects of

increased workload is fatigue. This may result in performance degredation and the need to restrict the duration of flights to remain within the limits of the pilot's capacities. Performance degredation and shorter flights negate the benefits of the helicopter's unique flight and maneuvering capabilites.

The design of the cockpit has a direct influence upon the pilot's ability to perform at full capacity. When the design is best suited to the operator, the result is the most efficient use of the pilot's resources as a controller in the system. Although there are many aspects of helicopter cockpit design that merit attention, the specific area of man-machine interface emphasized in the present study involves the primary control elements.

The object of the present study was to examine the use of a multi-axis side-arm controller for the primary flight control of helicopters. The research was directed toward understanding tracking behavior in a multi-axis manual control task to provide guidelines for the development and improvement of helicopter controller design.

In most present day helicopters, the control system consists of three primary elements; the cyclic, collective, and rudder (tail rotor) pedals. Figure 1 illustrates the controls and their arrangement within the cockpit. The cyclic is center-floor mounted controlling aircraft pitch and roll attitude utilizing the right hand in a fore-aft and side-to-side motion, respectively. The collective is left-side floor mounted and directs changes in altitude using the left hand in an up-down motion. The pedals control yaw movement, counteracting the effects of torque and controlling heading in hover maneuvers, and are 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. In these taxing operations the pilot's hands and feet are very active because of the frequent and simultaneous control activity required. In addition to the fact that the pilot's resources are consumed by 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 and fatiguing posture to maintain for any length of time. Other disadvantages of the current configuration include inefficient use of cockpit space, and slow egress in emergency situations.

A need clearly presents itself to consider an alternative configuration in an attempt to alleviate the problems associated with this control arrangement. One possible solution to this problem is to combine the primary control functions into a single and more rationally positioned control element; termed an "integrated controller". This potential solution was examined in the present study by comparing two alternative control configurations with the conventional arrangement. Both of the alternative control configurations are integrated controllers in that the primary control functions of pitch, roll, and yaw are integrated into a single side-arm mounted controller. The difference between the two controllers is that one is an isotonic controller with output proportional to displacement, and the other is an isometric controller with output proportional to force.

The integrated controller concept is not a new one and has been the subject of several investigations (Mathews, Arnold, & Thomas, 1976; Waugh & Stevens, 1976; Lippay, McKinnon, & King, 1981; Landis & Aiken, 1982; McGee, 1982; and Landis, Dunford, Aiken, and Hilbert, 1983). One of the problems with the research to date is that the integrated controllers have been studied with the addition of stability and control augmentation systems which mask the raw manual control behavior. This behavior needs to be understood before the system becomes too complex, and especially before such a controller is considered for implementation.

It was assumed that a multi-axis side-arm configuration is a more sensible arrangement, and would therefore result in better performance. This assumption was tested by measuring performance on a three-axis compensatory tracking task intended to portray some of the manual control difficulty encountered in helicopter flight. The specific hypotheses were: (1) Tracking performance will be better with the isometric integrated controller than with the conventional controls; (2) tracking performance will be better with the isotonic integrated controller than with the conventional controls; and (3) tracking performance will be better with the isometric integrated controller than with the isotonic integrated controller. The research approach is unique and allows for analysis of the interaction between the integrated controller and the human operator, as well as provides information required in the development of an optimum controller.

METHOD

Subjects were eight non-helicopter pilots who exhibited at least a minimum amount of psyho-motor ability as measured

by the Jex Critical Tracking Task (Jex, 1966). The design was a 3 (configuration) X 3 (difficulty) X 8 (subjects) completely repeated measures. The three configurations were conventional, isotonic integrated, and isometric integrated, and the three difficulty levels were easy, medium, and hard.

The experimental task was a 3-axis compensatory tracking task utilizing the pitch, roll, and yaw axes, shown in Figure 2. The portion of the display marked "fixed" appeared and remained stationary on the CRT. The cursor was superimposed on the fixed portion and moved as a function of the difference between the input forcing function and the output of the plant. The subjects were required to make the proper control input to null the error in all three axes. Trial length was 100 seconds with the first 10 seconds unscored. Pitch error was displayed along the y-axis, yaw error along the x-axis, and roll error as angular deviation about the center. As illustrated in Figure 3, subjects were tracking an independent K/S plant in all three axes.

For a given trial, random appearing error was generated by white noise passed through a second order filter for two of the axes. For the third axis, random appearing error was generated by a sum of sine waves input forcing function of the same bandwidth and amplitude as the white noise axes. The component sine waves used are provided in Table 1. The input was structured in this manner so that from a describing function analysis of the output, the presence of sine wave components (which are known) occurring in the two axes comprised of white noise provides a way of detecting and assessing the magnitude of cross-coupling which is known to be a problem with multi-axis controllers.

To impose some realism on the tracking task, the relationship among the axes was structured to simulate a representative flight segment. This consisted of making the pitch the most demanding, followed by roll, with the yaw control the least demanding in approximately a 3:2:1 relationship, respectively. This was accomplished by manipulating the bandwidth of the input forcing function to produce the highest frequency of error in the pitch axis, lower in the roll axis, with the lowest bandwidth in the yaw axis. This relationship is illustrated in the top portion of Figure 4.

Each experimental session consisted of nine trials, three of each of the three levels of difficulty. The difficulty levels were obtained by varying the bandwidth of the input forcing function for a given trial while maintaining the bandwidth relationship among the axes. The bandwidth relationships among difficulty levels and axes are

illustrated in Figure 4. Within a session of nine trials, the three trials of a particular difficulty level were comprised of the three possible combinations of white noise and sum of sines input. This structure provides the potential of studying cross-coupling for any difficulty level and for any axis combination.

Subjects performed to asymptotic levels on each configuration. At several intervals during the experiment, the subjects completed a subjective workload rating (Hauser, Childress, & Hart, 1982) in which they rated fourteen items from zero to 100 percent based on their experience in the experimental situation.

RESULTS

Analysis of variance was performed on six sessions during which asymptotic levels were maintained, with Root Mean Squared (RMS) error as the dependent measure. For a given trial, RMS error was calculated in each axis separately, then summed to provide the trial error measure. These sums for a session of nine trials were then averaged to provide the performance metric. The describing function analysis has not been completed but will be the topic of a subsequent paper.

Figure 5 shows that the bandwidth manipulations produced the expected trends. Performance was rank ordered nicely for axes and difficulty levels, and there appeared a similar trend for each of the axes over the three difficulty levels with pitch highest overall, and yaw lowest.

It was hypothesized that both side-arm configurations would prove superior to the conventional configuration, but in fact this did not happen as shown in Figure 6. The isometric controller proved superior followed by the conventional arrangement, with the isotonic condition resulting in the worst performance. There was a statistically significant difference between the isometric and the conventional configurations ($p < .01$), a significant difference between the isometric and isotonic configurations ($p < .001$), with no difference between the isotonic and conventional configurations.

This was a somewhat puzzling finding, so variability was examined in the hope of providing an explanation. As seen in Figure 7, the pattern of variability among the configurations follows that of mean performance, with the lowest variability in the isometric condition and the highest variability in the isotonic condition. Subjects apparently did not have the same sense of control with the

isotonic or conventional controls that they had with the isometric controller.

Shown in Figure 8 are the axis relationships within the three configurations. Here a rank ordering of means across the axes occurred in accord with the bandwidth manipulations, except in the isometric condition for the yaw axis. Although the yaw axis was the least difficult to control (the lowest bandwidth), the error matched that of the roll axis. It seems likely that cross-coupling was present between the roll and yaw axes in the isometric configuration, making the yaw control difficult. Definitive conclusions can only be drawn concerning cross-coupling upon completion of the describing function analysis.

Figure 9 shows the subject's ratings on the workload items for each configuration. The subjective results paralleled the objective results where the highest average workload ratings are paired with the isotonic condition, and the lowest ratings are paired with the isometric condition.

DISCUSSION

The conclusion would seem to be that because of its obvious superiority both objectively and subjectively in this experimental situation, the isometric side-arm configuration is the direction to pursue in terms of implementing a multi-axis controller in the helicopter. At this time a definitive conclusion would be premature because there are several issues that must be addressed.

First of all, the type of tracking task used may have favored the isometric controller. A compensatory tracking task was used in this experiment which requires frequent quick responses to maintain control; the isometric controller is well suited to this type of response. On the other hand, a pursuit tracking task which requires smoother input to pursue a moving target, may favor the isotonic controller which is more suited to this type of response.

Another issue is that the use of an input forcing function with a high bandwidth as used in this study, may have also favored the isometric controller. The force controller has an inherently higher bandwidth and can therefore track high frequencies more easily than displacement controllers.

Finally, although the gains were set on the three configurations to "feel" as equivalent as possible, the mechanical differences among the controllers make this very difficult. For instance, a partial explanation for poor performance with the isotonic controller may be the fact that the

breakout force was too high and/or the force gradient was too low, making for a somewhat clumsy controller.

The research reported was the first in a series of experiments planned to investigate issues present with the use of a multi-axis side-arm controller. The issues discussed above will be the first to be addressed. The main contribution of this first study was the methodology; the development of a multi-axis manual control framework that can be used to study, in detail, the man-machine interaction in multi-axis tasks. From this we will gain information and direction needed to fine tune our methods so as to make a solid contribution to the data base concerning integrated controllers.

REFERENCES

Hauser, J. R., Childress, M. E., and Hart, S. G. Rating Consistency and Component Salience in Subjective Workload Estimation. In Proceedings of the Eighteenth Annual Conference on Manual Control, Dayton, June, 1982.

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.

Landis, K. H. and Aiken, E. W. An Assessment of Various Side-Stick Controller/Stability and Control Augmentation Systems for Night Nap-of-the-Earth Fight. Presented to the American Helicopter Society, Palo Alto, June, 1982.

Landis, K. H., Dunford, P. J., Aiken, E. W., and Hilbert, K. B. A Piloted Simulator Investigation of Side-Stick Controller/Stability and Control Augmentation System Requirements for Helicopter Visual Flight Tasks. Presented at the 39th Annual Forum of the American Helicopter Society, St. Louis, May, 1983.

Lippay, A., McKinnon, G. M., and King, M. L. Multi-axis Manual Controllers A State-of-the-Art Report. In Proceedings of the Seventeenth Annual Conference on Manual Control, Los Angeles, 1981.

Mathews, M. S., Arnold, J. R., and Thomas, C. L. Integrated Controller Evaluation. US Army Aviation Engineering Flight Activity Project No. 75-24, 1976.

McGee, J. An Experimental Investigation of the Interference Between a Right-hand Sidearm Controller Tracking Task and a Left-hand Switch Operation Task. United Technologies Sikorsky Aircraft Document No. SER-510079, January, 1982.

Waugh, J. D. and Stevens, J. A. Helicopter Integrated Control (GAT-2H). US Army Human Engineering Laboratory Technical Memorandum 39-76, December, 1976.

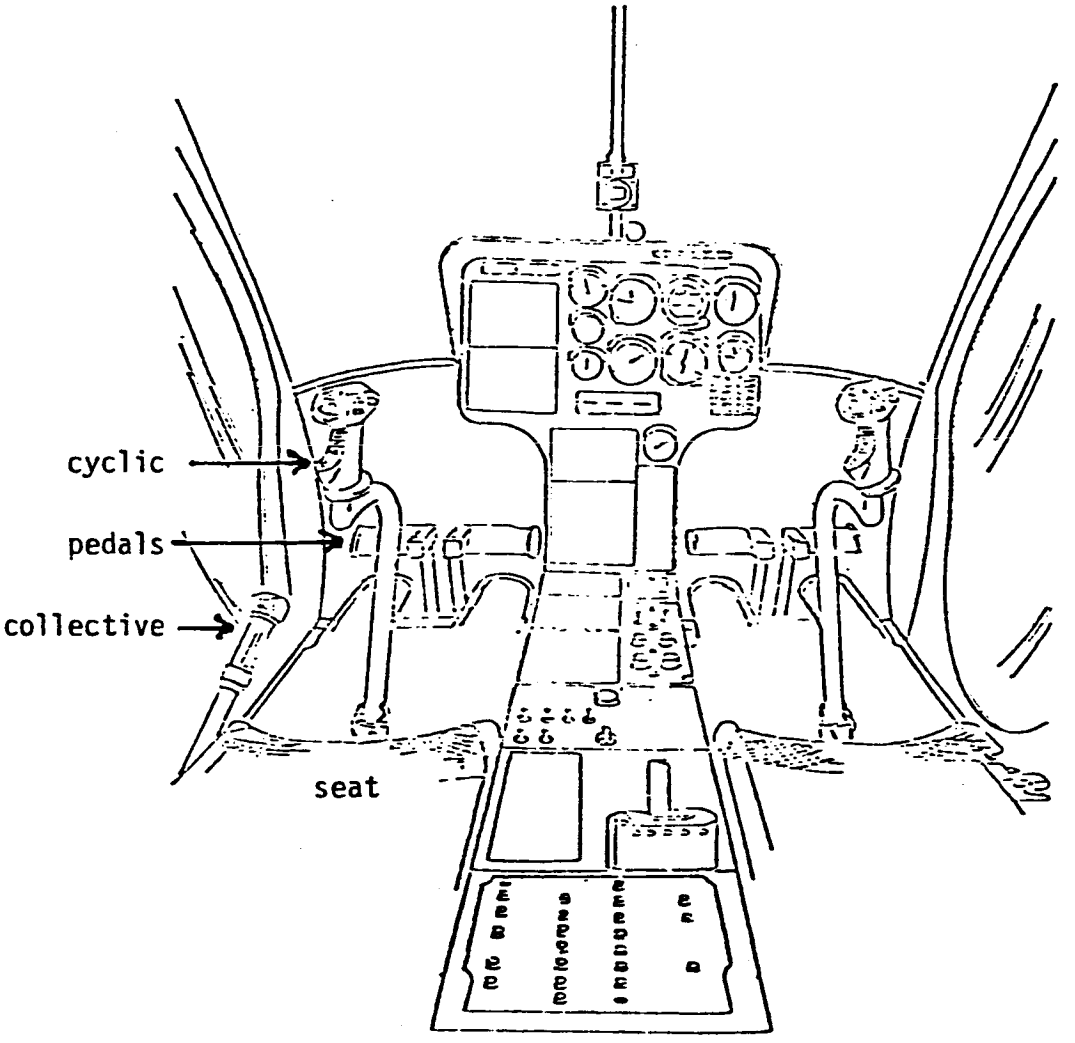


Figure 1: Conventional (Current) Helicopter Control Configuration.

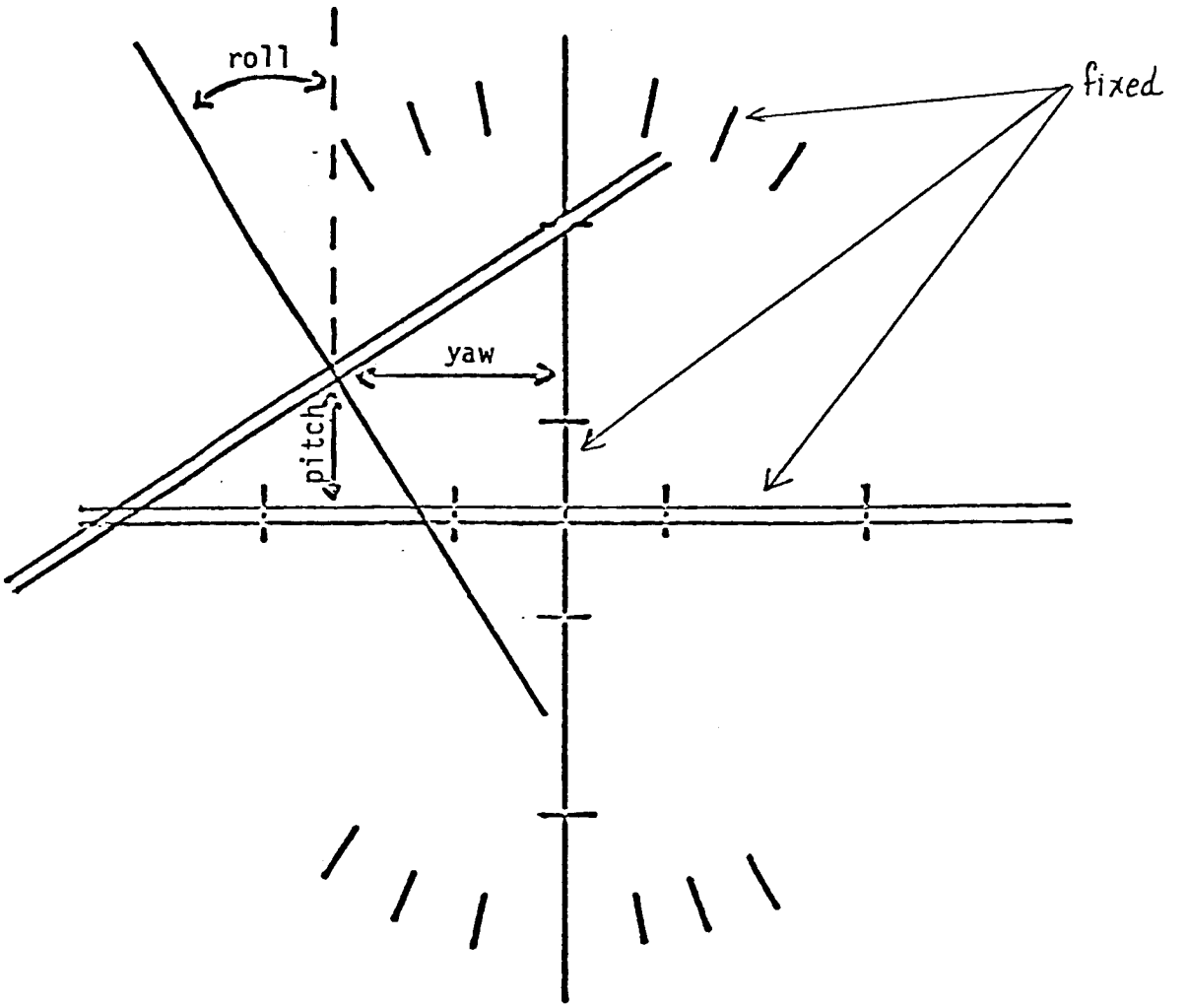


Figure 2. Experimental task display showing error in the three axes.

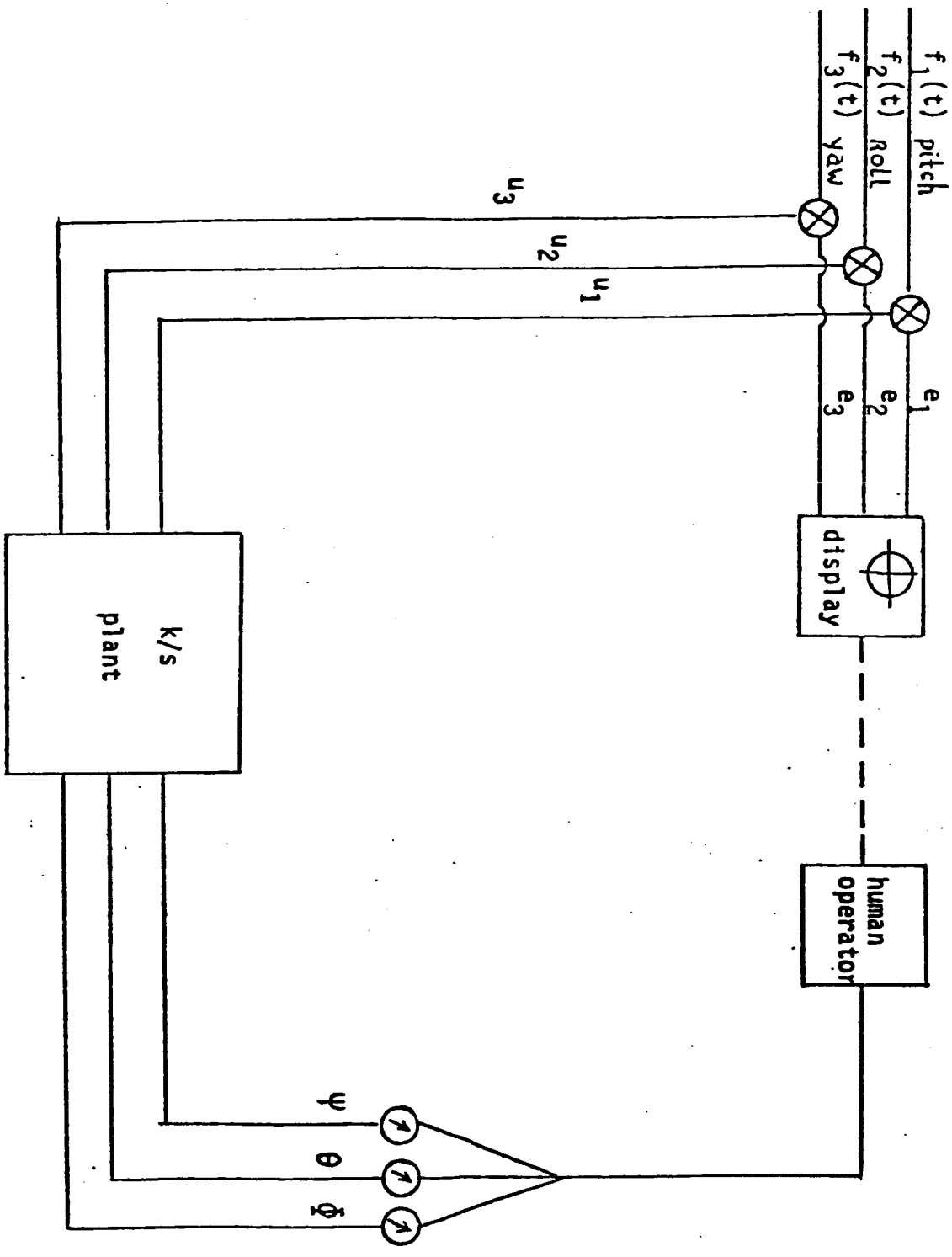


Figure 3. Block diagram illustrating the three independent axes.

sine wave	frequency	amplitude
1	.0785	1.0232
2	.2356	.8212
3	.3926	.4034
4	.6283	.2709
5	.8639	.1518
6	1.0995	.0958
7	1.3351	.0654
8	1.8063	.0591
9	2.4346	.0387
10	3.3771	.0259
11	4.1624	.0139
12	4.7907	.0087
13	5.5716	.0073
14	6.9898	.0068
15	8.4035	.0045
16	9.9742	.0033
17	14.0582	.0033
18	19.7129	.0019
19	28.1949	.0011
20	39.1902	.0008

Table 1. Frequency and amplitude of sine wave terms for the input forcing function.

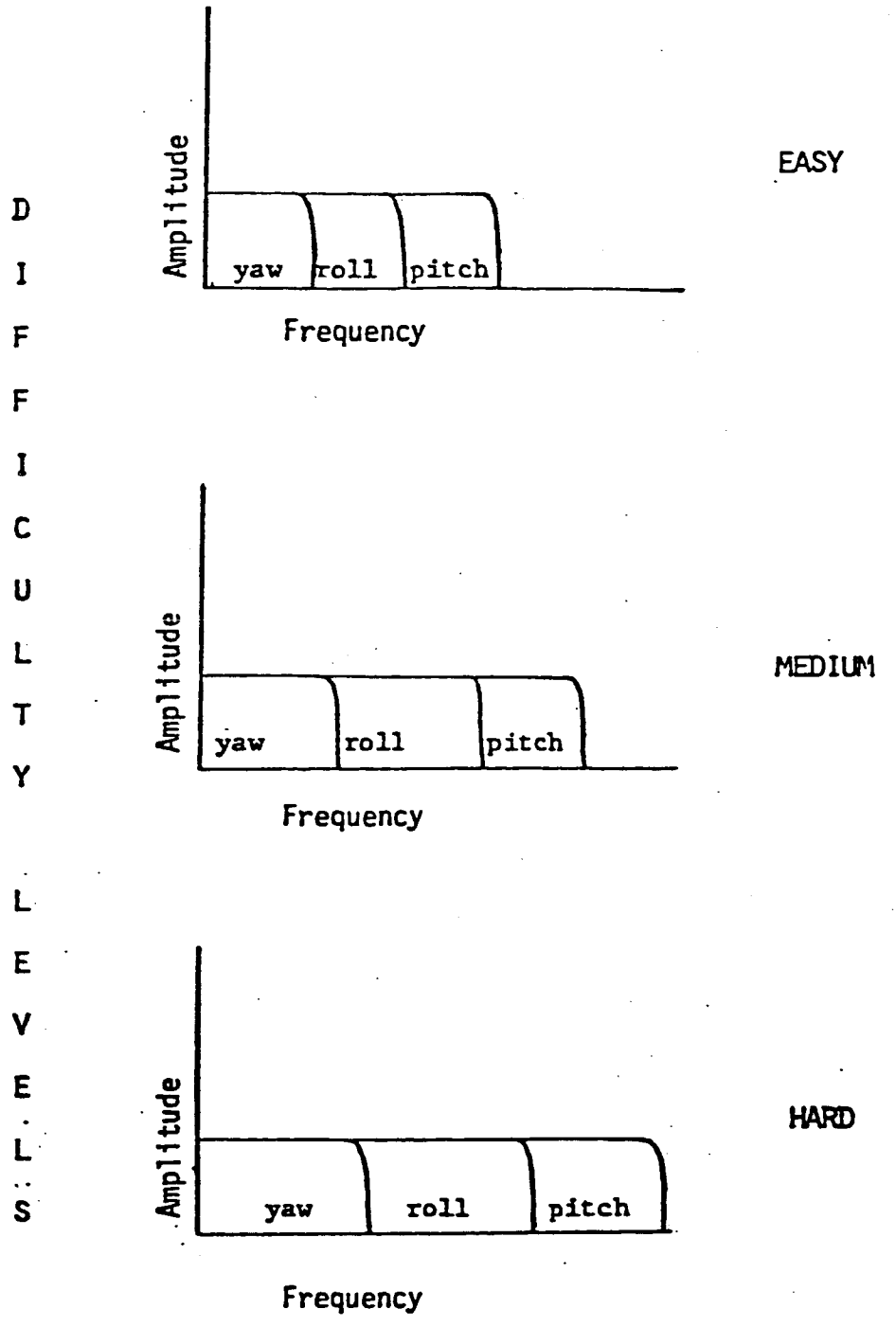


Figure 4. Bandwidth relationships among axes and difficulty levels.

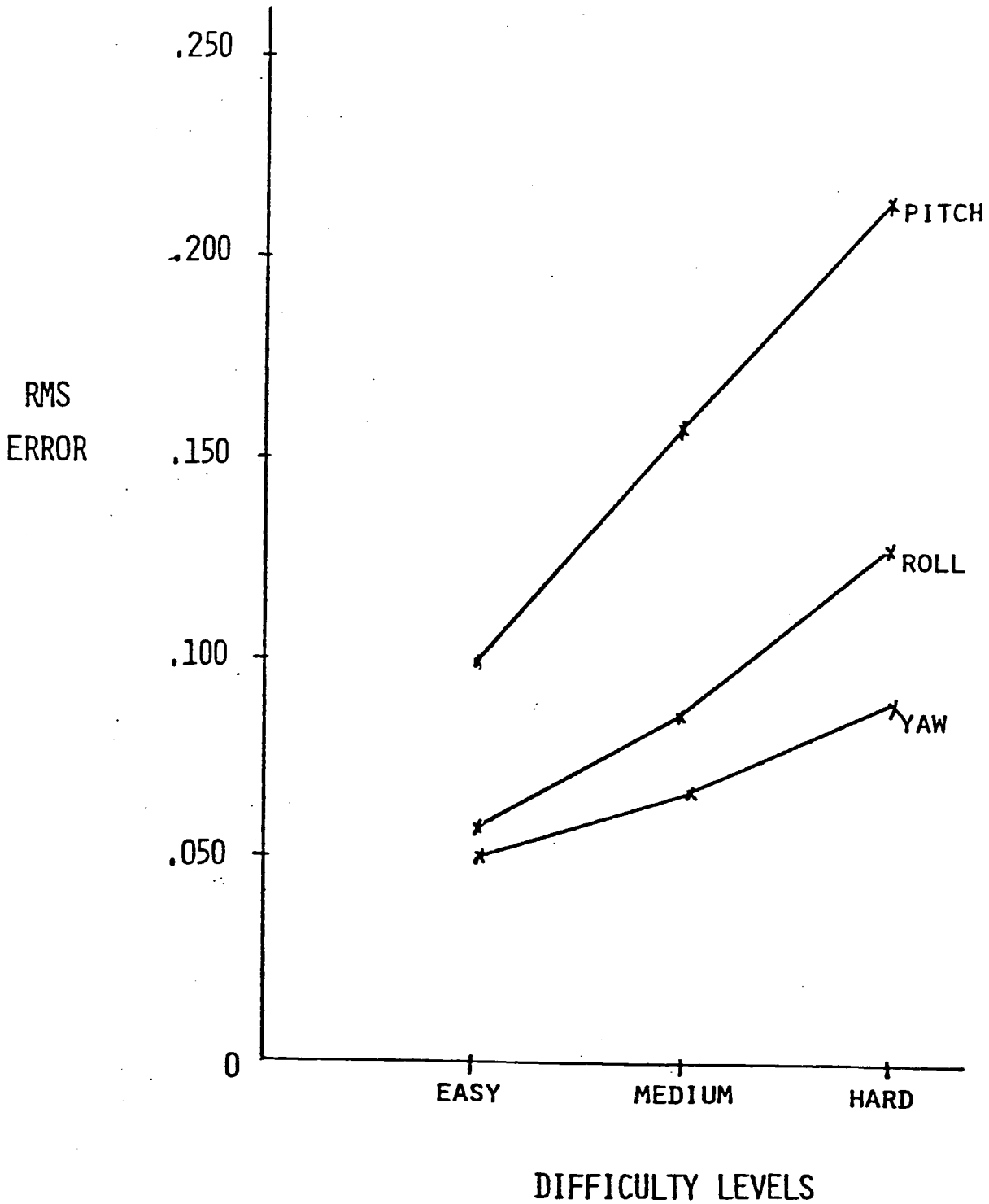


Figure 5. Mean RMS error for the 3 axes over the 3 difficulty levels collapsed across configuration.

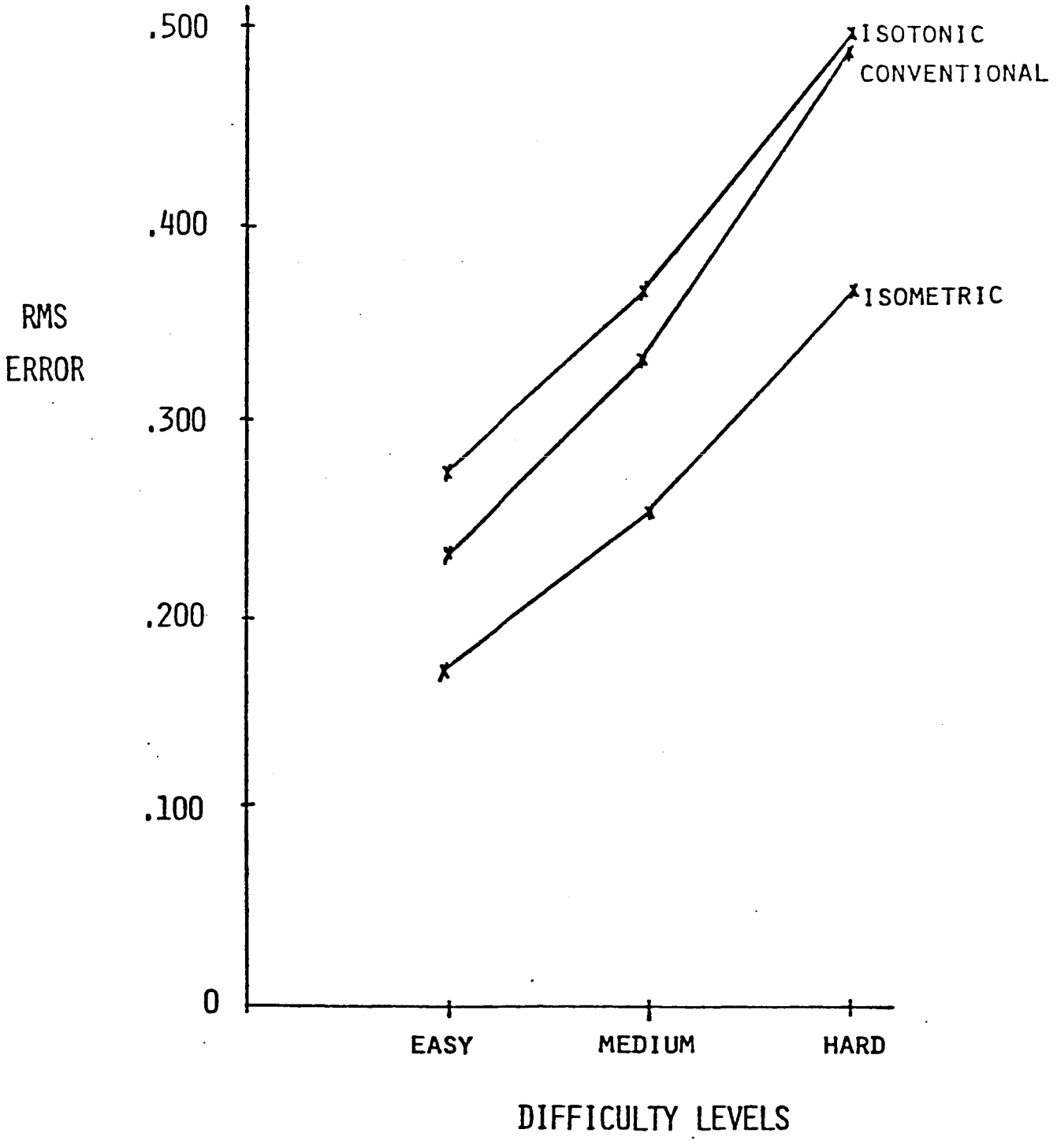


Figure 6. Mean RMS error for the 3 configurations over the 3 difficulty levels.

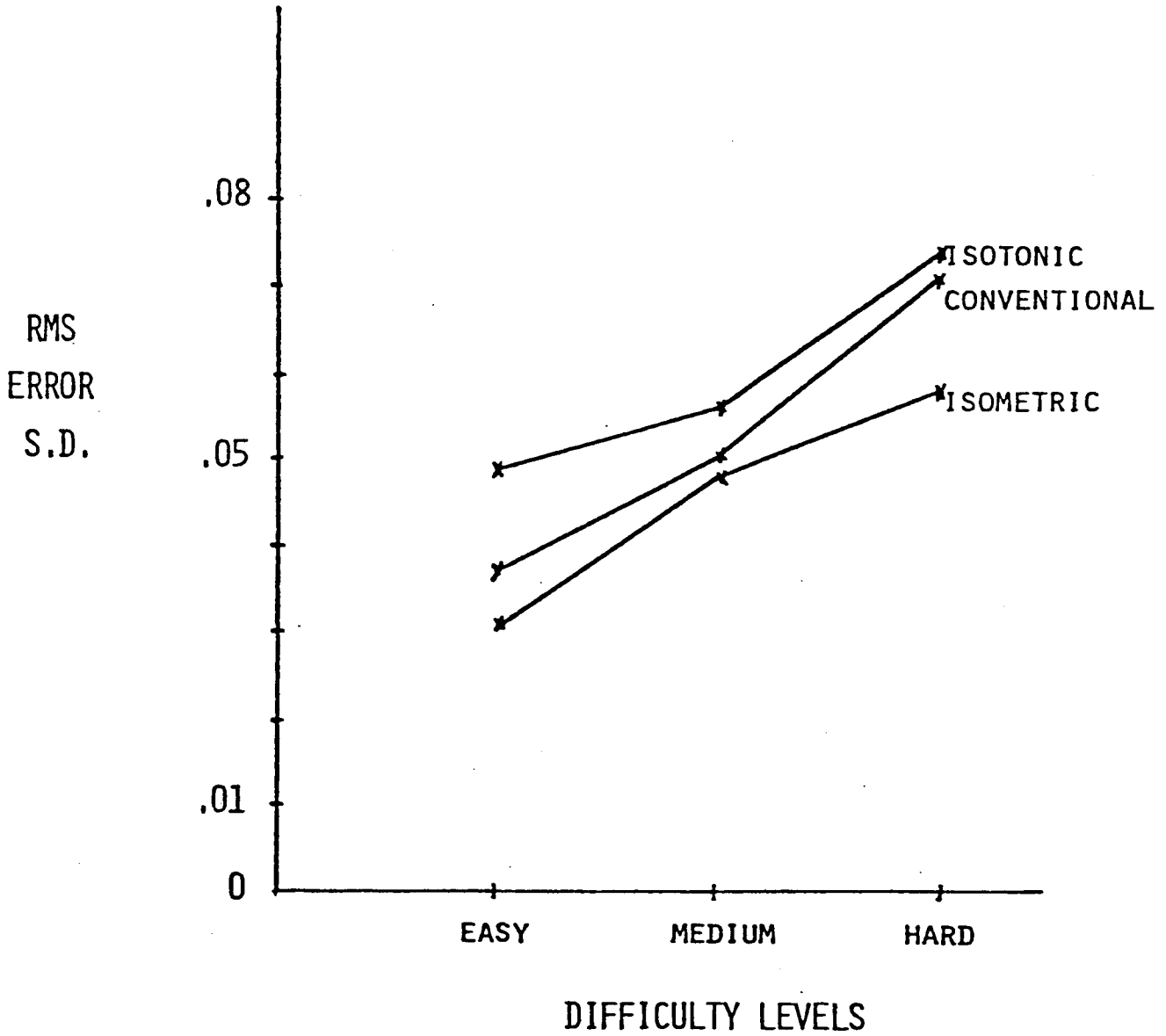


Figure 7. Standard deviation of RMS error for the 3 configurations over difficulty levels.

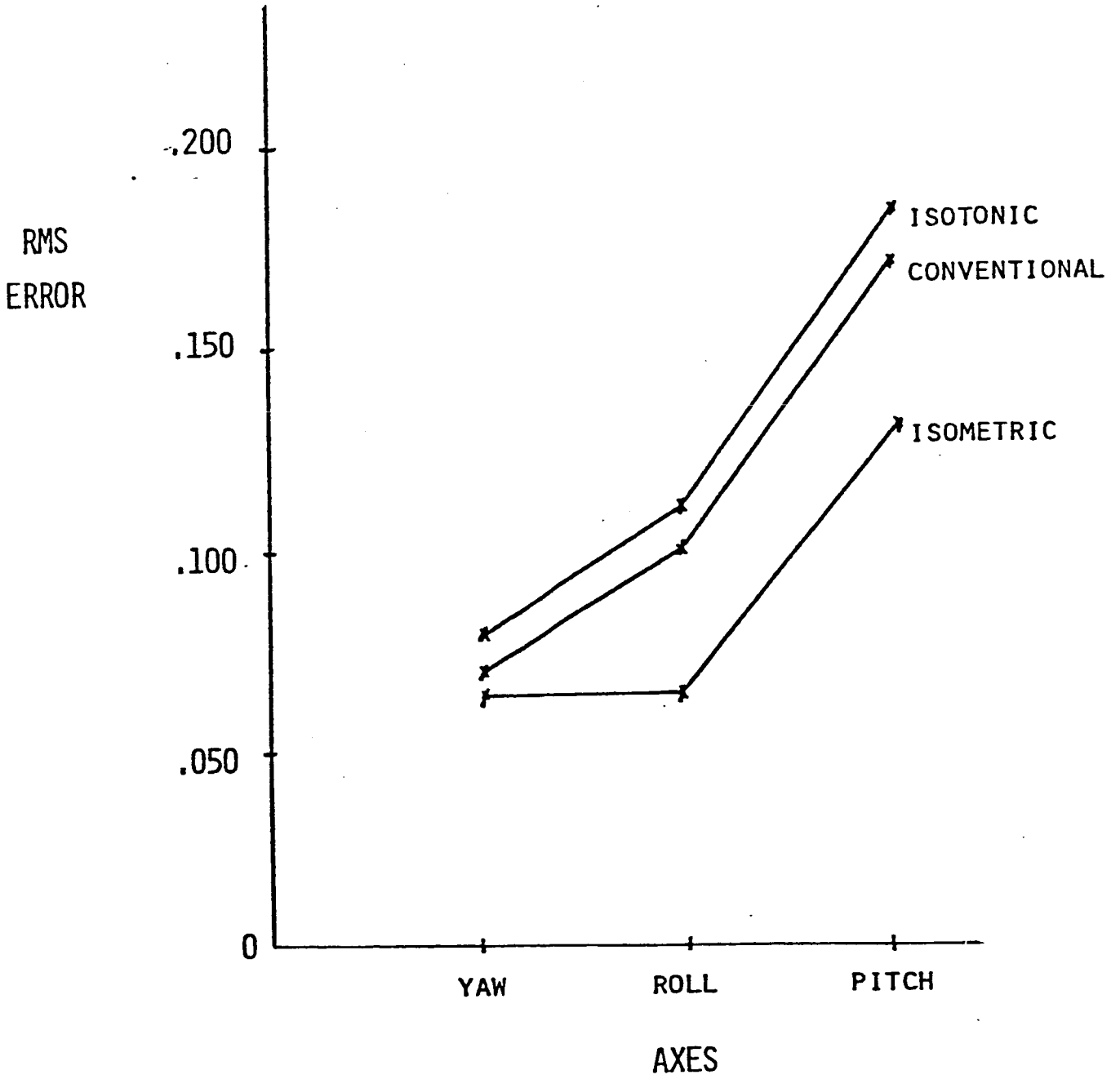


Figure 8. Mean RMS error for the 3 configurations over the 3 axes.

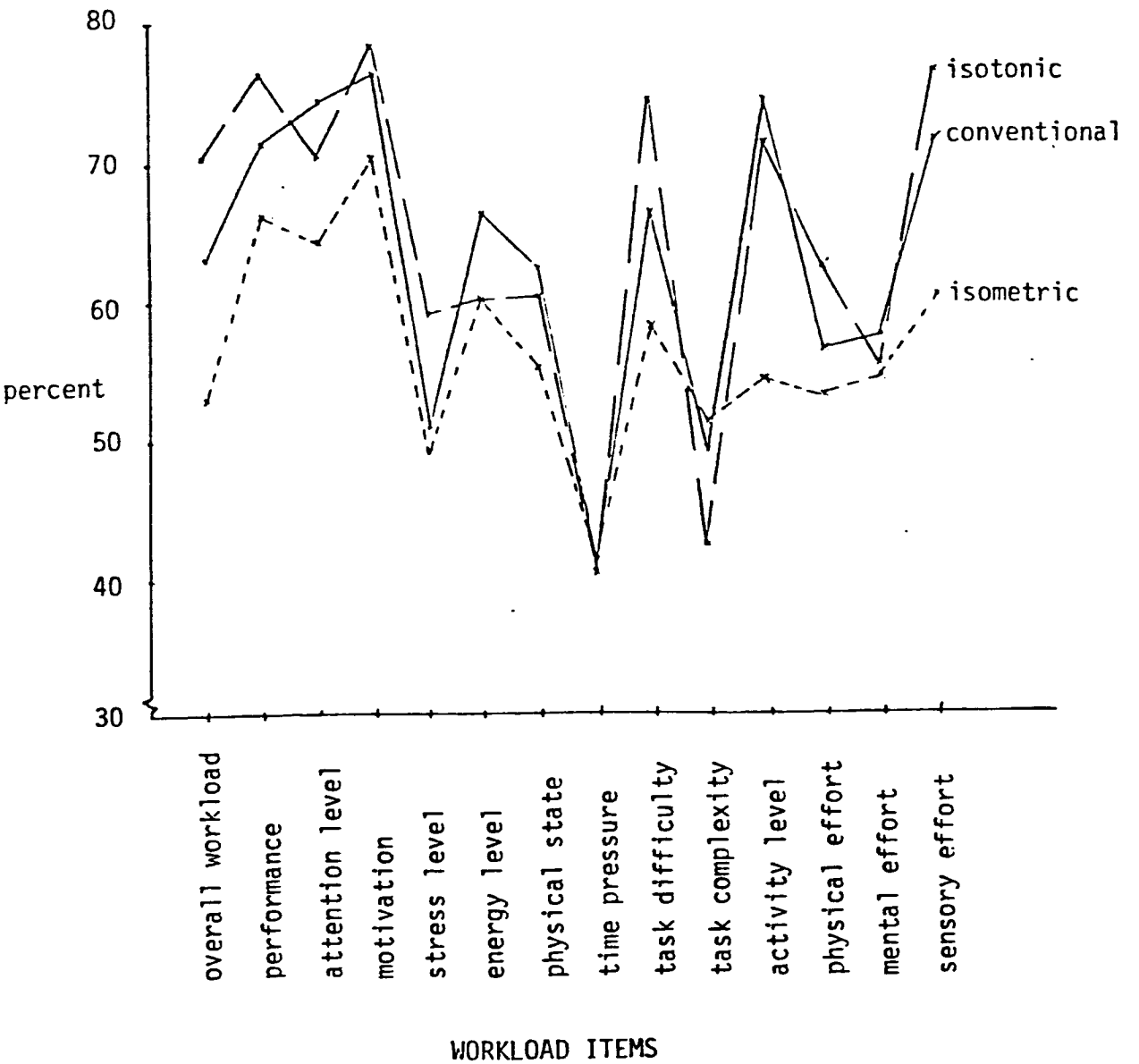


Figure 9: Averaged subjective workload responses for the three configurations.

Helicopter Pilot Response as a Function of Compatibility of the Control-Display Configuration

Kathleen M. Craig
San Jose State University
San Jose, CA 95192

E. James Hartzell
US Army Aeromechanics Lab
NASA/Ames Research Ctr.
Moffett Field, CA

Sherry L. Dunbar
San Jose State University
San Jose CA 95192

Abstract

The present research examined differences in reaction time and task completion measures as a function of control-display configuration in a helicopter cockpit environment. The traditional, contralateral arrangement, in which each control (cyclic/velocity; collective/altitude) is on the side opposite its corresponding instrument, was compared with an ipsilateral, or same side, configuration. The task required simultaneous manipulation of the altitude and velocity controls, and also control of the rudder as a secondary loading task. The mathematical paradigm used to generate levels of task difficulty for comparison and analysis was Fitts' law.

A standard ANOVA showed no significant difference in performance due to control-display configuration. Examination of the data, however, indicated that variations in dual task strategy among subjects may have masked the differences between the two configurations. The Fitts' law relationship held when the controls were looked at individually. A linear combination of difficulty levels was not adequate for the combination of tasks.

Introduction

Flying a helicopter is a demanding visual/motor task requiring a high level of skill and training. In rotocraft, the quality of man-machine interface is critical not only to the stable operation of the vehicle, but also to the safety of the human operator. Special conditions, such as night or adverse weather flying, leave the pilot almost completely dependent upon appropriate and reliable feedback from the aircraft's information display systems. Moreover, in nap-of-the-earth (NOE) flight, which involves flying as close to the earth's surface as possible, below the level of surrounding vegetation and terrain, the helicopter is often only split seconds away from collision with the landscape. These situations demand a man-machine interface that provides a manageable level of pilot workload, and critical to this is the most efficient correspondence between controls and instruments.

Controlling a helicopter requires both hands and feet. The altitude control, or collective, is a floor-mounted left-

hand control, which directs changes in height by an up-down motion. The longitudinal and lateral velocity control, the cyclic, is a center-floor-mounted control-stick, and is manipulated with the right hand by a fore-aft and side-to-side motion. The pilot's feet control the tail rotor by means of rudder pedals, keeping the helicopter on a desired heading during hover and counteracting the effects of torque in forward flight.

In most American-built helicopters the instrument panel is similar to that of its fixed-wing counterpart, with apparent disregard for the special location and function of helicopter controls. The traditional arrangement places the altitude gauge on the right side of the panel, and the airspeed gauge on the left side of the panel; this means that each instrument is contralateral to its corresponding control. This traditional 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 original PNVS, 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).

Stimulus-response compatibility research (Berlucchi, DiStefano, & Tassinari, 1977; Sampson & Elkin, 1965; Simon, 1969; Wickens, Mountford, & Schreiner, 1981) indicates that the present contralateral arrangement of gauges and controls may not be the most efficient configuration for the helicopter and may, in fact, add to pilot workload. A recent study (Hartzell, Dunbar, Beveridge & Cortilla, 1982) found a response difference of 74 msec in favor of an ipsilateral (same side) display arrangement when helicopter cyclic and collective controls were manipulated one at a time, indicating that the "incompatibility" of the contralateral configuration may cost the pilot, at the very least, valuable extra time.

Fitts' law. Fitts' law was the empirical relationship used to generate levels of difficulty for comparison and analysis in the present study. In discrete movement tasks, the law refers to the log linear relationship between the difficulty of a movement and the time required to complete it. The index of difficulty (ID) is defined as:

$$ID = \log_2 2A/W \qquad (1)$$

where A = distance from the home position to the center of the target (amplitude); W = width of target. Fitts' law may be used to predict movement time, that is:

$$MT = a + b \log_2 (2A/W). \qquad (2)$$

Figure 2 illustrates the systemmatic variation of movement time as a function of ID. This linear trend has been

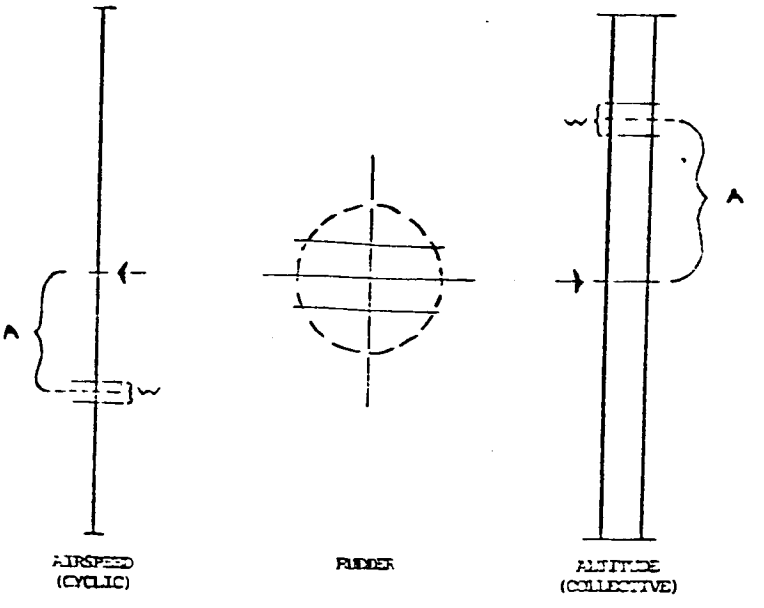


Figure 1. Modified Pilot Night Vision System (PNVS) display, contralateral configuration, showing amplitude (A) and width (W) of sample targets.

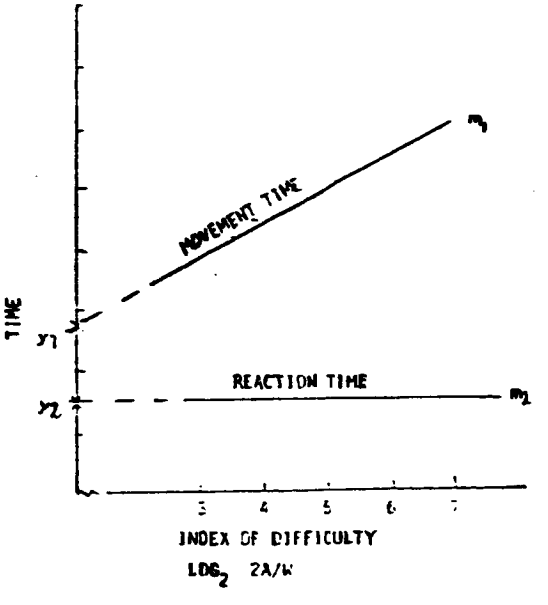


Figure 2. Illustration of Fitts' law relationship between index of difficulty and time. The slopes and y-intercepts of the regression lines are represented respectively by m_1 and y_1 (movement time) and m_2 and y_2 (reaction time). Reaction time is shown to be independent of difficulty level.

shown to hold in zero (Fitts & Peterson, 1964; Klapp, 1975), first (Jagacinski, Hartzell, Ward, & Bishop, 1978), and second (Hartzell et al., 1982) order control systems (position, velocity, and acceleration controls, respectively) for movements over an amplitude range of approximately 35mm to 150mm.

According to the second finding of Fitts and Peterson and others, the effects of ID on the initial reaction to the presentation of the stimulus will be quite small, and slopes of reaction time regression lines by ID will approach zero (Jagacinski et al., 1978). Therefore, the y-intercept as well as the average reaction time may be looked upon as a measure of the central tendency of response latency.

Multi-task performance. Kelso et al. (1979) conducted a study in which two Fitts' law movement tasks were performed simultaneously. The researchers found that subjects waited until they had prepared both movements, and executed them simultaneously. Reaction times and movement times depended, therefore, upon the complexity of the combination of movements, rather than on independent target IDs. Kelso et al. concluded that Fitts' law did not apply to two-handed movements.

The results of the above study suggest that, in dual-movement tasks, a different way of looking at difficulty levels may be necessary. In the present study, IDs for each target have been combined to form an estimate of difficulty for the combination of required movements. One of the purposes of the present research was to test the applicability of the Fitts' law relationship in a multi-task condition.

The hypotheses in the present study were:

(1) Fitts' law is a valid paradigm in a multi-task situation.

(2) In a multi-task design, an ipsilateral control-display configuration will result in significantly better performance than a contralateral configuration.

A helicopter pilot in NOE flying situations will often be required to make immediate, simultaneous changes in altitude, velocity, and heading, or any combinations of these. The task in the present study was designed to approximate the immediate control manipulations that would be required of a pilot in NOE flight.

Demonstration of the superiority of the ipsilateral control-display configuration in the experimental situation would indicate that the existing helicopter instrument arrangement may be neither as compatible nor as efficient as is desirable, and would suggest the need for possible cockpit modifications to improve pilot performance in helicopter maneuvers.

METHOD

Subjects

Sixteen paid subjects, with minimum ages of 18, were employed. Potential candidates, were tested initially on Jex's (1966) "Critical Tracking Task," which is designed to measure psychomotor ability. Eight subjects from each level of tracking ability (moderate and high, according to the Jex task) were randomly assigned to the ipsilateral or contralateral experimental condition. All of the subjects had normal or corrected to normal vision, and none was a helicopter pilot.

Apparatus

The apparatus was similar to that used in the previous Hartzell et al. (1982) experiment. The subject station, as seen in Figure 3, was modeled after the seat, control, and display arrangement of a helicopter. A seat, and cyclic and collective controls were mounted on a platform; in front of the seat was a 54-cm display oscilloscope, which was sloped back from the vertical by four degrees to reduce reflection. The rudder control consisted of foot pedals mounted on the platform. The controls for the cyclic, collective, and rudder were independent of each other. The position of each channel (control) was recorded every 10 msec.

The control dynamics for the cyclic and collective controls, in keeping with actual helicopter dynamics, were second order and first order, respectively. The equations for the dynamics were as follows:

$$\text{CYCLIC} \quad \frac{u}{\delta e} \cong \frac{-22.89\,(s + 1.24)}{[s^2+2(.75)(1.62)s+(1.62)^2](s+.15)^2} \frac{-g}{s} \cdot \frac{\theta}{\delta e}(s) \qquad (3)$$

$$\text{COLLECTIVE} \quad \frac{h}{\delta c} = \frac{6.66}{s(s + 0.3)} \left(\frac{\text{ft/sec}}{\text{in}}\right) \qquad (4)$$

The pilot may be viewed as a regulator whose function is to close two control loops by reducing errors to an acceptable level, as depicted in Figure 4. With the collective, the subject controlled rate of altitude change, and was required to stabilize at a designated altitude (target). Cyclic deflection caused a change in pitch attitude, and a concomitant increase or decrease in airspeed. The cyclic task was twofold in that the operator had to adjust the pitch to initiate a change in airspeed, and then stabilize the system at the designated velocity (target). The display (contralateral configuration) that appeared on the oscilloscope, with pointers and sample targets, is as shown in Figure 1. For the ipsilateral condition, the airspeed and altitude displays were reversed.

The task for the rudder control was a compensatory

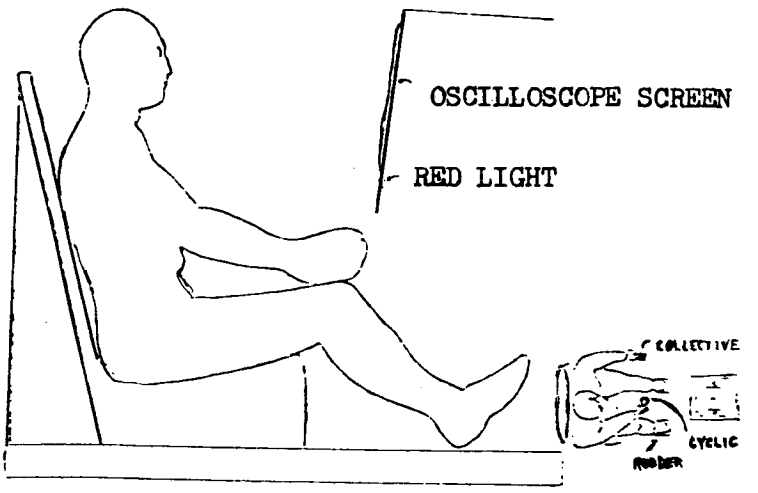


Figure 3. Subject station illustrating placement of controls and display (scale not exact).

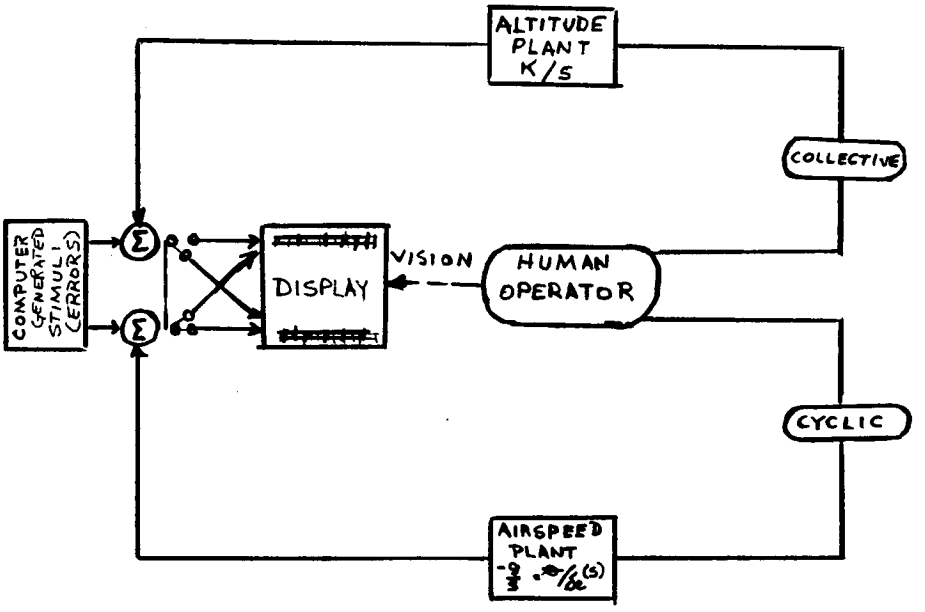


Figure 4. The task required the pilot to act as a regulator, closing two control loops, one for altitude and one for airspeed.

tracking task, involving a dynamic display that was always present throughout the block of trials. The task required the operator to null out the random-appearing error to keep the vertical cursor centered, and was intended to be a secondary loading task.

Target Generation

Targets for each primary task control were factorial combinations of three amplitudes, four target widths, and two directions (up or down from the center). The combinations of amplitudes and widths yielded six different indices of difficulty (calculated by the formula $ID = \log_2 2A/W$ and expressed in bits): 2.5, 3.3, 4.1, 4.9, 5.7, and 6.5. The sequences of target presentation for each of the controls were independent, and the resulting mixtures of altitude and airspeed targets were random. In order to calculate a preliminary estimate of difficulty for each trial as a whole, the IDs of the two controls were combined additively. Total time, or task completion measures for each trial were analyzed according to these combined IDs.

Procedure

Each of the two experimental groups was trained and tested on one control configuration, ipsilateral or contralateral. The secondary task (rudder control) included in both configurations was the same for the two groups. Subjects were seated in the subject station in a darkened room, and communicated with the experimenter through a microphone and speaker system.

Prior to the first trial of each session, the rudder display was activated. Subjects were 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 located at the bottom/center of the oscilloscope was a 2-second warning prior to the beginning of each trial. The primary display (Figure 1) was then activated. The altitude and velocity displays appeared simultaneously at the beginning of each trial, and disappeared following attainment of criterion as described below. If the criteria for both controls were not reached after 30 sec, the trial was terminated.

The subjects' task was 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. When this criterion had been attained for each primary target, and both pointers were within the target area, the trial was over. There was a 5-second interval between trials.

Subjects had two or three experimental sessions of 36 trials per day (6 trials for each index of difficulty), preceded by 10 practice trials. Subjects were given a 5-minute rest period between sessions, and received feedback after each

session to encourage rapid improvement of performance.

Subjects were trained until they were able to maintain asymptotic performance for eight sessions (17-25 hours of training). Asymptotic performance was defined as average variation in performance (reaction time and capture time) of 5% or less. Also, average standard deviations over these sessions could vary no more than 15% for reaction time measures, and 20% for capture times. Data for these 8 sessions or asymptotic performance provided the raw dependent measures.

RESULTS

A 2 x 2 x 8 factorial design (control configuration x control x sessions) with repeated measures on the last two factors was used (Winer, 1971). Four measures were taken on the primary tasks: RTs for cyclic and collective (response latency, from appearance of the targets to 2% deflection of each control stick); MTs for cyclic and collective (movement time, from RT to achievement of capture criterion; CTs for cyclic and collective (capture time, from appearance of the display to achievement of capture criterion); and TT (total time, from appearance of the displays to end of trial).

Mean reaction times, movement times, and capture times for both controls, and total times for each trial, were calculated for each subject's performance by index of difficulty for each of the last eight sessions. A linear regression line was fitted to each subject's data over sessions on each of the four dependent measures.

Regression analyses performed for each subject's reaction time data over sessions by index of difficulty yielded slopes that approached zero, and very low correlation coefficients. This was as expected, since previous findings have indicated that ID will have little effect on response latency.

Regression lines fitted to each subject's data over sessions on movement time for each control had correlation coefficients of .85 or greater (with one exception). Correlation coefficients of regression lines generated by task completion performance (total time) on 11 indices of difficulty were lower than expected (between .66 and .87), based on coefficients for individual controls. No further analysis was performed on total time data.

Given the results of the regression analyses for reaction time, movement time, and capture time, which were in accordance with Fitts' law predictions, the slopes and y-intercepts of the regression lines generated by data on reaction time and capture time as described above were used as dependent measures in four separate 2 x 2 x 8 factorial analyses (configuration x control x sessions).

The analyses that were of primary interest were the ANOVAs on reaction time y-intercepts and capture time slopes.

There were no significant differences in reaction time y-intercepts as a function of configuration. The only significant effect on capture time slopes was for control, $F(1) = 28.98$, $p<.01$. The controls effect was expected, since the plant dynamics of the two control systems were different.

Cell means from these two analyses, as well as overall mean reaction time and mean capture time, were also examined to give a clearer picture of the data. Graphic representations of these means are presented in Figures 5-8.

Figures 5 and 6 show that response latency is somewhat lower for the ipsilateral configuration, in terms of both the y-intercepts and the overall means. As illustrated by Figures 7 and 8, there appears to be an interaction in capture time, both for the slopes and for the overall means.

An examination of the data following the above analyses revealed that comparison between ipsilateral and contralateral times by control was not appropriate, since subjects varied as far as the order of response (i.e., which control was deflected first) and the order of capture. In light of this, several analyses were performed intended to give a better picture of differences in performance as a function of control-display configuration. A 2 x 2 x 8 ANOVA (configuration x order x subjects) was conducted. Order, in this case, refers to the first or second response, regardless of the control involved. A similar analysis was performed using first and second captures as the second factor.

As expected, there was a significant effect for order, $F(1) = 33.61$, $p<.01$ for RT; $F(1) = 66.95$, $p<.01$ for CT. Also, a significant interaction was found in both RT and CT analyses, $F(1) = 4.56$, $p\leq.05$ for RT; $F(1) = 5.00$, $p<.05$ for CT. In other words, the first response and capture were faster for the contralateral configuration; and the second response and capture were faster for the ipsilateral display. Figure 9 contains graphic representations of mean RT1 and RT2, and mean CT1 and CT2.

Given the significant interaction in both the reaction time (RT1 and RT2) and capture time (CT1 and CT2) ANOVAS, a t-test was performed on the difference between mean RT1 and RT2, or inter- response intervals, by configuration. In conjunction with RT1 and RT2, the inter-response interval, was seen as an indication of the manner in which subjects responded to the task as a whole. The difference was significantly larger for the contralateral configuration, t (14) = 2.14, $p < .05$, indicating a difference in response strategy as a function of display arrangement. A t-test was also performed on the intervals between mean CT1 and CT2 by configuration. Again, the interval was significantly larger for the contralateral condition, t (14) = 2.24, $p < .05$. An interpretation of these results will be given in the following section.

Root mean square error measures on the rudder control for

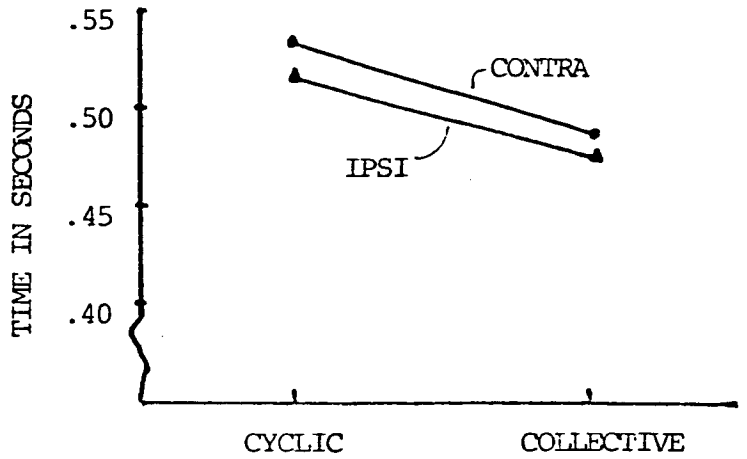


Figure 5. Cell means from analysis of variance on reaction time y-intercepts, reflecting faster RTs for the ipsilateral configuration.

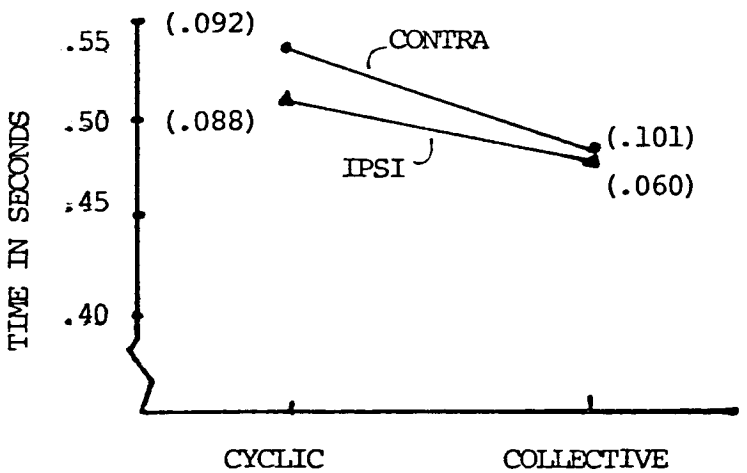


Figure 6. Overall mean reaction times. Note that times are lower for the ipsilateral configuration. Standard deviations are in parentheses.

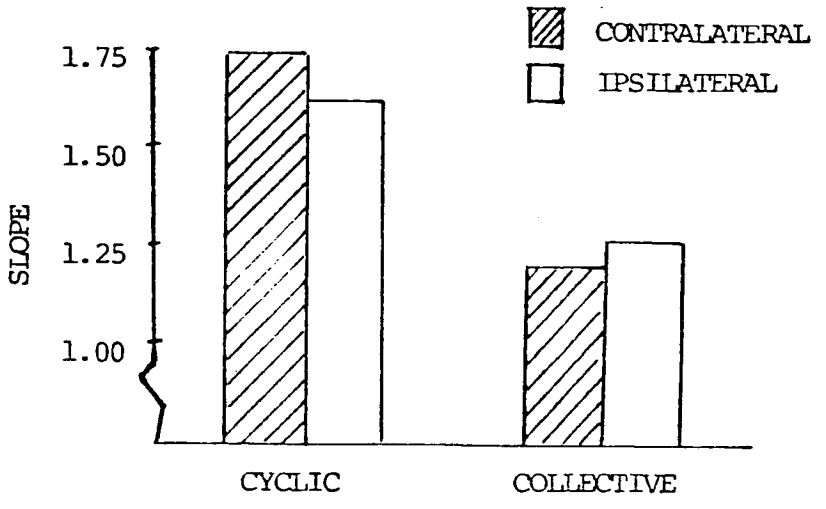


Figure 7. Cell means from analysis of variance on capture time slopes. Note that there is an apparent interaction, and that the difference in control dynamics is reflected in higher slopes for the cyclic control.

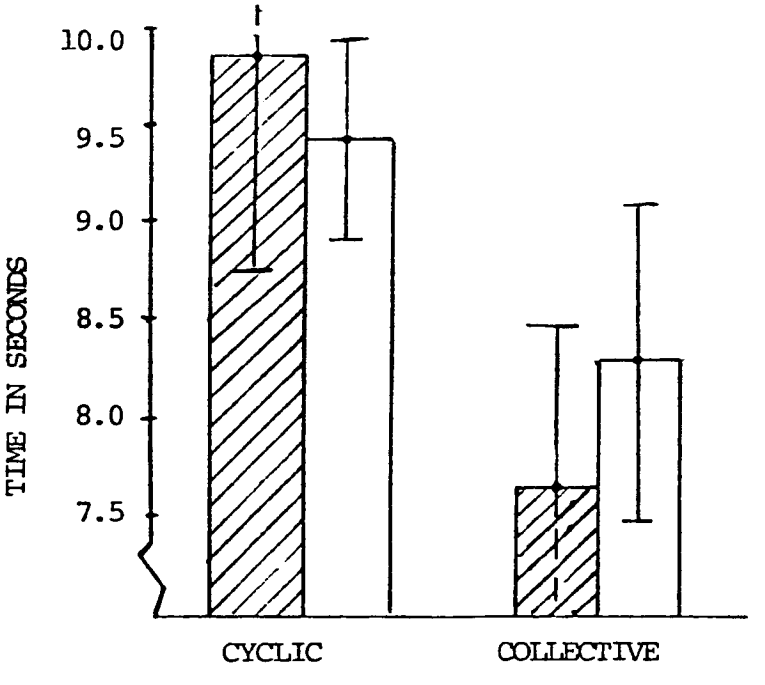


Figure 8. Overall mean capture times. Note that there is an apparent interaction. Standard deviations are represented by $\mathbf{I}$.

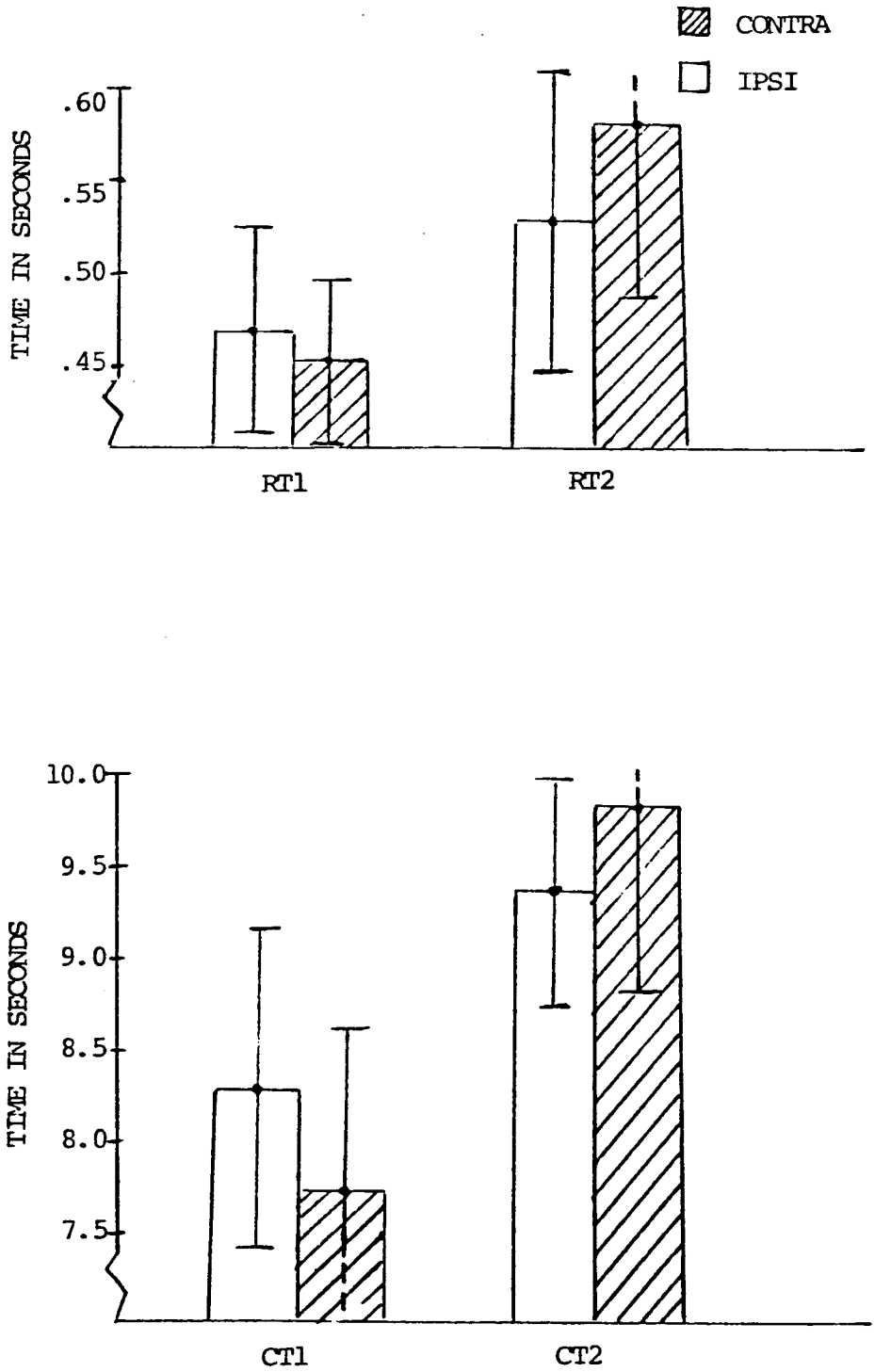


Figure 9. Illustration of first and second response (RT1 and RT2), and first and second capture (CT1 and CT2). Note that there is a significant interaction between configuration and order. Standard deviations are represented by I.

all subjects were consistently low (between 1.0 and 3.0), indicating that sufficient attention was given to this task. No further analysis was performed on the rudder data.

DISCUSSION

Fitts' law. The high correlation coefficients on regression lines generated by movement time performance on individual controls (cyclic and collective), unlike Kelso's findings, demonstrate the applicability of the Fitts' law paradigm in the present multi-task experiment when each task is considered separately.

During the course of the study, it became apparent that the additive combination of IDs for each task was not an appropriate procedure for establishing the IDs for the total trial. The lower correlation coefficients for total time regression lines suggest that the linear combination of Fitts' law IDs does not provide an adequate index of difficulty for the dual-task situation, and that a new relationship must be found.

In addition to differences in plant dynamics between the two control systems, subjective reports indicated that the difficulty of the dual-task trial was affected by several aspects that were unanticipated by the experimenter. Individual subjects reported that trials which should have been comparable in terms of linear combinations of IDs were not perceived as being equally difficult. For example, some subjects commented that combinations with both targets on the same side of center (up or down) were easier to capture than the same ID combinations with targets on opposite sides of center. Also, several subjects commented that any combination that included one wide target was easier than the same ID combination with narrower targets.

It is apparent that the method of estimating dual-task difficulty in the present experiment would have to take into account the dynamics of the control systems involved, the placement of the targets (up or down from center), and the demands of the concurrent tasks as described earlier.

Ipsilateral vs contralateral. The results of the first 2 x 2 x 8 analysis of capture time (configuration x control x sessions) showed a significant effect only for control. Due to the difference in control systems, the CT slopes for the cyclic were expected to be steeper than for the collective. However, the predicted difference between ipsilateral and contralateral capture time slopes by control did not appear. The initial analyses of reaction times showed no significant effects, although, as shown in Figures 5 and 6, mean reaction time y-intercepts and mean response latency with the ipsilateral configuration were lower for both controls.

The second 2 x 2 x 8 analyses, in which order of response or capture was taken into consideration, proved to be more informative and useful in interpreting the data. The significant interaction between condition and order was of particular interest because of the direction it takes. Although contralateral subjects were faster with the first movement and capture (RT1 and CT1), the ipsilateral subjects completed the second response and capture (RT2 and CT2) more quickly (Figure 9).

The interaction described above seems to have masked any main effects of display configuration. A re-examination of strategies used by subjects shed more light on the differences between the two experimental groups. One of the ways of discovering response and capture strategies was to compare the intervals between the first and second reactions, and between the first and second captures. As detailed previously, these intervals were significantly longer for the contralateral condition. The control- display arrangement, then, seemed to have an effect on the method of response and task completion.

The longer inter-response intervals for the contralateral arrangement indicate that the majority of contralateral subjects used the alternating response strategy (Damos, 1981). The strategy seemed to extend into the movement phase of the task, as the contralateral subjects also had longer intervals between capture of the two targets. The difference in these intervals for the two configurations could reflect a difference in task processing that is directly related to the type of configuration.

In the present study, subjects having the ipsilateral control- display configuration responded in a way consistent with results found by Kelso et al. (1979). Following Kelso's interpretation, the ipsilateral subjects waited until they had processed both stimuli (targets), and then responded to the task as a whole, following the first control response almost immediately with a movement of the other control.

The contralateral configuration, because of the incompatibility of the control-display arrangement, appears to make it difficult for the subject to develop a strategy for both controls; instead, subjects seem to look at and react to one target at a time. As a result, contralateral RT1 was faster than ipsilateral RT1, since it takes less time to process one stimulus than two varying stimuli (Kelso et al., 1979).

Once the operator has made the first response, however, he/she must then shift to the other stimulus (target), process it, and respond. Thus, the second reaction times (RT2) for the contralateral configuration are slower than the ipsilateral times because another complete processing phase must take place. This explains the larger inter-response interval with the contralateral configuration. When responses in both controls are called for, the strategy used with the contralateral condition is clearly not the fastest way to effect the desired

changes. The same argument may be made regarding capture times.

The contralateral configuration, then, is more likely to be seen in terms of two separate tasks. Because of this, the operator is more likely to treat them as such and perform the tasks in a disjointed, back-and-forth fashion. The effect of the display complexity will accumulate, since the operator will be faced with another incompatible S-R correspondence with every required change.

The question raised in the present study, as to whether the traditional contralateral control-display configuration is the most efficient for the helicopter, depends on whether a configuration that encourages a segmented processing and movement strategy is better than one that elicits a more integrated natural response. The data in the present research indicate that the traditional configuration may actually cause a detriment in performance when the pilot is called upon to make coordinated control inputs.

If the argument holds that displayed feedback in a contralateral configuration forces the operator to shift attention from one element to the other, without being able to process and effect changes simultaneously, then this is clearly not the most efficient arrangement. The initial lag between responses would extend into further control manipulations, causing a more serious detriment in performance as required changes became more complex. The final conclusions from the results of this study must therefore be that modifications of the traditional helicopter instrument/control configuration are indicated.

REFERENCES

Berlucchi, G., Crea, F., DiStefano, M., & Tassinari, G. Influence of spatial stimulus-response compatibility on reaction time of ipsilateral and contralateral hand to lateralized stimuli. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 505-517.

Damos, D. Individual differences in multi-task response. Proceedings of the Human Factors Society, 1981, 291-295.

Fitts, P. M., & Peterson, J. R. Information capacity of discrete motor responses. Journal of Experimental Psychology, 1964, 67, 103-112.

Hartzell, E. J., Dunbar, S., Beveridge, R., & Cortilla, R. Helicopter pilot response latency as a function of the spatial arrangement of instruments and controls. Proceedings of the Eighteenth Annual Conference on Manual Control, Dayton, Ohio, June, 1982, in press.

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., Mc Donnell, J. D., & Phatak, A. V. A "critical" tracking task for manual control research. IEEE Transactions on Human Factors in Electronics, 1966, 7, 138-145.

Kelso, J. A., Southard, D. L., & Goodman, D. On the nature of human interlimb coordination. Science, 1979, 203, 1029-1031.

Klapp, S. T. Feedback versus motor programming in the control of aimed movements. Journal of Experimental Psychology: Human Perception and Performance, 1975, 104, 147-153.

Sampson, P. B., & Elkin, E. H. Level of display integration in compensatory tracking. Perceptual and Motor Skills, 1965, 20, 59-62.

Simon, R. J. Reactions toward the source of stimulation. Journal of Experimental Psychology, 1969, 81, 174-176.

Wickens, C. D., Mountford, S. J., & Schreiner, W. Multiple resources, task-hemispheric integrity, and individual differences in time-sharing. Human Factors, 1981, 23, 211-229.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill Book Co., 1971.

Use of the AFTI/F-16 Twist Throttle as a Decoupled Controller

Kenneth E. Sanders
and
Mike E. Cope

General Dynamics Corporation
Fort Worth, Texas

Presented at the 19th Annual Conference on Manual Control
Massachusetts Institute of Technology
May 23-25, 1983

USE OF THE AFTI/F-16 TWIST THROTTLE AS A DECOUPLED CONTROLLER

ABSTRACT

Manual control of the longitudinal decoupled motions of a highly augmented aircraft such as the AFTI/F-16 required the development of a new controller. Control device motion for these uncoupled aircraft motions had to be compatible with the resulting aircraft motion and the new controller would ideally be located on either the throttle or the side-stick controller to meet the requirement of hands-on combat control. This paper dicusses in detail the criteria, rationale, and testing which produced the AFTI/F-16 twist throttle configuration for decoupled control application.

The development phases of the twist throttle are grouped into several categories. These are:

(1) Review of the F-16 fighter CCV program method for controlling decoupled motions and developing alternate methods.

(2) Studying these methods using conceptual mock-ups.

(3) Selection and testing the best configuration.

(4) Flight test.

Results of testing the twist throttle in centrifuge tests and the actual use of the throttle in a fixed base simulator are also discussed in detail.

INTRODUCTION

The AFTI/F-16 program is jointly sponsored by the Air Force, NASA, and the Navy, conducted by General Dynamics and is aimed at development and integration of emerging aircraft technologies to improve fighter mission effectiveness. The overall objective of the AFTI/F-16 program is to establish the benefits of integrating direct force and weapon line pointing, digital flight control, integrated flight and fire control, and advancements in pilot/vehicle interface in terms of fighter weapon system effectiveness and survivability. The AFTI/F-16 is a testbed aircraft for flight demonstration of these integrated technologies as well as for development of individual technologies is this and future programs. The current program definition is broken into two phases: development of a Digital Flight Control System (DFCS) and development of an Automated Maneuvering Attack System (AMAS). The DFCS phase objectives are to validate the practicality and benefits of a triplex digital multimode flight control system and supporting flight management displays for task tailored weapon system operation. This program phase will provide a proven aircraft system with flexibility to incorporate the AMAS modifications in the second phase to assess weapon delivery and survivability improvements and pilot acceptance of automated systems. The AFTI/F-16 is constructed from a Full Scale Development F-16 aircraft. Principal airframe structural modifications were the addition of twin, all moveable vertical canards to a modified engine

inlet structure, and addition of a dorsal fairing to house flight test instrumentation and other equipment. Figure 1 illustrates the structural configuration of the AFTI/F-16.

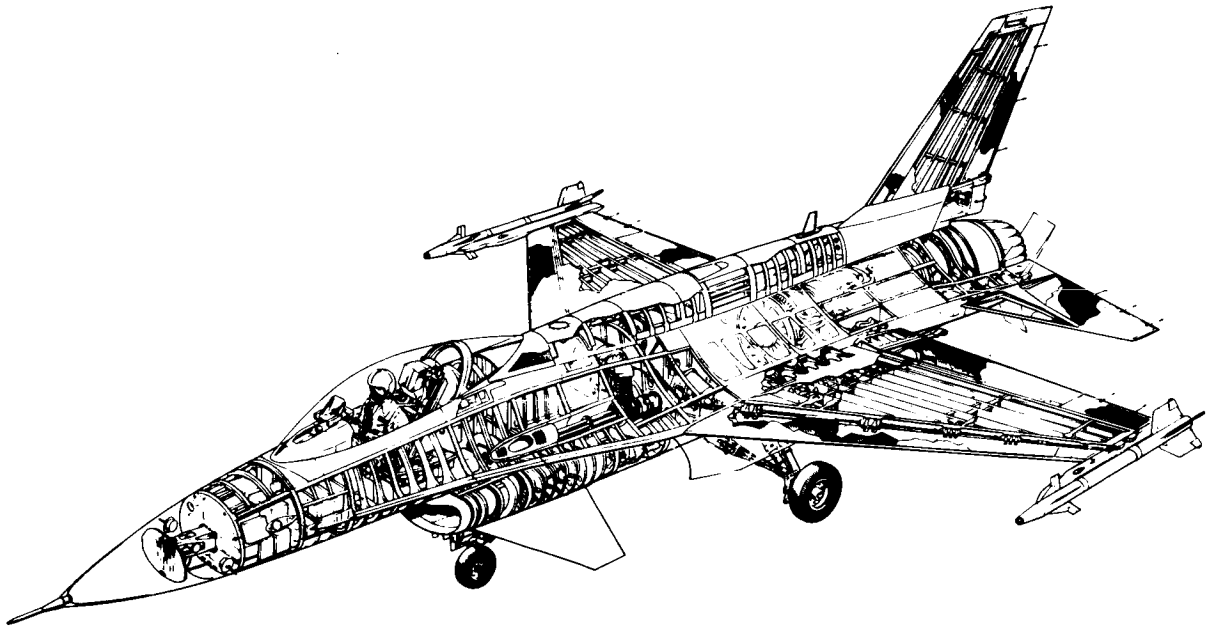
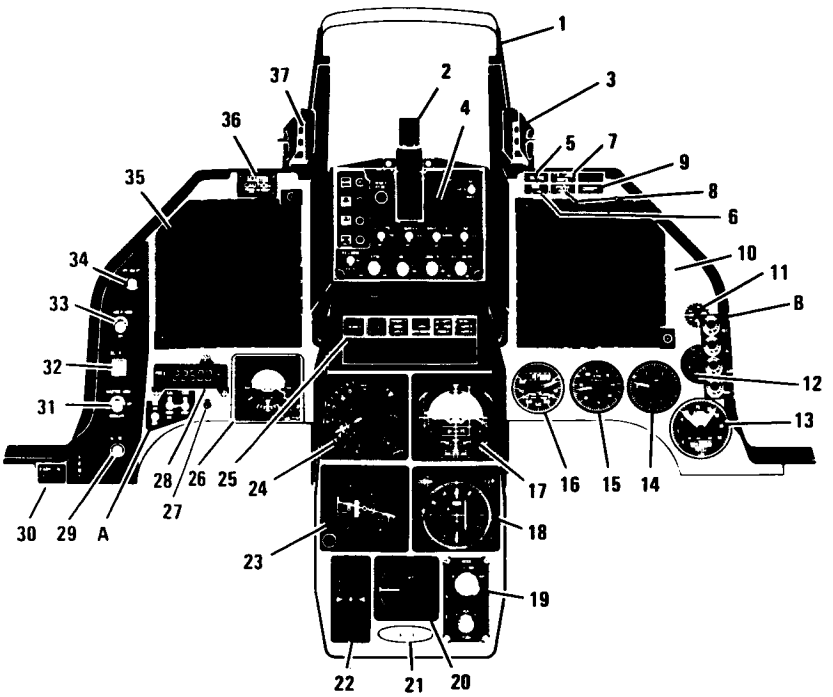


Figure 1. - AFTI/F-16 Structural Configuration

Significant revisions in the F-16A pilot display systems were made in order to host the DFCS multimode functions and to explore weapon delivery improvements through new concepts in pilot-vehicle interface. The revised displays are shown in Figure 2. Of particular interest are the multipurpose cathode ray displays which dominate the panels to the right and left of center in the figure. The CRT display generators can be programmed to provide virtually any type of alpha numeric or pictoral information and the displays give the pilot a new ability to



INSTRUMENTATION
A. Remote Instrumentation Control Panel
B. Instrumentation Status Lights

1. HUD Combiner Glass
2. Television Sensor
3. Air Refuel Status/NWS Indicator
4. HUD Control Panel
5. Eng Fire Warning Light (Red)
6. Dual F. C. Fail Warning Light (Red)
7. T.O. Land Config Warning Light (Red)
8. Hyd/Oil Press Warning Light (Red)
9. Canopy Warning Light (Red)
10. Righthand Multi-Purpose Display
11. Engine Oil Pressure Indicator
12. Nozzle Position Indicator
13. Fuel Quantity Indicator
14. FTIT Indicator
15. Tachometer
16. Fuel Flow Indicator
17. Attitude Director Indicator (ADI)
18. Horizontal Situation Indicator (HSI)
19. Instrument Mode Select Panel
20. Vertical Velocity Indicator
21. Rudder Pedal Adj Knob
22. Angle-of-Attack (AOA) Indicator
23. Altimeter
24. Air Speed/Mach Indicator
25. Threat Warning Indicator Control Unit
26. Standby Attitude Indicator
27. Marker Beacon
28. UHF Channel/Frequency Indicator
29. Alternate Release (Alt Rel) Button
30. Radio Call
31. Master Arm Switch
32. Sel Jett Button
33. Laser Arm Switch
34. IFF Ident Button
35. Lefthand Multi-Purpose Display
36. Master Caution Light (Amber)
37. Angle-of-Attack (AOA) Indexer

Figure 2. - AFTI/F-16 Displays

recognize and use data from the onboard equipment and the external sensors as well as a convenient capability to interrogate and configure the flight control system. The full cockpit modifications include new pilot's instrument panel, glare shield, and center pedestal used for displaying a new, improved field-of-view- head-up display (HUD) and the two interactive redundant multipurpose displays. The left and right console structures remained unchanged, but the panel arrangement within the consoles was modified for a more optium pilot-vehicle interface. These features are shown in Figure 3.

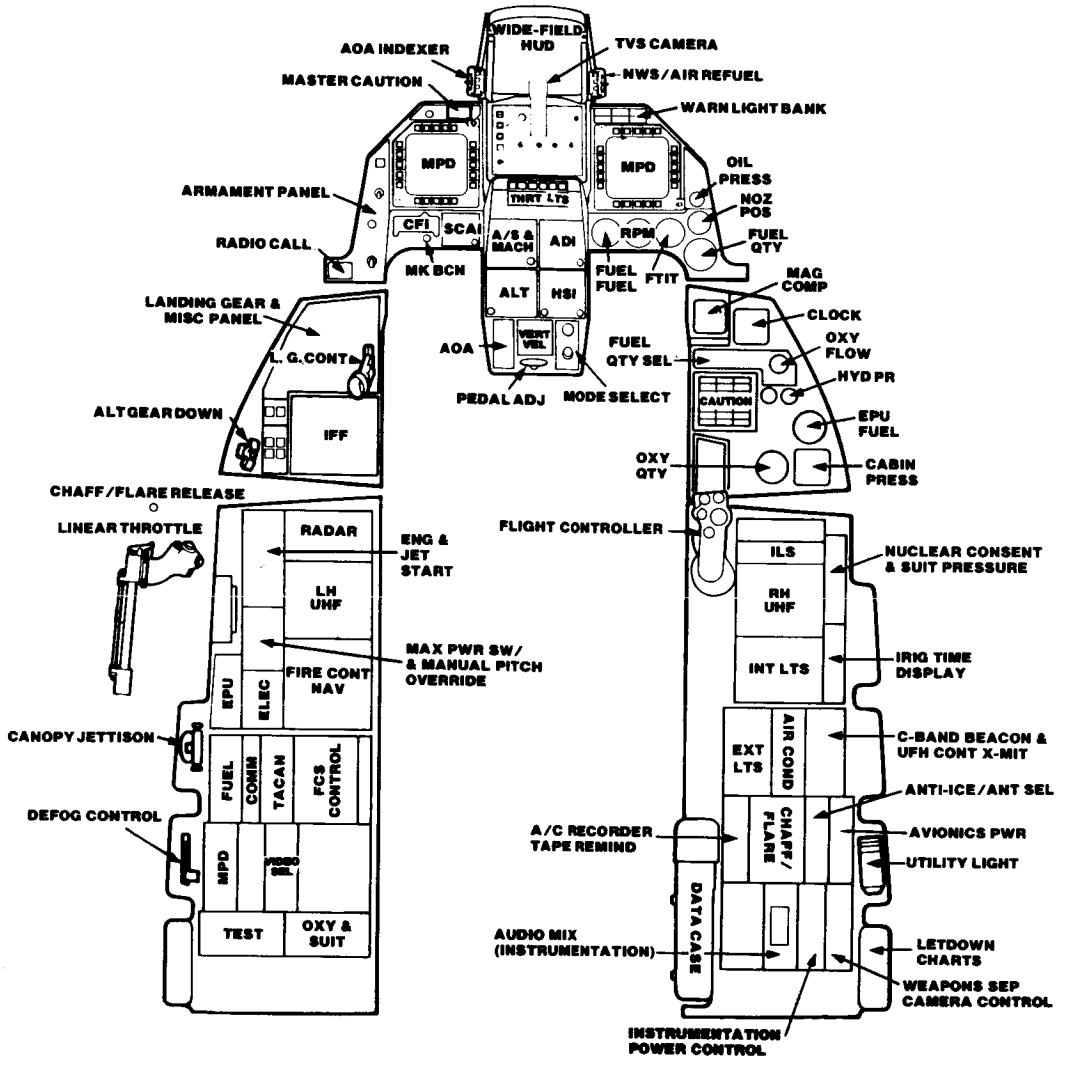


Figure 3. - Crew Station Arrangement

OVERVIEW OF MULTIMODE TASK TAILORED CONTROL LAW ARCHITECTURE

The concept of task-tailored integrated control systems has evolved over many years. However, practical application awaited the advent of digital flight control computers. An advanced development program utilizing the #1 YF-16 initiated in December 1973 (Reference 4), called the fighter Control Configured Vehicle (CCV) program. The project's objectives were to analyze, design, implement, and carry into flight test CCV concepts that had been identified from previous developmental programs as showing a potential for increasing a fighter's effectiveness. The aircraft was modified to carry an additional analog computer and twin vertical canards. Direct force capability was derived from simultaneous use of vertical canards and rudder or

symmetric flaperons and elevator. Seven unconventional modes were implemented with this capability allowing flat turns, decoupled normal acceleration control, independent longitudinal and lateral translations, uncoupled elevation and azimuth aiming and blended direct lift. The need for the task-oriented control laws and the inadequacy of the current design criteria were clearly indicated. It was also apparent that a blended technigue, mixing conventional and decoupled control, offered significant advantages. Both an analytic study and a piloted investigation on a large amplitude moving base simulator were conducted following the CCV flight test program to develop task-oriented control concepts employing the decoupled control features. Concepts such as blended conventional and decoupled control, weapon line stabilization and gust alleviation and rate command of independent fuselage pointing were found to be worthy of further evaluation. A variation of the multimode approach was indicated in which the pilot could be provided with various pre-determined combinations of conventional and decoupled control for each mission phase. The AFTI/F-16 program is the culmination of these areas of research. The AFTI/F-16 flight control laws are implemented with the following four major control modes: (1) Normal Mode (NRM), (2) Air-to-Air Gunnery (AAG) Mode, (3) Air-to-Surface Gunnery (ASG) Mode, and (4) Air-to-Surface Bombing (ASB) Mode. Each of these mode types is implemented for two types of operation: (1) standard (coupled) control and (2) decoupled control. Each of these modes is task-tailored to provide optimum flying qualities in the flight phase for which the modes are designed. General provisions that are implemented in each mode are as follows:

Normal Mode - The normal mode is used throughout the applicable flight envelope for takeoff, cruise, and landing and for the performance of secondary mission tasks, such as air refueling and formation flying. The mode is designed to provide a smooth ride, load factor gust alleviation and reduction of pilot workload during secondary mission segments.

Air-to-Air Gunnery Mode - The air-to-air mode is used throughout the air combat flight envelope to provide rapid maneuvering during target intercept and precise tracking. Control law design strategy for this mode is primarily based on optimizing tracking characteristics.

Air-to-Surface Gunnery Mode - The air-to-surface gunnery mode provides rapid and precise pointing for increased accuracy and survivability when ground targets are being strafed.

Bombing Mode - The bombing mode provides precise control of the aircraft velocity vector and improves gust alleviation responses in flight path to facilitate bombing accuracy and to enable the employment of effective control strategies in order to increase aircraft survivability.

Figure 4 illustrates the top level architecture of the multimodes that reside in the control system Operational Flight Program (OFP).

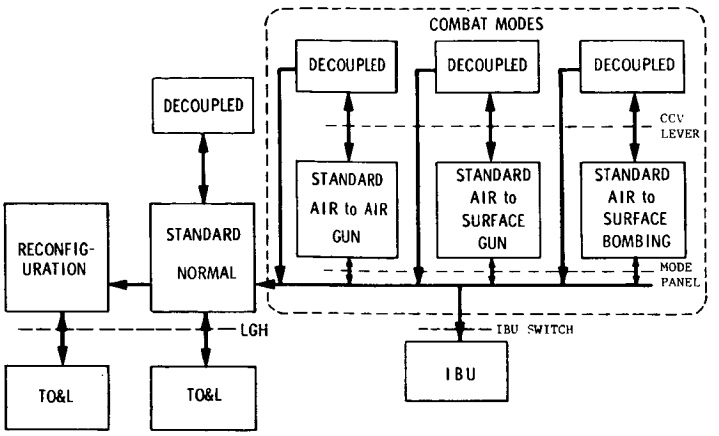


Figure 4. - AFTI/F-16 Control Law Multimode Structure

The primary mode in this architecture is the Standard Normal Mode. A special configuration of this mode is engaged when the landing gear handle (LGH) is lowered for optimal takeoff and landing characteristics. Also, reconfiguration of the control laws for dual control system sensor failures is automatically accomplished by the control system by modifying the Standard Normal Mode to operate without the failed sensor. The combat modes can all be engaged by depressing the proper mission phase button on the control panel of the HUD. Decoupled modes are engaged by first selecting the appropriate standard mode, and then depressing the CCV lever (paddle switch) on the sidestick. A special analog flight control system is mechanized in each flight control computer for use in the event that a generic software fault should cause failure of the digital control system. This mode is an Independent Back-Up (IBU) which is automatically engaged in the event that the digital system is unable to resolve indicated failures. It can also be manually engaged by the pilot by depressing an IBU button which is recessed in the top portion of the sidestick. The longitudinal control law command philosophy for optimized flight characteristics is illustrated in Figure 5. Preselected decoupled mode features shown in the figure are implemented with task utilization in mind.

CONTROLLER	MULITMODE			
	STD NORMAL	*STD BOMBING*	*STD ASG*	*STD AAG*
SIDE STICK (Pitch)	A_N COMMAND	A_N COMMAND	Q COMMAND	Q COMMAND
SIDE STICK (Roll)	ROLL RATE COM	ROLL RATE COM	ROLL RATE COM	ROLL RATE COM
RUDDER PEDAL	RUDDER DEFLECTION	FLAT TURN	FLAT TURN	FLAT TURN
THROTTLE (Twist)	NONE	NONE	NONE	NONE
	DECOUPLED	*DECOUPLED*	*DECOUPLED*	*DECOUPLED*
SIDE STICK (Pitch)	FPME*	FPME	PRME*	PRME
SIDE STICK (Roll)	ROLL RATE COM	ROLL RATE COM	ROLL RATE COM	ROLL RATE COM
RUDDER PEDAL	TRANSLATION	FLAT TURN	POINTING	POINTING
THROTTLE (Twist)	TRANSLATION	DIRECT LIFT	POINTING	POINTING

*FPME – FLIGHT PATH MANEUVER ENHANCEMENT
PRME – PITCH RATE MANEUVER ENHANCEMENT

Figure 5. - AFTI/F-16 Multimode Flight Controller Commands

Standard Normal Mode and Standard Bombing Mode are load factor "g" common systems and the other modes are pitch rate command systems in that steady state load factor errors are nulled or pitch rate errors are nulled by the feedback control law. The coupled response characteristics are different for each mode and are tailored to the task requirements of each mission phase.

Decoupled control of the AFTI/F-16 is achieved by independent control of normal force (normal acceleration of normal flight path) and aircraft pitch rate. By properly controlling the redundant surfaces in the longitudinal axis, a variety of decoupled aircraft motions are obtainable. Figure 6 defines the basic longitudinal decoupled responses in nomenclature which is currently accepted in the aircraft community. The AFTI/F-16 utilizes each of these motions in task-tailored combination in the various control modes.

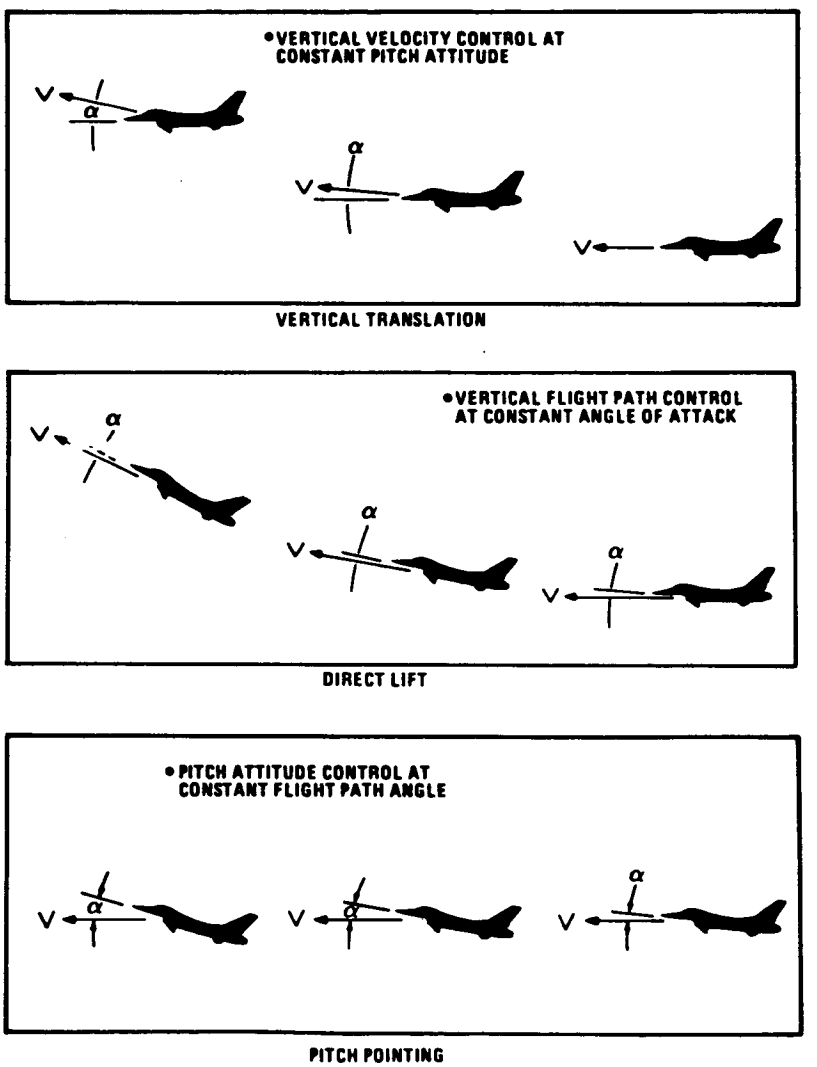


Figure 6. - Longitudinal Decoupled Motion

DEVELOPMENT OF A LONGITUDINAL DECOUPLED CONTROLLER

Prior to this program, most of the research and development efforts on 6 DOF aircraft motions have emphasized the control laws and flight control hardware necessary to obtain the various pure decoupled modes of response and not how the pilot is to manually control advanced 6 DOF features effectively in an operational environment. A brief preliminary discussion of the history of decoupled controllers at General Dynamics is in order, followed by a description of how a valid decoupled longitudinal controller was developed for the AFTI/F-16.

During the early phases of the AFTI/F-16 program, the control method for the new decoupled flight modes was studied. This was initiated with a review of the Fighter CCV program, in which a minature two axis force controller was installed on top of the F-16 sidestick controller for commanding decoupled modes. This arrangement is shown in Figure 7.

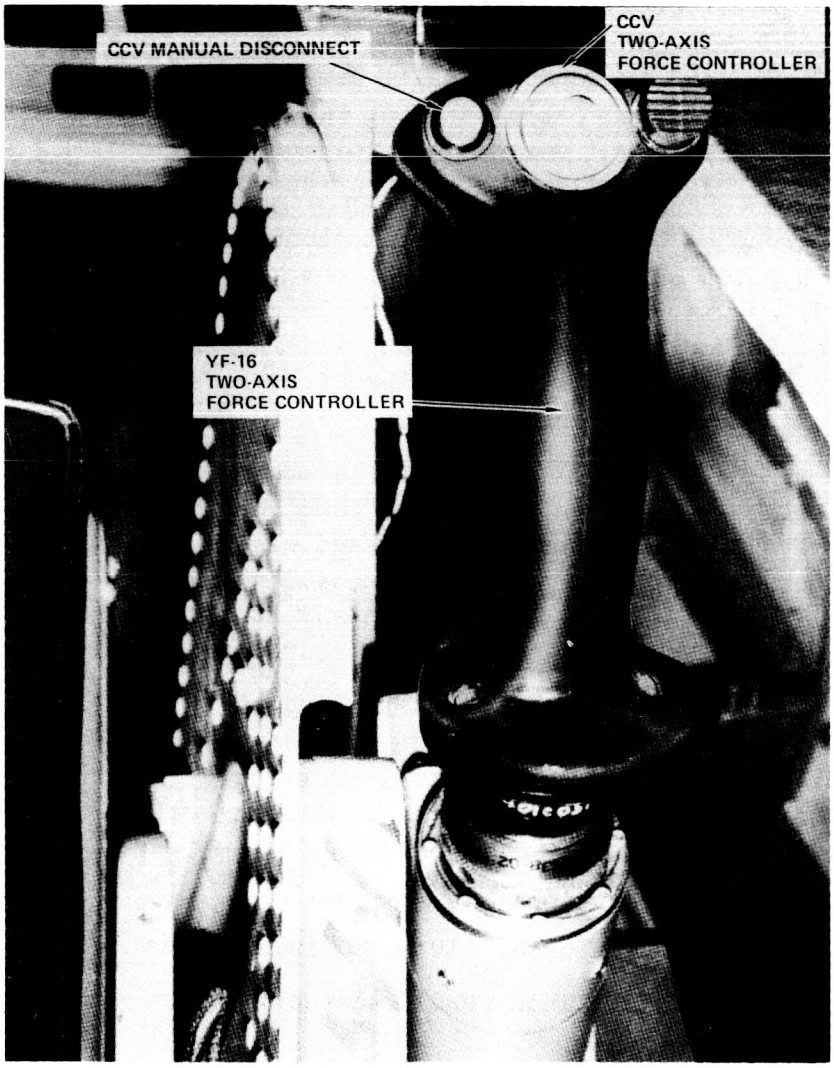


Figure 7. - Fighter CCV Force Button

The "coolie hat" switch on the sidestick controller worked satisfactorily, but the pilots had tendency to saturate the pointing system with the switch and then fly the attitude with the normal sidestick controller pitch control. The emphasis of this design was to provide an expedient controller solution and therefore these "add ons" to the existing control system were cost effective, but were not optimized for operational effectiveness.

The CCV program demonstrated the feasibility of decoupled control and provided an indication of its usefulness, but it was obvious that an additional research effort was necessary to develop the decoupled controllers into an operationally suitable implementation.

Based on crew mockups and the results of the fighter CCV longitudinal controller evaluation, two conclusions were reached for the AFTI/F-16. First, the secondary longitudinal controller must be removed from the primary longitudinal sidestick controller to eliminate the cross coupling effect seen during the fighter CCV project and secondly, another separate, controller could not be reasonably installed into the cockpit due to space limitations. As a result of these and the operationally desirable concept of "hands on controls at all times", the conclusion that adding the control function to the already existing throttle and using a control motion that would correspond to the resulting aircraft motion was the best approach.

Several approaches were condsidered, including 1) a finger operated momentary switch, 2) a slip ring around the throttle operated by a twist of a finger, and 3) a twist of the entire throttle grip. Study of these methods, using conceptual mockups, established that twisting the entire throttle grip was the preferred approach. Twisting the entire throttle grip provided a motion directly analogous to the desired aircraft motion and flexibility for development of grip motion magnitudes and force gradients.

After selection of the controller was concluded, development of the current throttle for the AFTI/F-16 aircraft progressed with consideration given to the following subjects: 1) flight control suitability, 2) power adjustment suitability, 3) throttle location and range, 4) grip shape and switch locations.

Flight Control Consideration - The use of the throttle grip for controlling the longitudinal decoupled control axis was studied to establish the force gradient and range of rotation to be used. Based on discussions with N.A.S.A. space physiology specialists and the results of force gradient tests using the General Dynamics test pilot community, the following values were selected as lower and upper limits:

a. total range of rotation: 20^{o} to 40^{o}

b. maximum rotational force: 10 to 30 inch-pounds

c. break-out force: 6 inch punds (nominal)

The prototype throttle was designed to accomodate any of these values by simple changes of torsion springs and location of mechanical stops. The fuel system parameters values selected for flight testing were determined as a result of simulator and centrifuge evaluations.

Power Adjustment Consideration - The existing F-16 throttle operates with a conventional, throttle quadrant arc motion. The twist grip features impose two new considerations on the engine power adjustment motion of the throttle. A linear motion is preferred on AFTI/F-16 for power adjustment to preclude inadvertent twisting of the grip. Also, total throttle throw must be compatible with pilot reach with the hand rotated to the maximum twist grip deflection position. The F-16 throttle also requires a lifting motion of the grip to enter the afterburner range. To command idle cut-off, the throttle grip must be lifted while pulling a release trigger. These features are compatible with the proposed AFTI/F-16 linear throttle motion, and were compared to other approaches. For afterburner entry, however, the lift feature was believed less desirable than a detent cue when the pilot is under high normal accelerations. Pilot interviews indicated good general acceptance of the F-5 throttle which provides a stong detent cue on entering afterburner and a lighter detent cue on exiting the afterburner range. The action requires no lifting motion, only an increase in the normal advance/retard forces. To advance the F-5 throttle the detent requires a total force of 7 pounds and to retard requires 4.5 pounds and these criteria were used for the AFTI/F-16 design. The normal advance/retard forces outside the detent area are essentially unchanged from the F-16, and coupling of the linear throttle motion into the F-16 throttle linkage was done so as to not affect there forces. To prevent inadvertent entrance into idle cut-off, the F-16 feature that requires pulling a release trigger was retained. The requirement for lifting the throttle grip was determined to be unnecessary and was eliminated. In advertent actuation of the trigger was prevented by adding a guard in front of the trigger.

Throttle Location and Range - The throttle mid-range position is unchanged from the existing F-16 location. As a result of the linear throttle, a new throttle range of translation of seven inches was developed. This was verified thru an anthropemetric study and AFTI/F-16 cockpit mock-up.

Grip Shape and Switch Locations - The F-16 throttle includes numerous other functions as shown in Figure 8. The compatibility of the F-16 throttle switch arrangement, the AFTI/F-16 requirements and various grip shapes were studied. No shapes or switch arrangements were found that were superior, to the current F-16 configuration and therefore, the standard F-16 grip shape was adopted for the AFTI/F-16.

The new twist grip throttle controller was tested by AFTI project pilots in two situations, the Air Force Aerospace Medical Research Laboratory (AFAMRL) human centrifuge and the General Dynamics, Research and Engineering (R&E) AFTI/F-16 simulator, prior to aircraft installation and subsequent flight testing.

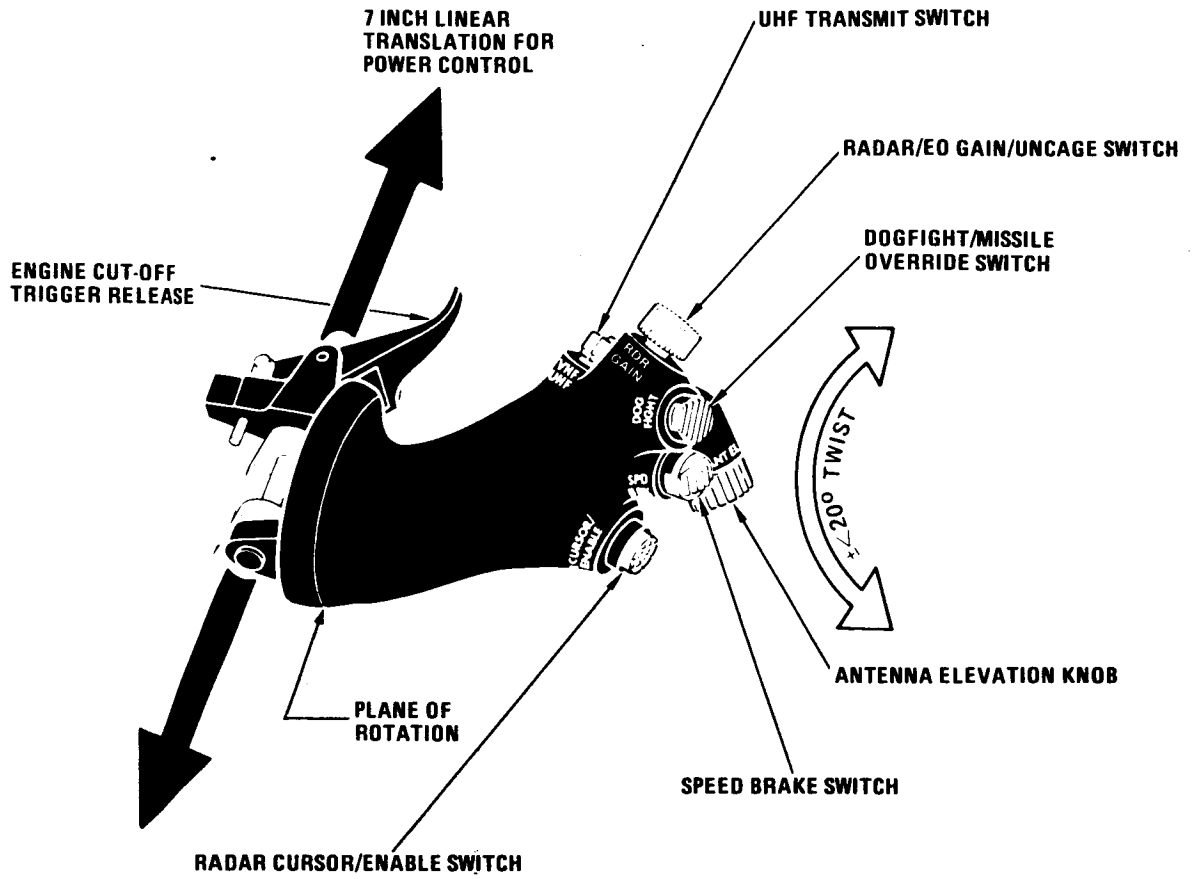


Figure 8. - AFTI/F-16 Throttle Twist Grip Controller

The AFAMRL centrifuge testing was accomplished to determine the presence and extent of any biodynamic coupling of control inputs into the throttle controller as a result of dynamic accelerations (Gz and Gy). During this testing, the subject pilots were permitted closed loop control of the Z axis acceleration at levels up to 6 Gz and y axis accelerations at levels up to 2 Gy. The results indicated that no significant errors were observed under conditions other than relatively high sustained Gz. Twist throttle errors in the pitch up direction were seen as high as 12%, with pitch down errors shown in Figure 9, reaching as high as 18%.

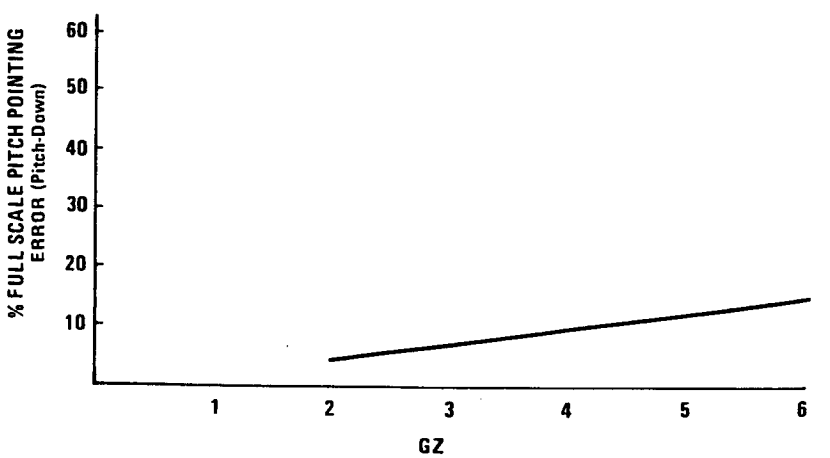


Figure 9. - Throttle Pitch Pointing Errors Under Normal Acceleration (+GZ)

The throttle tested in the General Dynamics R&E simulator was the culmination of the previous design work and testing. The simulator objective was to validate the throttle controller development. Mechanical design (force gradients, breakout forces, etc.) and controllability were evaluated by comparison against a baseline controller, namely the existing sidestick controller, in performing similiar tracking task with all other interactions controlled. The testing consisted of each of the project pilots performing multiple repetitions of single axis tracking of an air-to-air target.

The target motion consisted of a complex sinusiodal (varying amplitude and frequencies) function in the vertical axis only, with the pilots task to track this target with each of the two controllers using a fixed aiming reticle. The comparison, shown in Figure 10, of the controller errors while performing this tracking task demonstrates that the throttle controller was an effective controller and that with minor modification to mechanical design could be implemented into the aircraft for flight testing.

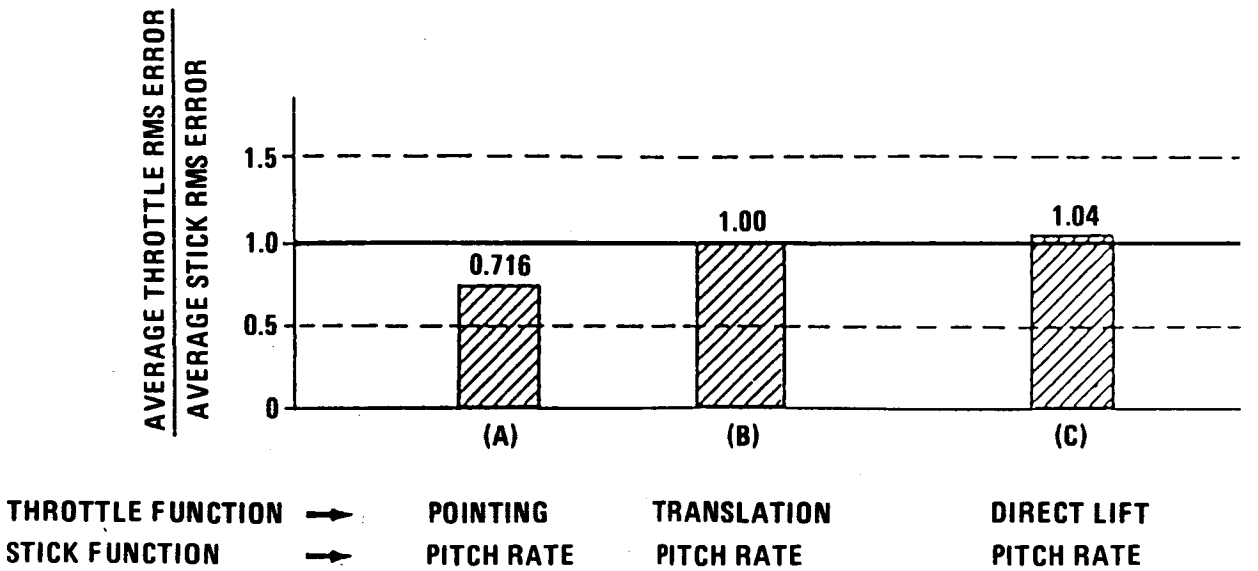


Figure 10. - Controller Tracking Errors

FLIGHT TEST RESULTS

The AFTI/F-16 is now in the flight test portion of the DFCS phase of the program, and therefore, any conclusions made about the effectiveness of the throttle twist configuration in an operational environment as a hands-on combat controller are only preliminary. There are, however, valid points to be made from the data collected to date.

I. It has been found that breakout forces should be increased based on inadvertent pilot inputs during periods when the twist grip throttle was not being used as a longitudinal controller. An alternative to higher breakout forces would be a latch switch to inhibit inadvertant throttle rotation during non-twist controller activities.

II. It has been found that a major criteria for judging controller effectiveness is the ability of the pilot to use the controller in a consistant linear fashion. Each longitudinal decoupled mode is designed to operate with different authorities. The authority limits built into the direct lift and translation modes have been sufficient to gather accurate data on the twist throttle controller effectiveness. However, the pitch pointing authority is very small. Although the airplane response is quick and well damped, the magnitude of the response is small and therefore the pilot is forced to treat the throttle controller as an on-off device rather than a proportional controller.

III. The twist throttle controller design has met many of the objectives and requirements imposed by the AFTI/F-16 program. This is especially true in meeting the "hands-on control at all times" requirements. However, the primary objective that drove the design, i.e. seperating primary coupled aircraft motion from secondary decoupled aircraft motion, is being evaluated during current flight test efforts in an operational environment.

SUMMARY & CONCLUSIONS

The twist grip throttle controller design used on the AFTI/F-16 has gone through many developmental and evolutionary phases. These phases have been:

(1) Review of F-16 fighter CCV program

(2) Candidate definition and preliminary evaluations

(3) Effectiveness evaluations

  (a) Centrifuge tests

  (b) Evaluations on fixed base simulator

The results of the evaluation has in general shown the twist throttle controller to be a successful design. The results of the evaluation have shown the twist throttle to provide:

(1) Easy of operation

(2) Consistent aircraft motion

(3) Required hands-on control

(4) Means to eliminate the need to add a new controller to the cockpit

(5) Smaller breakout forces than needed

(6) Good pilot acceptance

(7) Good marks when used in the direct lift and translation modes

The flight test program for the DFCS phase of the AFTI/F-16 program is continuing and will be completed in July of 1983.

References

1. Anon. "Human Engineering Crew Station Design Report," AFTI/F-16 Report No. 20PR064, General Dynamics Corp., 9 April 1980

2. Anderson, D. C., Smith K. L., and Watson, J. H., "AFTI/F-16 Advanced Multimode Flight Control System Development Concepts and Design," AIAA Paper 82-1571 CP, August 1982

3. Barfield, A. F., Van Vliet, B. W., and Anderson, D. C., "AFTI/F-16 Advanced Multimode Control System Design for Task-Tailored Operation," AIAA Paper 81-1707, August 1981

4. Barfield, A. F., (AF Flight Dynamics Laboratory/Wright-Patterson AFB, Ohio) "Flight Experience with Manually Controlled Unconventional Aircraft Motion," 14th Annual Manual Control Conference, Los Angeles, California August 1978

5. The Air Force Aerospace Medical Research Laboratory - Biodynamic and Bioengineering Division - Acceleration Effects Branch (AFAMRL/BBS), "Investigation of the Effects of Gy and Gz on AFTI/F-16 Control Inputs, Restraints, and Tracking Performance," August 1981

Assessment of a Trajectory Guidance Display

Barbara Bernabe, Ph.D.
Florida Institute of Technology

Abstract
Informal Presentation

Test pilots have the unique task of flying an aircraft to the outer limits of its flight envelope and holding that maneuver for a duration extended enough for data gathering. For difficult maneuvers both getting "on condition" and maintaining a test point for even a few seconds may be close to or completely unmanageable.

One inherent difficulty in flying a difficult maneuver lies in having to sequentially sample information from a number of sources spread out over the instrument display and translate those discrete pieces of information into manual control inputs which in turn damp the error signal for the individual parameters specified by the flight profile. This may be considered to be a form of perceptual/cognitive workload. To the extent that perceptual and cognitive processing may be simplified either through the reduction of perceptual scanning requirements and/or by minimizing focal attentional requirements, manual control difficulties may also be attenuated. One way to accomplish this is to present the necessary information in an integrated manner which simultaneously displays error from the desired state for all the specified parameters of the test condition. Such a display called a Flight Trajectory Controller was developed and underwent preliminary testing at NASA Dryden Flight Research Facility.

Display Components

The display consists of horizontal and vertical crossbars which appear similar to glideslope/localizer needles, and a slide pointer. The indicators are driven by alogrithms which compare current vehicle state with desired trajectory. This information is then synthesized to provide error signals to the pilot, in the form of deviation of indicator components from the display center. The horizontal and vertical needles provide pitch and roll stick control information respectively, while the side pointer provides throttle information. Errors are damped by following each of the indicator needles with control stick movements. Thus, all needles function as fly to commands, and the pilot's task essentially becomes one of tracking and centering the indicators with stick and thrust control commands.

Experimental Test

In two simulated maneuvers, performance, as measured by time to achieve condition and elapsed time on condition, was facilitated by the trajectory controller when compared with performance using conventional instrumentation. In addition, pilot workload, as measured by performance on two secondary tasks was significantly reduced in the trajectory display condition. General guidelines for the utilization of display in reducing manual control difficulty will be discussed.

Reaction Time Studies of Separtion Violation Detection with Cockpit Traffic Displays

Patrick T. Lester and Everett A. Palmer
NASA Ames Research Center
Moffett Field, CA

Abstract

In phase one of this study 12 pilots made judgments or current of future horizontal separation violations with either ground-referenced or ownship-referenced predictors of other aircraft future position on a cockpit traffic display. Pilots had significantly shorter reaction times when using the ownship-referenced predictor than when using the ground-referenced predictor. Pilots also had more errors with the ground-referenced predictor. These effects were strongest when there were four other aircraft on the display.

In phase two, 12 pilots made judgments of current or future vertical separation violations with three types of data tags. Pilots took longer to respond and made more errors on the data tag which showed the other aircraft altitude and vertical velocity than on data tags which showed predictions of the other aircraft future altitude.

Introduction

Cockpit traffic displays are being evaluated as a potential aid to the current Air Traffic Control system to maintain or decrease the required separation of aircraft, with no loss of safety, in the National Airspace System as expected air traffic increases. The display is a cathode ray tube based plan-view display that presents the pilots' own aircraft (ownship), routes, and other aircraft in its vicinity (Verstynen, 1980).

Past studies of cockpit traffic display research have dealt with pilot judgments of whether a displayed aircraft would pass in front of or in back of ownship (for a summary of these studies see Palmer et al, 1980). In these studies, traffic aircraft were programmed to miss ownship by 1 nm or less, and the pilots knew that a separation violation would occur. But how will pilots perceive traffic aircraft that will pass further away from ownship horizontally or vertically, but are closing on ownship's position?

The experimental task reported here required the pilots to decide whether or not there would be a horizontal or vertical separation violation with any of the aircraft on the display. The objective of the experiments was to evaluate the usefulness of various displayed information in helping pilots make these judgments.

Method

Procedure

Subjects. Twenty-four commercial airline pilots participated as paid volunteers (12 for each phase of the experiment). The subjects were obtained through the NASA Ames subject pool.

Hardware. Three pilots were tested at one time in three separate fixed-based cockpit simulators with static cockpit traffic displays generated by a general purpose stroke-writing computer graphics system. The system was controlled by a Digital Equipment Corporation PDP 11/70 computer coupled with an Evans and Sutherland Picture System II.

Display Conditions. The display variables held constant in both experiments were the route marking (3 nm wide), a 3 nm radius range ring around ownship's position and a map range of 15 nm. All aircraft were in straight and level flight. Ownship was always displayed with a ground-referenced predictor (See Figure 1).

Phase 1

Task. The subject's task was to view the display and answer the question: "Will any aircraft pass within 3 nautical miles of ownship, provided that all aircraft will maintain their current speed and heading?" The subject indicated his response by pushing forward on a response lever for "YES" or pulling backward for "NO".

Predictor Conditions. The ground-referenced predictor shows where the traffic aircraft will be in 60 sec relative to the ground (Figure 1). In this example, the tip of the 60 sec ground-referenced predictor on Golden Gate 190 (GG190) indicates that in 60 sec GG190 will be less than 3 nm from ownship if both aircraft maintain their present speed and course. GG190 will come within 3 nm of ownship, while Golden Gate flight 534 (GG534) and United Airlines flight 527 (UA527) will not come within 3 nm and are traffic aircraft. The pilot would respond "YES" to the GG190 aircraft.

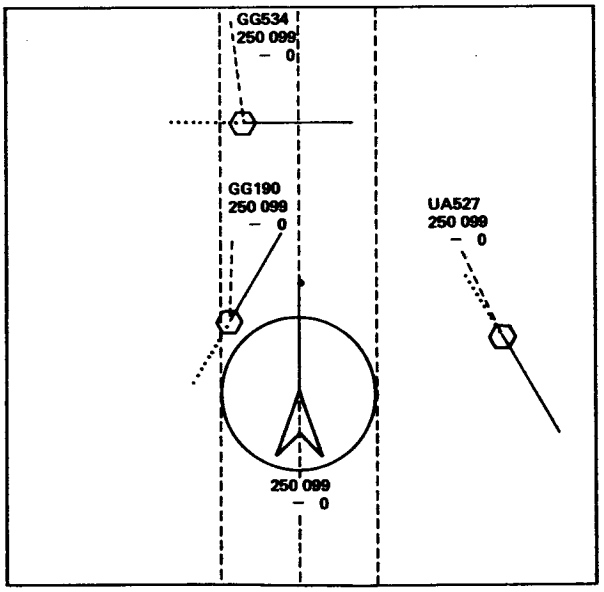


Figure 1. The cockpit traffic display presented with the ground referenced predictor.

The ownship-referenced predictor shows where the traffic aircraft will be relative to ownship. That is, it takes into account ownship velocity. Whenever a traffic aircraft's ownship-referenced predictor is in or will project through the 3 nm range ring of ownship, a horizontal separation violation has occurred or will occur. In this example, GG 631's predictor is penetrating ownship's range ring and will be a separation violation (Figure 2).

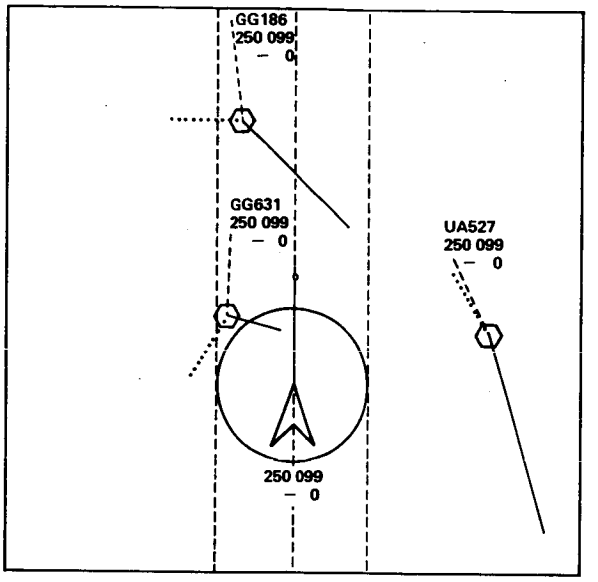


Figure 2. The cockpit traffic display presented with the ownship-referenced predictor.

Trials. Subjects were given a written introduction and a dynamic demonstration of the traffic display. They received one block of 30 practice trials for each of the predictor conditions. Practice trials had two traffic/intruder aircraft. At the beginning of each trial, a "computer ready" light flashed. When each pilot was ready, he pushed a "ready" button on the response lever. When all of the pilots had responded, the intruder aircraft appeared. Promptness of the response was stressed as well as the accuracy of the response. When all of the pilots had responded, each received feedback as to the correctness of their response. Response time feedback was given to the nearest 0.1 sec. In an attempt to control for conflict between speed and accuracy, response times were not displayed when the pilot did not respond correctly. The display did, however, tell the pilot when he had made an incorrect response. When all of the pilots had indicated their readiness by pushing the ready button, the new trial began.

Subjects received one block of 30 trials of one, two, and four aircraft (in addition to ownship) for each of the predictor conditions. There was only one programmed intruder in the traffic. The intruder and non-intruder (traffic) were taken from a set of preprogrammed traffic files.

The experimental encounter parameters were: final miss distances of 1, 2, 4, 5, 6, and 7 nm; headings of +/- 30, 60, 90, and 150 degrees from ownship;

and time to minimum miss of +/- 0, 30, 60, and 90 sec. Negative miss distances were used for background (traffic) aircraft and meant that the aircraft was past the point of the minimum miss distance. All intruder and traffic aircraft were at the same altitude as ownship.

In 50% of the trials one of the aircraft would (if it remained on a straight course) come within 3 nm of ownship or already be within 3 nm of ownship. In the remaining 50% of the trials, one or more aircraft appeared on the screen but none came closer than 3 nm from ownship and were merely traffic aircraft.

Subjects were tested in two conditions of aircraft predictors, ground-referenced and ownship-referenced predictors on traffic/intruder aircraft.

All conditions were counter balanced across groups of three subjects. Each set of subjects received one block of 30 trials for each of the display conditions. Trials were run at a rate of two per minute. Subjects received six blocks of 30 trials of approximately 15 min each. There was a 15 min break between the ground-referenced and ownship-referenced predictor conditions.

Results
Phase 1

Error Rate. An analysis of variance (ANOVA) was performed. Significant differences ($p < 0.001$) were found between both the predictor types and the number of aircraft. Figure 3 indicates the effect of aircraft density on the ownship referenced predictor.

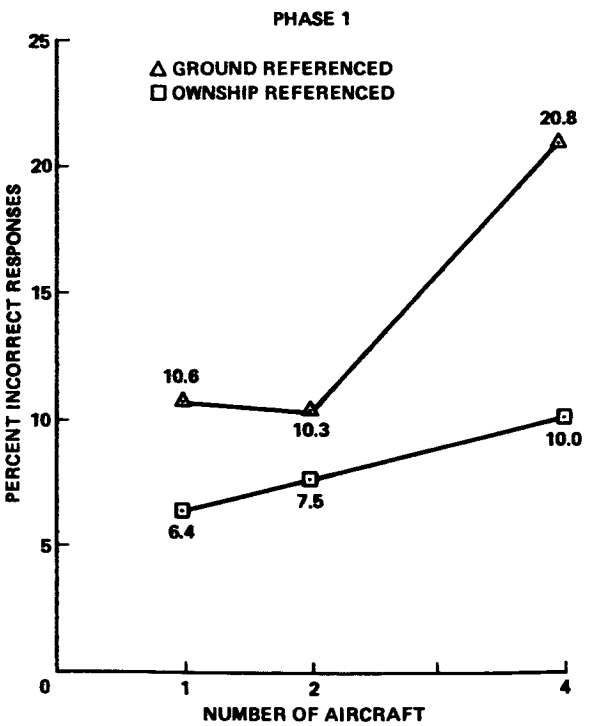


Figure 3. Error rate performance with the ground-referenced and ownship-referenced predictors.

The mean overall error rate was lower for the ownship-referenced predictor than for the ground-referenced predictor. This effect was most noticable at the higher aircraft density of 4 aircraft.

Response Times. An ANOVA showed significant differences ($p < 0.001$) in response times for predictor type and number of aircraft. The correct mean response time for all subjects was calculated. Figure 4 indicates that subjects had an increase in response time with an increase in aircraft density for both the ground-referenced and ownship-referenced predictors.

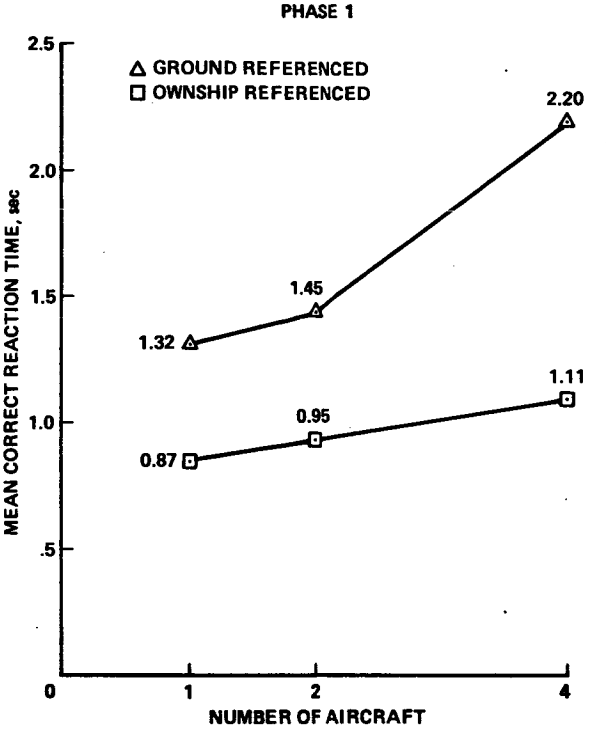


Figure 4. Response time performance with the ground-referenced and ownship-referenced predictors.

With the ownship-referenced predicator, however, the subjects had overall faster response times than with the ground-referencd predictor. This effect was most noticable at the higher aircraft density of 4 aircraft. Standard errors of the mean response times were small, ranging from 0.01 to 0.23 sec.

Phase 2

Task. The subject's task was to view the traffic display and answer the question: "Will any aircraft be within 1000 ft vertically of ownship when they are at their closest point to ownship, provided that all aircraft will maintain their speed, heading, and vertical velocity displayed?"

Trials. All display hardware and display conditions were identical to those in Phase 1. In 50% of the trials an aircraft would (if it remained on a straight course) come within 1000 ft of ownship or already be within 1000 ft vertically of ownship. In the remaining 50% of the trails, one or more traffic aircraft would appear on the screen but not come closer than 1000 ft vertically from ownship. Thus, 50% of the responses were "YES" and 50% were "NO" when answered correctly by the pilot.

The experimental encounter parameters were: speed of 250+/-30 knots; heading 90+/-45 degrees from ownship; time to minimum miss of +/-0, 30, 60, 90 sec; and vertical miss distances of +/-333, 667, 1333, 1667, 2000, and 2333 ft. Negative times to minimum miss were used only on traffic aircraft. The aircraft were displayed with ground-referenced predictors. All intruder/traffic aircraft were programmed to pass between +/-0 and 1 nm of ownship and would be horizontal sepration violations.

Subjects were tested with three types of data tags giving altitude information. The information displayed on the "normal" data tag included airline and flight number (Golden Gate flight 631, ground speed (280 knots), altitude (8300 ft), and vertical speed (ascending 2000 ft/min) of the intruder aircraft.

GG631
248 083
+ 20

The "absolute" data tag shown below, displayed an aircraft's current altitude (8300 ft) and its altitude when it is in closest proximity to ownship (9300 ft). Current altitude was displayed with a leading zero (0) digit to differentiate it from the absolute altitude. Ground speed was not displayed.

GG631
083
93

The "relative" data tag (below) showed the aircraft's current altitude (8300 ft) and the difference in altitude in hundreds of feet when the aircraft was in closest proximity to ownship (700 ft below ownship). Ground speed information was not displayed.

GG631
083
- 7

Trials were run at a rate of two per minute. Subjects will receive six blocks of 36 trials of approximately 15 min each. There were 15 min breaks between the data tag conditions.

Results
Phase 2

Error Rate. An ANOVA showed significant differences ($p<0.001$) among data tag types and number of aircraft for the error rate measure. Subjects made a singificantly greater number of errors with the normal data tag than with the absolute or relative data tags (Figure 5).

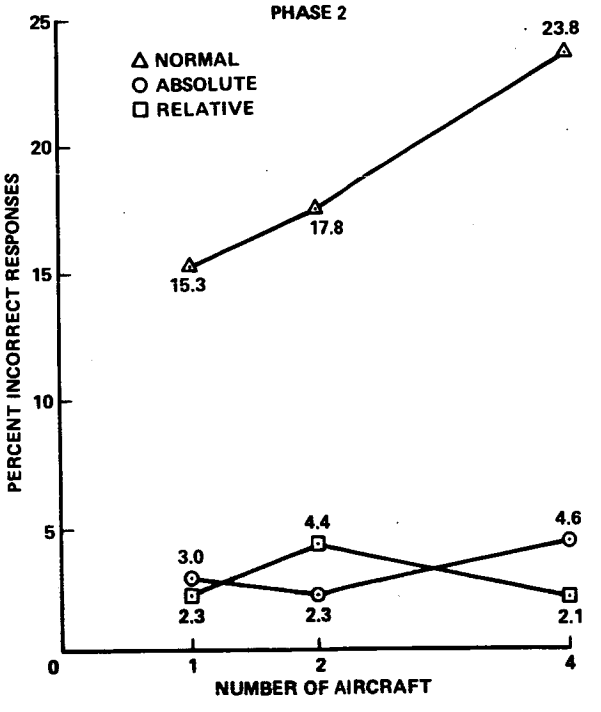


Figure 5. Error rate performance of the three types of data tags.

There was little difference between the absolute and relative data tags. There was an increase in errror rate with an increase in aircraft density for the normal data tag.

Response Time. An ANOVA showed significant differences ($p<0.001$) in data tag types and number of aircraft for the response time data. The total mean response time was found for each of the correct trials of the three data tag conditions. Figure 6 shows the great differences in the mean reponse time between the normal data tag and the absolute and relative data tags.

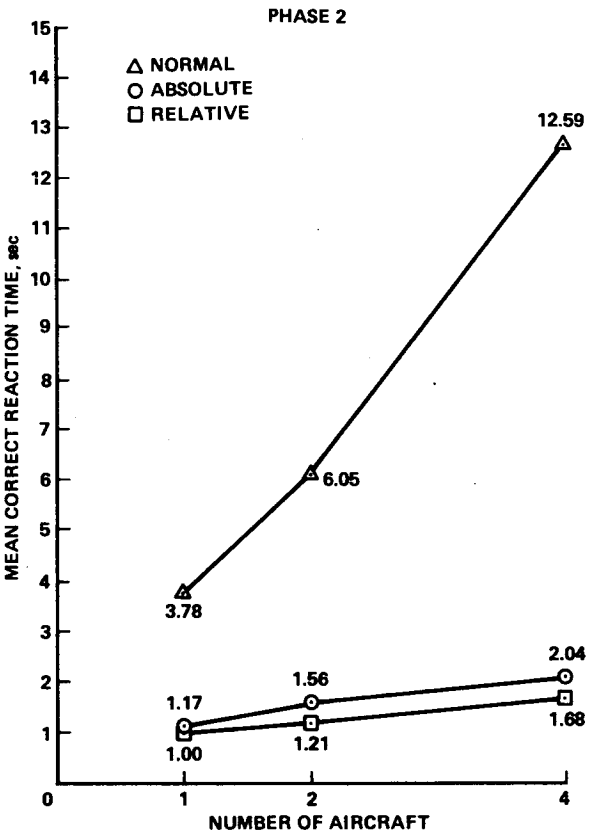


Figure 6. Response time performance with the three types of data tags.

The normal tag gave greater response times than the absolute or relative data tags. There was an increase in response time for all of the data tag types with an increase in aircraft density. There was little difference in response times between the absolute and relative data tag types. Standard errors of mean response times were minimal, ranging from 0.02 to 0.13 sec.

Discussion

Pilots stated that ground-referenced information of traffic would be helpful at certain times, but preferred ownship-referenced information of traffic for a quick check of horizontal traffic separation.

Ground-referenced information is particularly useful in monitoring whether or not a following aircraft is maintaining adequate separation.

The ownship-referenced predictor on traffic aircraft caused some confussion when used in conjunction with ownship's ground-referenced predictor. In debriefings, some pilots stated that they were comparing the

60 sec ground-referenced predictor of ownship to the 60 sec ownship-referenced predictors of the traffic aircraft and were coming up with erroneous responses. It is possible that this problem may be eliminated with further training.

The normal data tag showed the poorest performance in a task to determine vertical separation of ownship and traffic aircraft. Although there were no significant differences between the absolute and relative data tags in either the reaction time or error rate measures, in debriefings, pilots thought that the relative information was the best.

The information provided by the absolute and relative data tags contains predictive information of vertical separation distances of traffic aircraft. This information may not be enough for other situations. Pilots, in fact, stated in debriefings that they preferred to have the vertical speed information displayed as it was on the normal data tag and have the relative information on demand if the situation called for monitoring of a potential separation problem.

References

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, October 1980, 605-620.

Verstynen, H. A. Potential roles for the cockpit traffic display in the evolving ATC system. Society of Automotive Engineers (SAE-800736), Warrendale, PA, 1980.

MODELING THE EFFECT OF LOUD BLASTS AND DISTURBANCES ON HUMAN'S TRACKING ABILITY*

Susan M. Krolewski and Jonathan Korn

ALPHATECH, Inc.
3 New England Executive Park
Burlington, Massachusetts 01803

Abstract

A computer model designed to generate sample path time histories of human tracking error in different antitank systems was developed in [1,2]. This paper focuses on the development of a model to predict the effect of auditory disturbances, such as loud blasts, on an operator's tracking ability. The model was based on the assumption that the decrease in tracking performance following a disturbance is attributed to an increase in human motor noise. Analysis of data supplied by Army Materiel Systems Analysis Activity (AMSAA) demonstrates that model predictions are consistent with the characteristics and magnitude of actual tracking errors.

Introduction

A computer model was developed [1,2] to evaluate the performance of manual tracking of several different antitank systems. This model is based on the Optimal Control Model (OCM) of Human Response [3]. The computer antitank model (ATM) is capable of generating Monte Carlo sample paths of human tracking error for both azimuth and elevation axes. Data supplied by Army Materiel Systems Analysis Activity (AMSAA) was compared to model predictions with excellent agreement.

In this work, the ATM is extended to predict the reaction of an operator to a nearby explosion. Typically the gunner reacts to such a disturbance by "flinching". During a flinching episode the operator's ability to track is momentarily impaired. This results in high tracking errors with high variability. This paper focuses on the development of a "flinching" model designed to predict an operator's response to such disturbances. Ensemble data from flinching experiments conducted by AMSAA is used to determine the model parameters.

* Work supported by Army Materiel Systems Analysis Activity, Aberdeen Proving Grounds, Maryland under contract number DAAK-11-82-C-0015

Model Formulation

The model is based on the assumption that the decrease in tracking performance, during a flinching episode, can be modeled by an increase in human motor noise rather than a basic change in his control strategy. The increased motor noise is characterized by a bias and increased covariance during the period immediately following the blast. Both components decay exponentially. The modeling approach is formulated in the following.

In the OCM, the operator's control signals, u_A and u_E, for azimuth and elevation respectively, are defined by

$$\tau_n \dot{u}_i(t) + u_i(t) = u_{ci}(t) + v_{ui}(t) \quad i = A \text{ or } E$$

where τ_n is the neuromotor time constant, u_c is the commanded control signal and v_u is the human generated motor noise. Nominally the motor noise is zero-mean Gaussian with covariance

$$V_{ui}(t) = V_{ui}^{\circ}(t) + \pi\rho_{ui}E\{u^2_{ci}(t)\} + \pi\rho_{ij}E\{u^2_{cj}(t)\} \qquad \begin{matrix} i, j = A \text{ or } E \\ i \neq j \end{matrix}$$

where $V_{ui}^{\circ}(t)$ is the additive part, ρ_{ui} is the motor noise ratio coefficient, ρ_{ij} is the interaxis crossfeed and u_{ci} is either the commanded displacement or rate.

During flinching, an exponentially decaying bias is added to the operator's motor noise. The "flinching" motor noise, $v_{ui}^*(t)$, is given by

$$v_{ui}^* = v_{ui}(t) + m_i(t) \qquad i = A \text{ or } E$$

The bias, m(t) is defined by the expression

$$m_i(t) = \begin{cases} 0 & t < t_f \\ M_i e^{-(t - t_f)/\tau_f} & t \geq t_f \end{cases}$$

where M_i are constants dependent on the magnitude of the blast, t_f is the time of flinch, and τ_f is the recovery time constant.

The motor noise ratio coefficients are also increased during flinching. The "flinching" motor noise ratio coefficients, ρ_{ui}^*, are defined by

$$\rho_{ui}^* = \begin{cases} \rho_{ui} & t < t_f \\ 1 + Ke^{-(t - t_f)\tau_f} & t \geq t_f \\ \rho_{ui} & \end{cases}$$

where K is a constant dependent on blast magnitude.

Data and Model Predictions

Human tracking data from controlled field tests was supplied by AMSAA for purpose of model tuning and validation. The data consisted of azimuth and elevation tracking error time histories including a flinching episode. For each of these runs the temporal mean, μ_E, and standard deviation σ_E was calculated. These quantities were averaged to find the ensemble statistics $\bar{\mu}_E$ and $\bar{\sigma}_E$.

To illustrate the model's ability to reproduce flinching, model predictions were generated for comparison with the data. Figure 1 compares the azimuth axis ensemble statistics for nine runs of the DRAGON data with the statistics for 15 runs of modeled DRAGON results. The values chosen for M_A, K_A, and τ_f were 14.0, 15.0, and .5 seconds respectively. In each case flinching occurs at six seconds.

The model predictions have the same duration and magnitude as the DRAGON data. The major discrepancy is the oscillatory behavior of the model response. Further tuning of the DRAGON parameters should eliminate this problem.

Conclusions

The flinching model was developed to enable the prediction of tracking errors following a loud explosion or other auditory disturbances. Flinching results in higher tracking errors with high variability. This degradation in tracking performance was modeled by a temporary increase in motor noise bias and variance. Data obtained from AMSAA was reduced and compared to model predictions. The model's ability to replicate the magnitude, duration and variability of tracking errors during flinching was demonstrated.

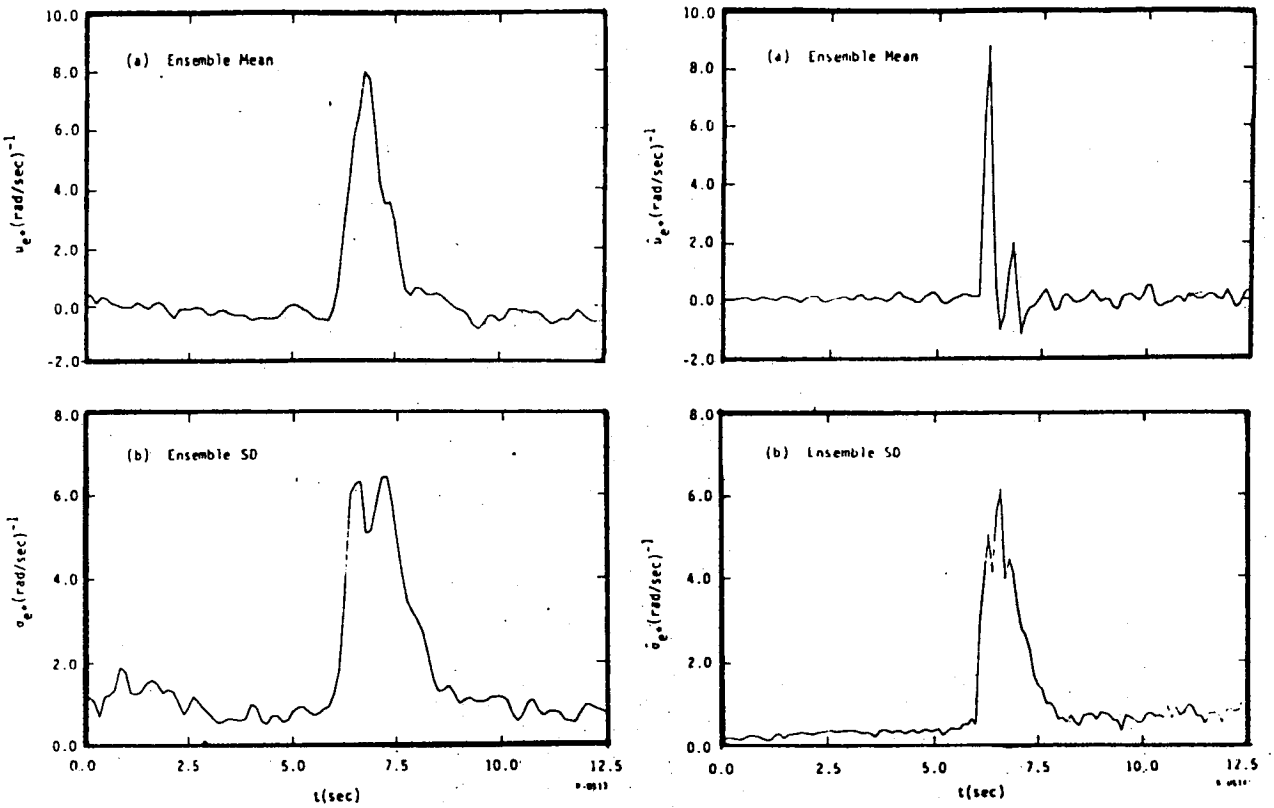


Figure 1 comparison of DRAGON azimuth tracking error data with model predictions.
(a) Data (b) Model

REFERENCES

1. Kleinman, D.L., "A Model for Generating Human Tracking Error in Antitank Systems," ALPHATECH Report Number TR-113, ALPHATECH, Inc., 3 New England Executive Park, Burlington, Massachusetts, August 1981.

2. Korn, Jonathan, "Human Operator in the Context of Tank and Antitank Systems," Proceedings of the Conference of IEEE Systems, Man and Cybernetics Society, Seattle, Washington, October 1982.

3. Kleinman, D.L., S. Baron and W.H. Levison, "A Control Theoretic Approach to Manned Vehicle Systems Analysis." IEEE Transactions on Automatic Control, Volume AC-16, No. 6, 1971.

ADAPTATION OF TRACKING BEHAVIOUR TO DISTURBANCES WITH TIME-VARYING CHARACTERISTICS

TOM BÖSSER
Psychologisches Institut
der Westfälischen Wilhelms-Universität
Schlaunstr.2
D-44 MÜNSTER
West-Germany

ABSTRACT. A tracking-task requiring the subject to employ an asymmetric cost-function induces a nonlinear tracking-strategy, which also depends upon bandwidth of the input. This task is used to study adaptation of tracking strategy to time-varying input signals. Results indicate that subjects do not base choice of a strategy on an estimate of the statistics of the disturbance, but rather elicit a sequence of discrete responses. Linear-filter analogies of tracking behaviour have to be considered limited to the domain of tasks where quadratic cost-functions can be assumed.

INTRODUCTION.

Tracking behaviour is a compound task comprising perceptual and motor processes, motivation and learning. Our focus is mainly on motivational processes, which we assume to be involved in the weighting given to outcomes of actions, in other words, they represent the criteria or cost-functions for optimization of performance. We see the functional structure of the processing involved in tracking behaviour as shown in Fig.1: Perceptual-motor performance represents the basic limitations of performance, prediction is used to compensate for these. But also there are processes with the function of finding the best performance, given limitations of the processing system as well as constraints set by the outside world.

It has been suggested before that the linear tracking behaviour found in many studies is a special case, a function of a quadratic error-criterion beeing induced by instruction (BÖSSER 1982, 1983). Previous studies have shown that a nonlinear tracking-strategy arises when an asymmetric error-criterion is induced by instruction or payoff-conditions. The nonlinear strategy interacts with the spectral contents (bandwidth) of the disturbance: With higher bandwidth of the disturbance the nonlinear component in the tracking behaviour increases.

Tracking strategy under these conditions adapts to the statistical properties of the disturbance. This has been found in the studies

mentioned above, and may also be inferred from studies of car-following (DREYER 1979, BÖSSER 1980). With an asymmetric cost-function a control strategy may be successful, which accepts a larger RMS-error, but reduces large errors in one direction away from the target. This process of optimization requires, apart from knowledge of the cost-function, an estimate of the statistics of the disturbance.

The tracking-task which we have previously employed to investigate nonlinear tracking strategies can be used to test adaptation of tracking strategy to varying bandwidth.

In order to derive the principles upon which the present investigation is based, some of the results obtained previously (BÖSSER 1982, 1983) have to be summarized: Theoretical considerations emanate from the consideration of cost- or utility-functions employed for weighting the error in a tracking task. The statistical methodology available and also general theoretical considerations suggest the use of a quadratic cost-function: Error is weighted symmetrically, large deviations from target are given unproportionally larger weight than small errors - these seem to be reasonable assumptions for many tasks. It can also be seen from Fig.2 that a quadratic cost-function represents a reasonable approximation to other symmetric cost-functions (proportional and fixed bilateral limit), but not to an asymmetric cost-function. The assumption of a symmetric cost-function is one of the basic assumptions of estimation theory.

Whereas many tasks, eg lateral control of a car or aeroplane, induce a symmetric weighting of error, certain tasks imply asymmetric cost-functions: This is the case with distance-control in driving, perhaps operation of cranes and landing an aeroplane. (This research was instigated by the observation that distance-control in driving is not described well by the usual linear describing-functions.)

In our experiments a task with asymmetric error-criterion was employed, where a standard compensatory tracking-task is augmented by a unilateral 'borderline' to the display (Fig.3). A number of experiments have shown that cost-functions of subjects can be variied in this experiment effectively by instruction or by pay-off conditions being made dependent on different weighted sums of the two error-conditions. As a consequence of this asymmetric cost-function, subjects assume a nonlinear control-strategy which is apparent from two types of parameters:
- a mean-error away from the borderline is introduced, increasing with bandwidth of the input-signal and
- nonlinear dynamics occur which can be estimated from the second-order WIENER-kernel.

These two properties of the control strategy need not be independent of one another, it is conceivable that the mean-error arises as a consequence of the asymmetric control strategy. Time courses of track, output and linear estimate based on the transfer-function are shown in Fig.4 to illustrate the effect of the experimental condition.

The fact that the mean-error observed is a function of the bandwidth of the disturbance can be employed to identify the adaptation of tracking-behaviour to a disturbance with time-varying properties. We expect to be able to see from the way the mean-error trails the variation of the bandwidth of the disturbance in which way subjects adapt to time-varying spectra. Previous results from ROUSE &GOPHER (1977), CURRY & GOVINDARAJ (1977) and others using the Kalman-filter as a model suggest a view which would make us expect that the subject tries to use the information presented to him in an optimal fashion for adapting his tracking-strategy.

EXPERIMENTS

A number of experiments were done with a standard compensatory tracking-task augmented with a borderline as shown in Fig.3, the controlled system was of zero-order. Disturbances were generally generated from ternary pseudo-random signals, which were filtered digitally with zero-phaseshift filters. The procedure was designed to give us test-signals with good statistical properties for the identification of the nonlinear components of the control-strategy. Trials lasted for 308 secs, a large number of subjects (students of psychology) participated in the experiments, they were trained extensively before experimental sessions.

EXPERIMENT 1. Adaptation to a signal with time-varying bandwidth was investigated, the aim beeing the estimation of the time-course of adaptation. For this purpose time-varying signals were generated. The principle of the construction of signals with time-varying bandwidth is shown in Fig.5. Two signals z1, z2 with a different bandwidth and 1/f characteristics are generated as described above. A weighting-function g is used, which may be deterministic, e.g. a sine-function or pseudo-random, and the time-varying signal is generated as a weighted sum of these two signals according to

$$x(t)=z1(t)*g(t)+z2(t)*(1-g(t))$$

In the experiments two pairs of base-functions (0.1 / 0.28 Hz and 0.2 / 0.8 Hz) and two types of weighting-functions (sine wave of .16 Hz and a three-level pseudo-random function with clock-interval of 1.92 secs were employed.

SUBJECTS. Eight subjects participated in this experiment.

RESULTS. Data were analyzed by estimating the transfer-characteristics of the human in the frequency-domain using standard FFT-procedures (BENDAT & PIERSOL 1971). The nonlinear strategy of subjects should be apparent from the mean-error beeing dependent on the weighting-function of the signal and can be seen from the coherency between weighting-function and output-signal. (Weighting-function and base-functions beeing uncorrelated.)

Coherency for sine- and pseudo-random weighting-functions is shown in Fig.6: In both instances there is a high coherence, higher for the input-signal composed from higher-bandwidth base-functions. This corresponds well with previous results, which have shown the nonlinear component in the control-strategy increasing with bandwidth of the disturbance. This result shows clearly that tracking behaviour adapts to the bandwidth varying in time.

Adaptation of the tracking-strategy should require the subject to collect some information about the statistical properties of the random disturbance. The amount of information required determines the time-lag between variation of input-bandwidth (weighting-function) and mean-error in the output signal. In order to obtain an estimate of the time-course of adaptation the time-delay was estimated from the maximum value of the cross-correlation. The values obtained are shown in Tab.1 for the 8 subjects in the experiment. There is essentially no time lag greater than the lag associated with tracking behaviour itself.

We have to conclude from these results that adaptation of tracking strategy does not occur in the fashion assumed, i.e. does not seem to be based on an estimation of the statistical properties of the input-signal as assumed by a Kalman-filter model.

EXPERIMENT 2. This leaves us with a hypothesis which we attempt to test in another experiment: The change of bandwidth is associated with a change of speed of the input-signal, the change in speed may be used as a cue for changing tracking-strategy.

In order to test this, test-signals were devised which are composed of a low-bandwidth pseudo-random signal and a number of randomly occurring steps of the type shown in Fig.7. These steps have been designed such as not to modify the mean of the signal, and also they represent asymmetric disturbances in the sense that there is a slow movement of the error-pointer in one direction - so slow that it is not discriminable from the low-bandwidth background-noise - and a high-speed movement in the other direction. Response of subjects is analyzed by averaging over steps, as the times of occurrence of these is random in relation to the background-random disturbance, averaging will be a consistent estimate of the response to these steps.

PROCEDURE. A large number of combinations of random-disturbances with different bandwidth and of step-functions with different amplitudes and durations were presented to subjects.

SUBJECTS were 8 students of Psychology.

RESULTS. Responses are shown for various experimental conditions (combinations of bandwidth of random-component and amplitude of step). Under the assumption that the subject adapts his tracking behaviour as a consequence of a high-speed movement of the error-pointer we would have to expect a mean-error following the step as indicated in Fig.7 by a broken line. It is obvious from the results shown that an adaptation of

this type is not observed in our data. However, comparison of responses to steps with positive (towards borderline) and negative (away from borderline) steps (Fig.8) shows a highly asymmetric resonse to the step-functions: In both cases there is no adaptation in the sense expected, rather the response corresponds to the response typical of step-tracking, where the delay of the human processing system generates the particular error-function as observed here.

The response to steps towards/away from the borderline varies in the different amount of under- and overshoot: With a step towards the borderline an overshooting response is a strategy nearer to optimum. In addition to this the time-lag of response is different, with a shorter lag for the error going towards the borderline. The remnant was calculated by using the transfer-function to estimate the linear response to the input signal. The remnant for positive and negative going steps shows also that the difference in response is essentially due to a nonlinear response (Fig.8).

From these results we have to conclude that subjects do not optimize their tracking-strategy on the basis of an estimate of the statistical properties of the disturbance in the tasks investigated. It seems rather as though they use a multiple number of motor-programs and select from these according to input-conditions. COSTELLO (1968) has previously suggested a similar model, where large deviations from target elicit nonlinear responses. The role of the cost-functions in a model of this kind is to modify the conditions eliciting fast responses, which is demonstrated by overshoot depending on the weighting of error in our experiments.

DISCUSSION

The data obtained requires us to reconsider the usefulness of linear-filter analogies (MCRUER & KRENDEL 1974, KLEINMANN, BARON & LEVISON 1970) as models of human manual control behaviour. The data suggest that manual control behaviour under certain conditions at least should be seen as a sequence of discrete motor programs, which would encompass discrete responses and thus be a much more general model. The main difficulty for such a view is the fact that in the tracking task a number of processes are confounded: Motor control, perception, motivation, prediction and optimization. A strategy for investigating these processes should be aimed at separating these diverse processes, which differ by their time-constants and by the theories and methodes appropriate for their scientific investigation.

Also at this moment a methodological framework permitting the integration of knowledge about discrete motor programs and the other processes involved in continuous tracking is not in sight.

Tracking behaviour is limited by a number of factors, thus limits in performance can not be attributed to specific processes. This is not a grave shortcoming as long as interest in tracking is motivated by the attempt to control specific processes with structure similar to the task

investigated experimentally. This state of affairs makes it difficult, however, to integrate models of tracking behaviour with other theories of motor-behaviour.

Considering describing function models we do not conclude that these are incorrect, rather they are limited to a domain of tasks with very specific properties, namely quadratic or at least symmetric cost-functions, a task-structure forcing the subject into a mode with responses tightly feedback-controlled and varying continously (as opposed to discrete) in time and amplitude. Also a disturbance with quite specific properties is implied. Tasks of this type occur in control of vehicles, which has been the domain of prominence of linear-filter theories of human behaviour.

Process- and supervisory control behaviour, on the other hand, is not modelled well with models of this type, whereas on the other hand the function we call "optimization" should be common to both tasks.

Optimization is where the task-specificity of the popular models is most obvious. In principle there are two approaches to optimality: 'Normative' and 'inverse'. Normative optimality is the case where the optimality criteria are specified - which is the case in linear-filter models, where the cost-function is specified as the minimum of the quadratic (RMS-)error. The question commonly asked in conjunction with questions of optimality is: Is, or to what extent, is the human optimal in his performance? Inverse optimality, in contrast, is characterised by the assumption of optimality principles governing control behaviour, but cost-functions maximized beeing unknown.

It is obvious that engineers have been looking at human control behaviour from the normative point of view. What is the point of view of empirical science - psychology, physiology - here? For the motor- and perception processes the criteria for good performance can be specified and can be assumed to be known, eg in the form of the quadratic error criterion. This is not likely to be the case in a general manner for the optimization processes involved, and the empirical point of view would be to identify those cost-functions.

We do not believe that research on tracking behaviour per se will be a major topic in the future, the main difficulty beeing the confounding of different processes. It will be a major aim to determine how these processes can be separated, and which processes can be investigated profitably using the tracking task. The choice of methods in further research on manual control behaviour depends upon the purpose of these investigations: The scientific inquiery of motor behaviour will have to consider the relationship of tracking to discrete motor-programs, the most likely model is to consider tracking a sequence of discrete responses. For problems of man-machine interaction motor performance is not a major limitation any more, due to increasing automation. Here other problems dominate, in the field of driving behaviour, where manual control is still used exclusively, the problem of traffic-safety is not considered to be caused by limitations of motor-performance, but by

incorrectly set points for motor performance. This involves the estimation of very low probabilities for extrema, specifically those leading to accidents. This is also involved in the type of tracking behaviour we have investigated here, and which we call the process of optimization. A unified view of this processes should be helpful in practical problems in a wide range of tasks.

LITERATURE

Bendat,J.S. & A.G.Piersol. Random Data: Analysis & Measurement Procedures. Wiley:N.Y. 1971.

BÖSSER,T. Dangerous behaviour in speed- and distance control of the automobile driver on the motorway. (In German) Research Report, Münster 1981.

BÖSSER,T. Tracking performance with varying error-criteria. Proc. IFAC Conf. on Analysis, Design and Evaluation of Man-Machine Systems, Baden-Baden 1982, 427-434.

BÖSSER,T. Control theory models of tracking behaviour and higher level control functions. Psychological Research 1983 (in print).

COSTELLO,R.G. The surge-model of the well-trained human operator in simple manual control. IEEE Trans. Man-Machine Systems, MMS-9, 1968, 2-9.

CURRY,R.E. & GOVINDARAJ,T. The human as a detector of changes in variance and bandwidth. Proc. 7. Annual Conf. Manual Control, NASA:Moffett Filed 1977, 217-221.

DREYER,W. Längsdynamische Untersuchungen des Regelkreises Fahrer-Fahrzeug. Ber.525, Inst. Fahrzeugtechnik TU Braunschweig 1979.

KLEINMANN,D.L., S.BARON & W.H.LEVISON. An optimal control model of human response. Automation 6, 1970, 357-369.

McRUER,D.T. & E.KRENDEL. Mathematical models of human pilot behaviour. NATO, AGARDograph Nr. 188, 1974.

ROUSE,W.B. & D.GOPHER. Estimation and control theory: Application to modeling human behaviour. Human Factors 19, 1977, 61--618.

Elke Melchior participated in this research. This work was supported by Deutsche Forschungsgemeinschaft.

Fig. 1.
Levels of processing involved in tracking

* **OPTIMIZATION**

  (Allocation of resources to maximize a given cost-function or set of cost-functions)

* **PREDICTION**

  (Compensate for perceptual-motor limitations)

* **PERCEPTUAL-MOTOR PROCESSES**

  Inherent limitations:

  - **Reaction time**
  - **Accuracy**
  - **Organization**

Fig. 2.
Various types of cost-functions

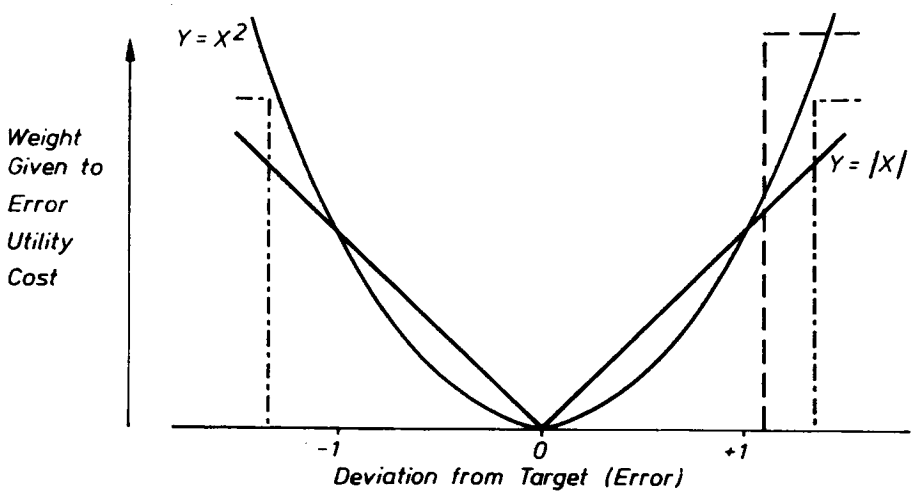


Fig. 3.
Schema of apparatus and display used in experiments

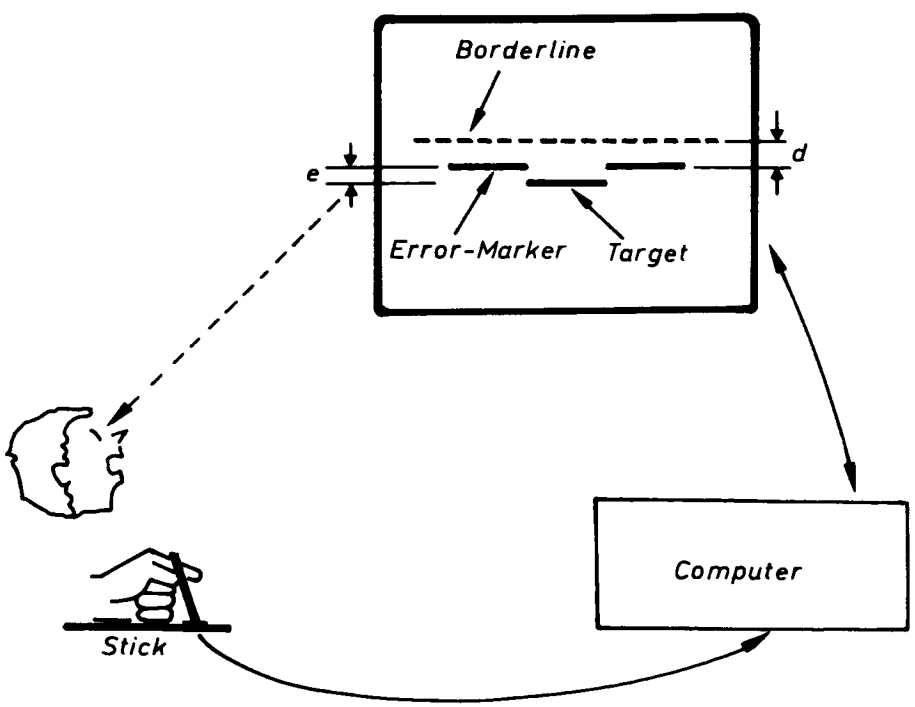


Fig. 4.
Time-histories observed with asymmetric cost-function

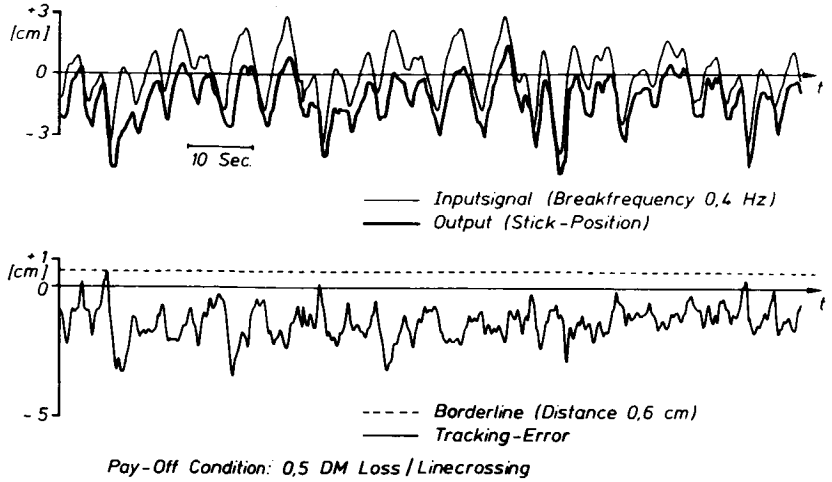


Fig. 5.
Construction of time-varying test-signals

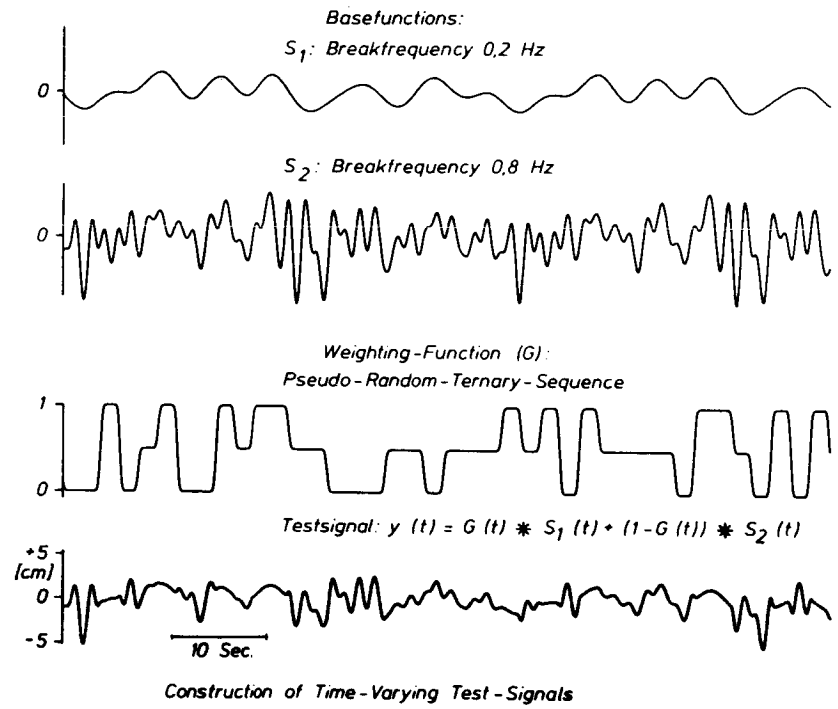
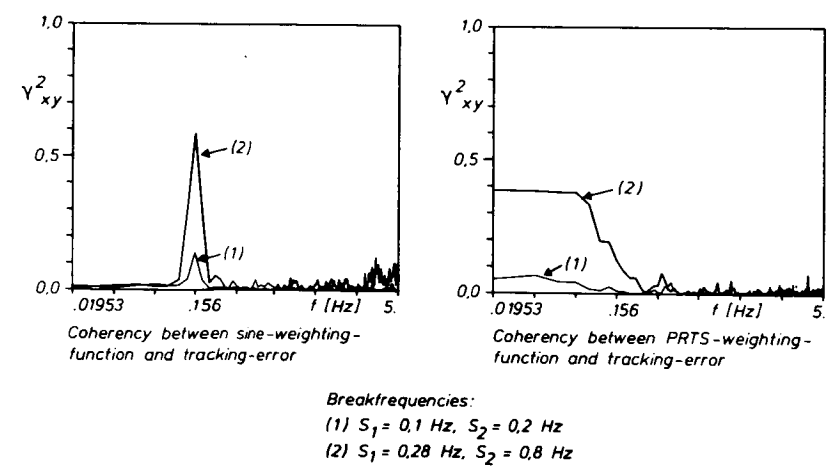
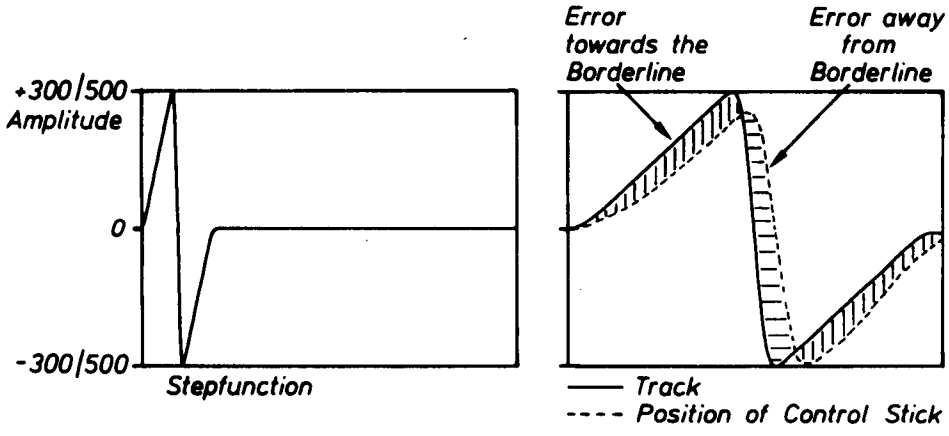


Fig. 6.
Coherences between weighting - function and output signals





BREAKFREQUENCY OF ADDED RANDOM SIGNAL	AMPLITUDE OF STEP-FUNCTION		
	26 STEPS/308 SEC + 500	+ 300	80 STEPS/308 SEC + 300
0.14 HZ	(1)	(2)	(3)
0.28 HZ	(4)	(5)	(6)

Table 2

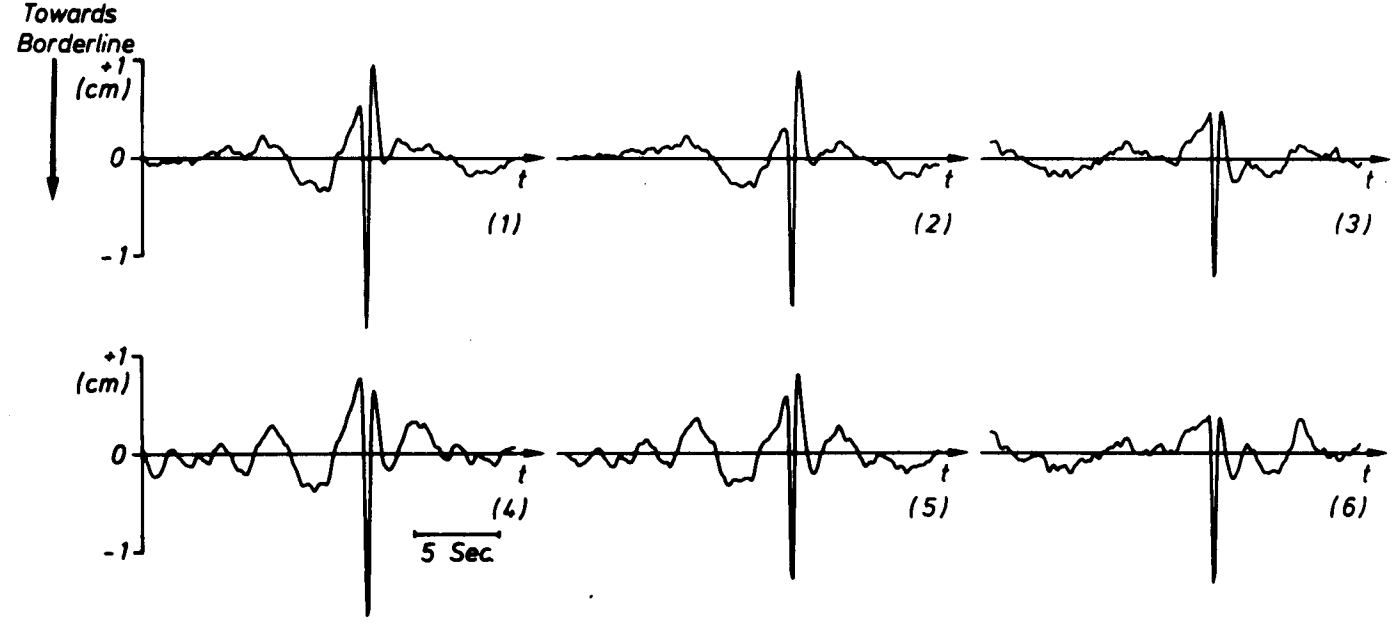


Fig. 7. Stepfunction employed and schematic of relationship between disturbance, stick-position and error. Average error observed with conditions as summarized in Table 2.

Tab. 1.
Time-lags between weighting-function and output signal

Subject	PTRS-Weighting Function		Sine-Weighting Function	
	(1)	(2)	(1)	(2)
1	0.9	0.6	0.6	0.6
2	0.9	0.7	1.1	0.8
3	0.6	0.6	0.6	0.8
4	0.8	0.7	0.8	0.8
5	0.9	0.8	0.8	0.8
6	0.9	0.9	0.8	0.8
7	0.7	0.9	0.9	0.9
8	0.8	0.9	1.1	1.1

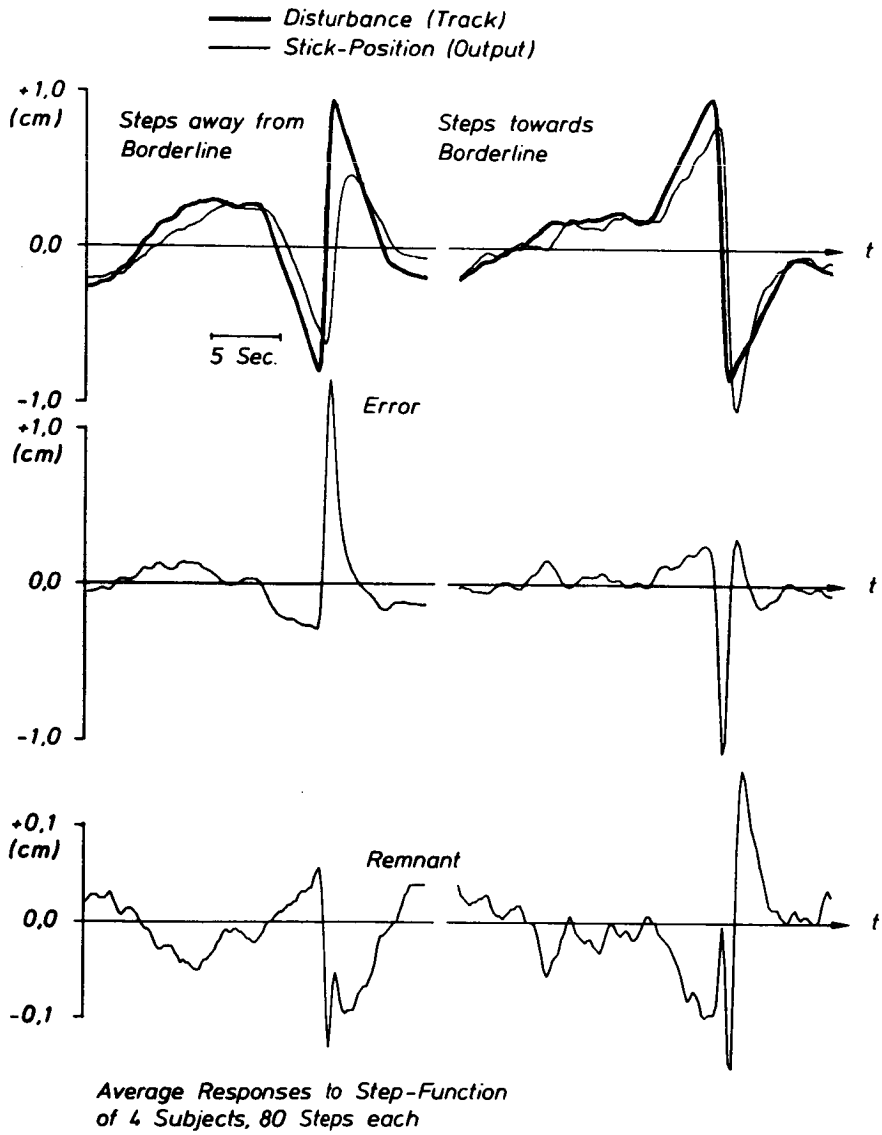


Fig. 8. Average response to steps away from / towards borderline. Remnant is calculated after subtracting output predicted by use of weighting-function.

PREVIEW-FOLLOWER METHOD FOR MODELING CLOSED-LOOP VEHICLE DIRECTIONAL CONTROL

Konghui Guo
The First Motor Works of China
The Changchun Motor Vehicle Research Institute
Changchun, China

Paul S. Fancher
Transportation Research Institute
The University of Michigan
2901 Baxter Road
Ann Arbor, Michigan 48109

PREVIEW-FOLLOWER METHOD FOR MODELING CLOSED-LOOP VEHICLE DIRECTIONAL CONTROL

Konghui Guo
Paul S. Fancher

ABSTRACT

A preview-follower concept is presented for modeling the closed-loop system consisting of a roadway vehicle controlled by a driver.

A theoretical technique for identifying the driver/vehicle system and optimal driver parameters, e.g., feedback constants, open-loop gain, and derivative time constant, is presented. It is found that a satisfactory range of preview time can be chosen on the basis of a dominant eigenvalue analysis. Theoretical results are used to demonstrate the importance of the system parameters appearing in an "ideal follower" model.

Computer simulation results are compared with test measurements and demonstrate that simulation predictions show satisfactory agreement with test data for a heavy truck/experienced driver combination.

FOREWORD

This informal paper describes the initial efforts of a research investigation undertaken by Mr. Guo while he was a Visiting Scientist at The University of Michigan Transportation Research Institute. It is anticipated that upon his return to the Changchun Motor Vehicle Research Institute in Changchun, China, Mr. Guo will be pursuing matters related to the influences of driver characteristics on vehicle control.

The ideas presented herein are speculative and based upon Mr. Guo's thorough understanding of vehicle dynamics plus his perception of the process of learning to drive. In this sense, the results of these initial efforts may be classified as establishing reasonable hypotheses rather than as providing proven findings. Nevertheless, the hypothesized model of driver control appears to be useful for aiding the vehicle dynamicist in examining questions related to the directional performance of the driver-vehicle system in obstacle-avoidance maneuvers.

Paul S. Fancher
Transportation Research Institute
The University of Michigan

INTRODUCTION

Problems pertaining to the directional control of vehicles have been studied for several decades. Very complete vehicle models with many degrees of freedom have been built to evaluate directional performance. However, it appears that a basic obstacle in utilizing the results from vehicle models has been a lack of understanding of driver control characteristics.

As a result, various researchers have been working on driver models and road/driver/vehicle closed-loop systems. In this regard, McRuer and his colleagues have done a substantial amount of work (e.g., see [1,2]). In pioneering work, the preview capability of the driver has been avoided by treating responses to external disturbances rather than responses to path inputs. Recently, several kinds of so-called "optimal preview closed-loop models" have been proposed. To simulate driver behavior, Reddy and Ellis [3] offered a method that instantaneously freezes the position of the vehicle, and proceeds with driving forward for a certain time increment (look-ahead time) to see if the error is acceptable. An acceptable steering-wheel angle which satisfies a defined error threshold is obtained by an incremental search strategy. Continuing in this step-by-step manner, driver behavior is simulated.

At about the same time, MacAdam [4,5] presented a different way to solve the problem, even though the basic idea is somewhat similar. MacAdam's method seems more powerful and more accurate than Reddy and Ellis's approach because a direct optimization method is used to determine steering-wheel angle according to path input, current vehicle state, and current steering input.

Recently, two new approaches, called the "Optimal Acceleration Strategy" and the "Optimal Preview-Follower Concept," have been used to study closed-loop driver control [8]. In these new approaches, the modeling of driver behavior is based on the standard hypothesis that the steering behavior of the driver is always aimed at minimizing the error between the desired course and the trajectory of the vehicle. In reference [8], it is shown that the "Optimal Acceleration Strategy" leads to the same result as that of the "Optimal Preview-Follower Concept." This paper provides a detailed discussion of the "Optimal Preview-Follower Concept."

Corresponding to the natural sequence of accumulating experience by a driver, modeling of driver behavior at low vehicle speed with a single-point preview assumption is studied first, after which the "Optimal Preview-Follower Concept" is introduced. Then, this concept is extended and applied to general cases. It is shown that predicted results are in satisfactory agreement with test data.

1. A QUASI-STATIC PREVIEW-FOLLOWER MODEL

It is conceivable that the steering behavior of a driver is formed and improved gradually as the driver accumulates experience. Quasi-static operation seems to be the basic experience in forming driver characteristics. The reasons for this are:

1) Drivers get their first experiences from low-speed driving where almost no vehicle dynamics are involved.

2) Drivers have much more practice at turning slowly than at turning rapidly.

In these cases, the curvature of the trajectory is primarily dependent upon steering-wheel angle, viz.,

$$\frac{1}{R} = \frac{\delta_{sw}}{iL} \tag{1.1}$$

or

$$\ddot{y} = \frac{V^2}{R} = \frac{\delta_{sw}}{iL} V^2 \tag{1.2}$$

where

δ_{sw} - steering-wheel angle

L - wheelbase

y - lateral position of the vehicle

V - vehicle velocity

i - steering ratio

If at moment, t, (see Fig. 1.1) the vehicle has a current state $y = y(t)$ (and $\dot{y} = \dot{y}(t)$), and if it is assumed that the driver is (1) looking ahead to the path, f(t), at a distance, d, in front of the vehicle (single-point preview assumption) and (2) applying a value of steering-wheel angle, δ^*_{sw}, which will generate a lateral acceleration, $\ddot{y}(t)^* = \delta^*_{sw} \cdot V^2/iL$, then, after a period of look-ahead time, $T = d/V$, the new position of the vehicle is assumed to be given by the following equation:

$$y(t+T) = y(t) + T\dot{y}(t) + \frac{T^2}{2}\ddot{y}(t)^* \tag{1.3}$$

Furthermore, if it is assumed that the driver expects that the vehicle will follow the desired path exactly, then the driver expects that

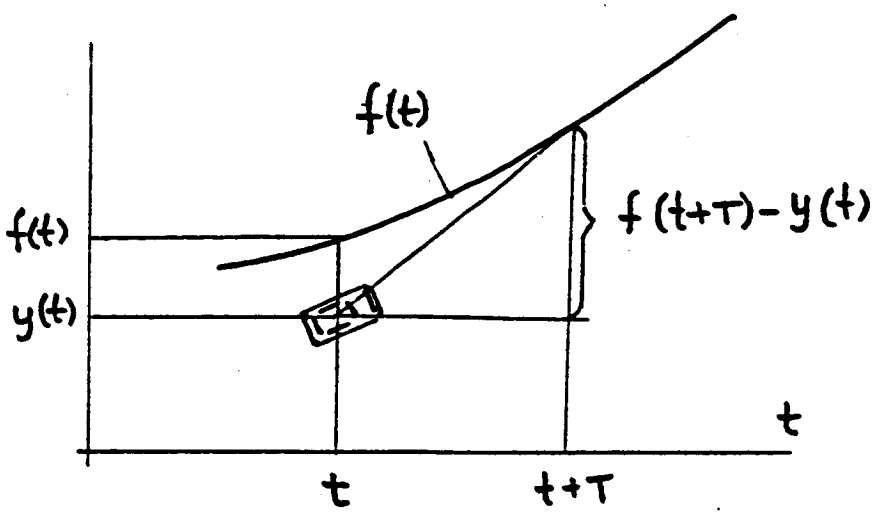


Figure 1.1. Vehicle state and preview input.

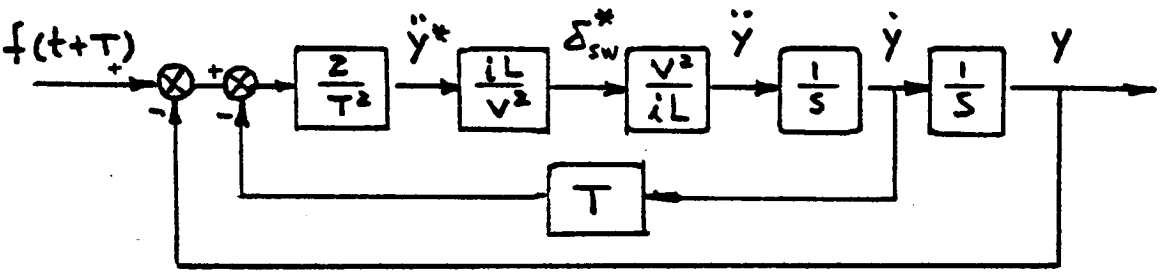


Figure 1.2. A block diagram of the closed-loop system for quasi-static directional control.

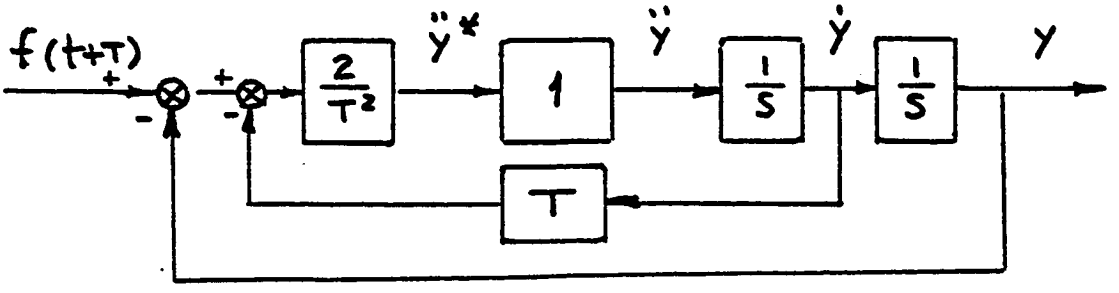


Figure 1.3. A simplified block diagram.

$$y(t+T) = f(t+T) \tag{1.4}$$

This is called the "single-point preview assumption" (more general preview windows will be discussed later).

Equations (1.1), (1.2), (1.3), and (1.4) yield

$$\ddot{y}(t)^* = \frac{2}{T^2}\left(f(t+T) - y(t) - T\dot{y}(t)\right) \tag{1.5}$$

or

$$\delta^*_{sw}(t) = \frac{2iL}{V^2T^2}\left(f(t+T) - y(t) - T\dot{y}(t)\right) \tag{1.5a}$$

This means that, for minimizing the error between the desired path and the vehicle trajectory, the driver endeavors to operate the vehicle according to Equation (1.5a).

The block diagram of this system is shown in Figure 1.2. Note that the two blocks in the middle of this block diagram (i.e., the vehicle characteristics, $\ddot{y}/\delta_{sw} = V^2/iL$, and the driver estimating function, $\delta^*_{sw}/\ddot{y}^* = iL/V^2$, can be cancelled and the block diagram can be simplified as shown in Figure 1.3.

The transfer function of the closed-loop system is

$$\frac{y}{f}(s) = \frac{e^{Ts}}{\frac{T^2}{2}s^2 + Ts + 1} \tag{1.6}$$

This is a second-order system with a lead time, T. The undamped natural frequency is $\omega_o = \sqrt{2/T}$, and the damping factor is $\zeta = \sqrt{2}/2 = 0.707$. Figure 1.4 shows the step input responses for different values of look-ahead (lead) time, T. The overshoot of the responses have the same value, 0.0425, however, the response errors are different for various look-ahead times. It is seen that, for this single-point preview model, the closed-loop system has a satisfactory tracking performance and it seems that tracking performance will be better for shorter look-ahead times.

This conclusion is true for the system without considering driver delays and dynamics effects of the vehicle. However, if a driver delay or the dynamic effect of the vehicle ("vehicle delay") is included, the rule for look-ahead time will no longer be "the shorter the better." For systems including time delays (either driver's or vehicle's), shorter look-ahead times will cause more oscillation of the system. In fact, there exists a lower limit of look-ahead time below which the system will become unstable.

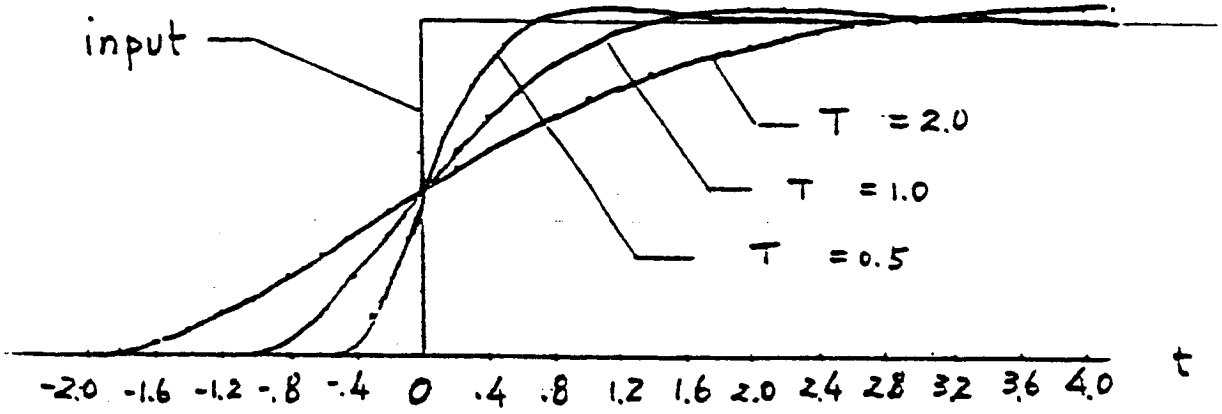


Figure 1.4. Responses of the simplified system to a step path input.

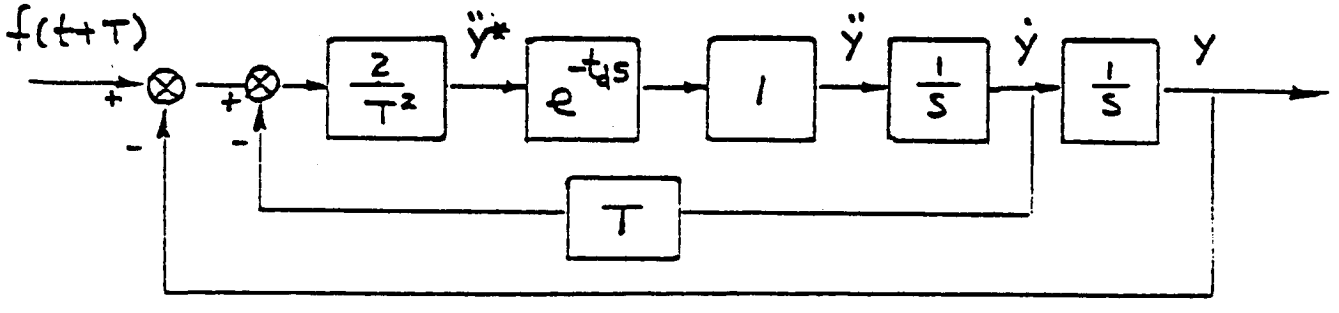


Figure 1.5. Quasi-static system with driver response delay.

Suppose that a driver delay, t_d, is included in the system. The block diagram of this system is shown in Figure 1.5. The transfer function can be expressed as follows:

$$\frac{y}{f}(s) = \frac{e^{Ts}}{\frac{T^2}{2} e^{t_d s} s^2 + Ts + 1} \tag{1.7}$$

Using a first-order approximation for $e^{t_d s}$, i.e.,

$$e^{t_d s} \simeq (1 + \frac{T}{2}s)/(1 - \frac{T}{2}s) \tag{1.8}$$

and

$$\frac{y}{f}(s) = \frac{e^{Ts}(1 - \frac{t_d}{2} s)}{\frac{T^2}{4} t_d s^3 + (\frac{T^2}{2} - \frac{t_d}{2} T)s^2 + (T - \frac{t_d}{2})s+1} \tag{1.9}$$

the characteristic equation of the system is

$$A_3 s^3 + A_2 s^2 + A_1 s + 1 = 0 \tag{1.10}$$

where

$$A_1 = T - \frac{t_d}{2} , \quad A_2 = \frac{T^2}{2} - \frac{t_d}{2} , \quad A_3 = \frac{T^2}{4} t_d$$

Using Routh's criterion of stability, the stable conditions are:

$$\left.\begin{array}{l} 1, \quad A_1 > 0 = T > t_d/2 \\ 2, \quad A_2 > 0 = T > t_d \\ 3, \quad A_1 A_2 > A_3 = T > 1.707\ t_d \end{array}\right\} \tag{1.11}$$

Obviously, condition 3 is dominant. The eigenvalues of the system with t_d = 0.2, 0.4 sec. and different look-ahead times are computed and shown in Figure 1.6 (the conjugate root and a nondominant real root are not shown). Figure 1.6 shows that, as the look-ahead time, T, gets shorter, the system will become more oscillatory; and when $T < 1.707\ t_d$, the system will become unstable.

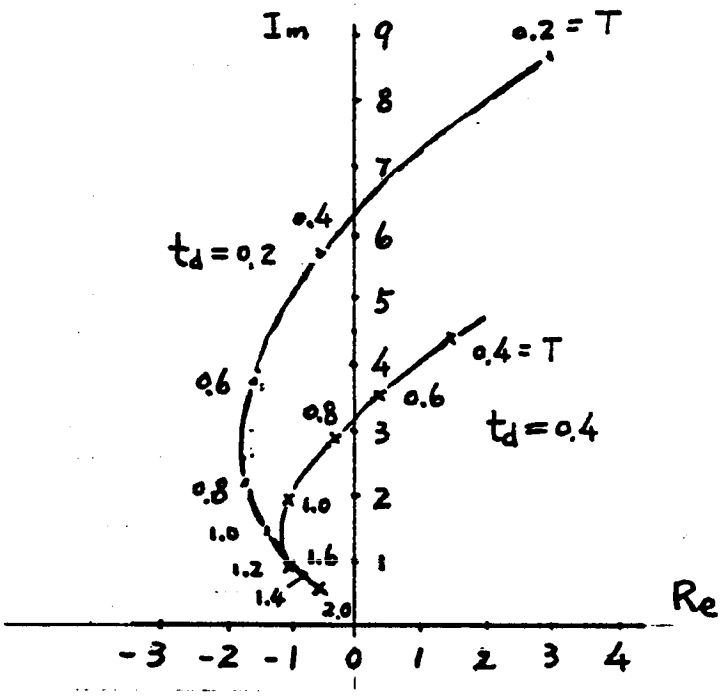


Figure 1.6. Loci of the dominant root of a quasi-static system with various driver delay and look-ahead times.

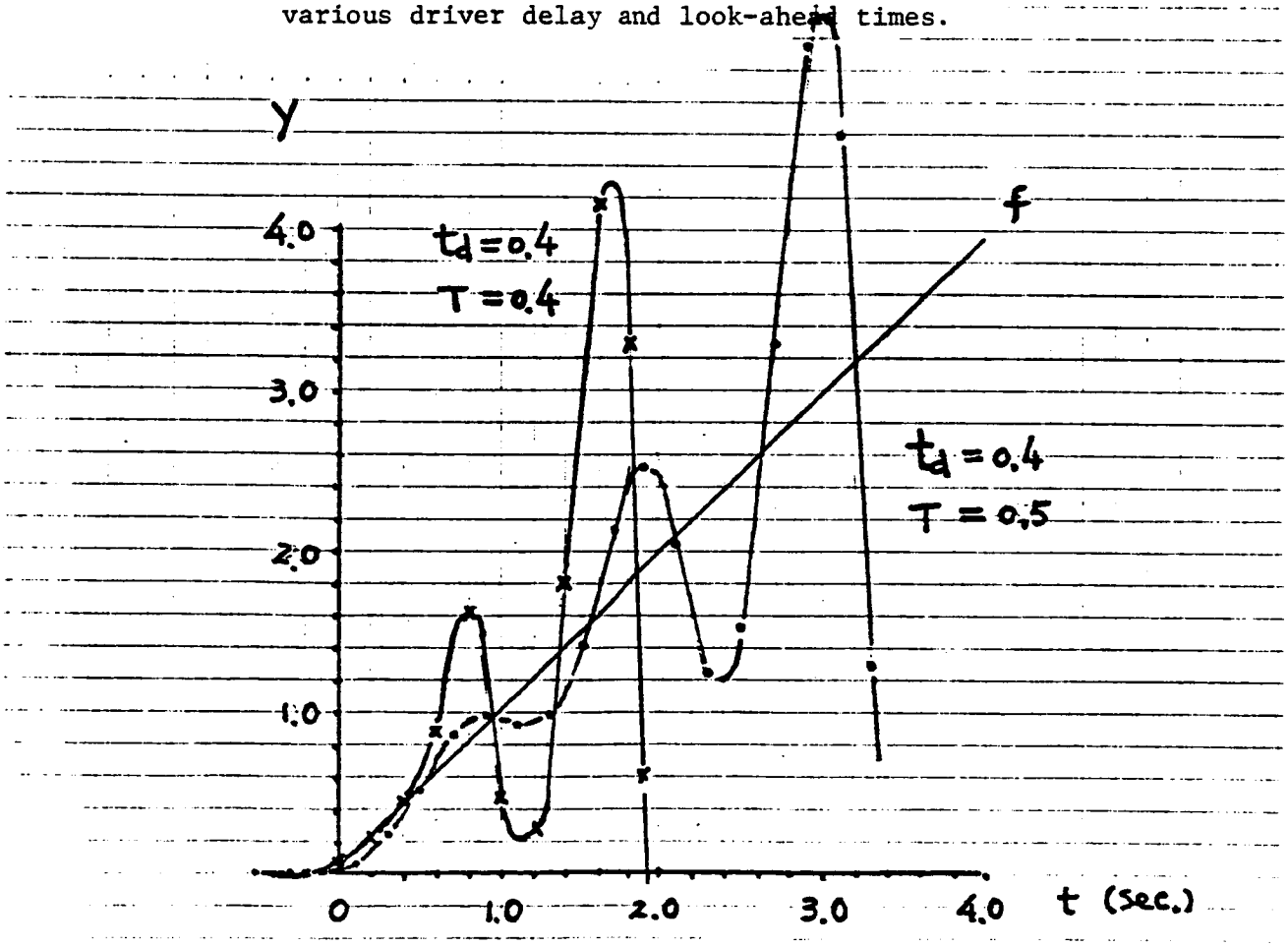


Figure 1.7. Responses of the system to a ramp path input when the stable condition (1.11) is not satisfied.

It is also seen in Figure 1.6 that the eigenvalues of the system with $t_d = 0.4$ sec. are located on the right side of those with $t_d = 0.2$ sec. Hence, larger time delays correspond to reduced system stability. Simulation results showing unstable ramp input responses are presented in Figure 1.7.

Notice that Equation (1.2) holds either for low vehicle speed conditions or for high vehicle speed with low frequency (smooth turning) operation, if the vehicle itself is stable (that is, if a steady-state gain exists). Since a driver accumulates experience primarily in a quasi-static control mode, it is conceivable (further arguments will be presented later) that Figure 1.5 and Equation (1.7) provide a basic structure for modeling closed-loop directional control. Nevertheless, in a general case where vehicle dynamics are involved, the block showing unit gain (located near the center of Fig. 1.5) will be replaced by a vehicle transfer function, $\ddot{y}/\delta_{sw}$ (s), multiplied by a driver estimating/correcting function, $\tau(s)$, which will be chosen to represent a reasonable estimate of the driver's ability to close the control loop.

2. DEFINITION OF THE IDEAL PREVIEW-FOLLOWER CONCEPT

It is interesting to note from Equation (1.5) that the trajectory of the vehicle, y(t), is not exactly the path input, f(t), because $(1 + Ts + T^2s^2/2)$, the denominator of the transfer function, is not exactly equal to e^{Ts}. Expanding e^{Ts} by using Taylor's series,

$$e^{Ts} = 1 + Ts + \frac{T^2}{2} s^2 + \frac{T^3}{3} s^3 + \text{---} \qquad (2.1)$$

one can find that the denominator of the transfer function $(1 + Ts + T^2s^2/2)$ is just the first three terms of the expansion of e^{Ts}, and the tracking error is caused by the terms of s^3 and higher order.

Now suppose that the driver tries to perform a steering input such that the vehicle has a cornering motion in which $\dddot{y}(t+\tau) \equiv \dddot{y}^*(t)$ is a constant during the interval $\tau = 0$ to $\tau = T$, so that,

$$y(t+T) = \int_0^T \int_0^T \int_0^T \dddot{y}(t)\, d^3$$

$$= y(t) + T\dot{y}(t) + \frac{T^2}{2} \ddot{y}(t) + \frac{T^3}{3!} \dddot{y}(t) \qquad (2.2)$$

If the driver expects that the vehicle will be driven so that $y(t+T) = f(t+T)$ exactly, then

$$\frac{T^3}{3!}\dddot{y}(t)^* = f(t+T) - y(t) - T\dot{y}(t) - \frac{T^2}{2}\ddot{y}(t) \tag{2.3}$$

The block diagram of the system will be as shown in Figure 2.1. The transfer function can be written as

$$\frac{y}{f}(s) = e^{Ts}\frac{1}{1 + Ts + \frac{T^2}{2}s^2 + \frac{T^3}{3!}s^3} \tag{2.4}$$

In fact, the first multiplier e^{Ts} is a transfer function describing the preview effect. (This transfer function may take a different form if the driver is not looking ahead at just a single point.) The other multiplier $(1 + Ts + T^2s^2/2 + T^3s^3/3!)^{-1}$ is a transfer function for a "following effect" (which again may take a different form if a different preview function is employed).

Generally, the transfer function of the whole system can be considered as a multiplication of two parts:

$$\frac{y}{f}(s) = P(s) \cdot F(s) \tag{2.5}$$

where P(s) is the transfer function of the previewer, which has a Taylor's expansion as

$$P(s) = P_o + P_1 s + P_2 s^2 + P_3 s^3 + \text{---} \tag{2.6}$$

and F(s) is the transfer function of the follower, the inverse of which has a Taylor's expansion as

$$F(s)^{-1} = a_o P_o + a_1 P_1 s + a_2 P_2 s^2 + a_3 P_3 s^3 + \text{---} \tag{2.7}$$

where a_o, a_1, a_2 --- are constants depending on the dynamics of the vehicle and the parameters of the driver.

The n^{th} order follower is defined as follows. If the transfer function of a closed-loop dynamic system can be expressed by Equations (2.5), (2.6), and (2.7) such that

$$\begin{aligned} a_i &= 1, \quad \text{for} \quad i = 0,1,2,\text{---}n \\ a_i &= 0, \quad \text{for} \quad i > n \quad , \end{aligned} \tag{2.8}$$

then the system is called an n^{th} order follower system. One might conjecture that the higher the order of the follower system, the less the tracking error will be. However, in fact, this is not always true. A follower of too high an order may cause an overly oscillatory response and corresponding tracking error.

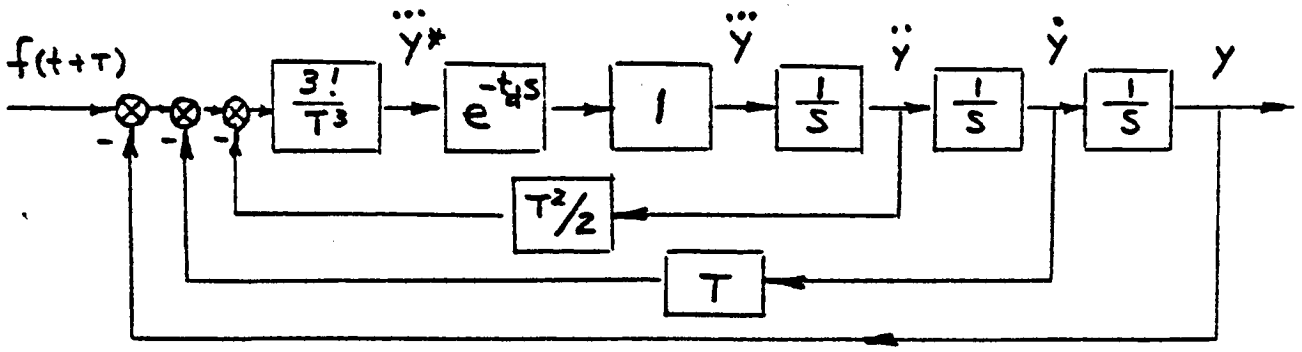


Figure 2.1. A block diagram of the closed-loop directional control model based on the "optimal $\dddot{y}$ strategy."

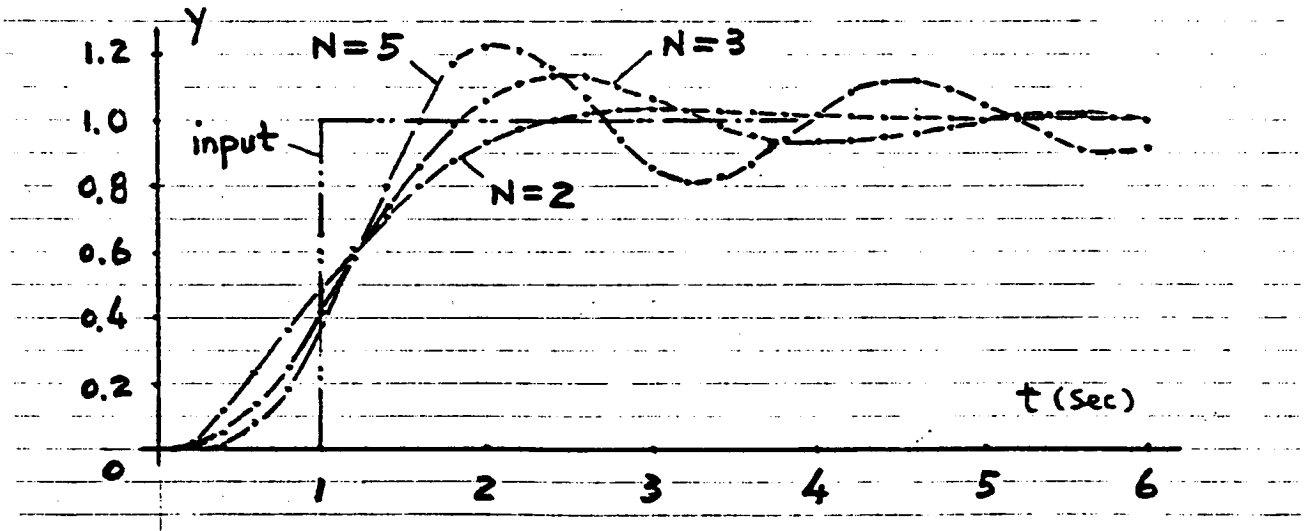


Figure 2.2. Step input responses of ideal followers of different order.

Figure 2.2 shows the step input responses of different order followers for single-point preview control systems, illustrating that the second-order follower and the third-order follower have tracking performance preferable to higher-order followers for the requirements of both damping and response error.

Figures 2.3(a) and (b) show computer simulation results for the total square error (TSE) corresponding to six seconds of the unit ramp input response for different time delays, t_d, and look-ahead times (T = 1.0 sec. and T = 2.0 sec.). It is seen from Figure 2.3 that, from the viewpoint of TSE, followers of higher than fourth order appear to have no advantage for tracking tasks (let alone, their low damping performance).

From the viewpoint of steady-state error, the second-order follower system can guarantee the steady error to be zero if the path input, f(t), has the feature that the $\lim_{t\to\infty} \dddot{f}(t) = 0$. However, for a path input with $\lim_{t\to\infty} \ddot{f}(t) \neq 0$, the steady-state error will not be zero unless the system has a higher-order follower (see Appendix A of Reference [8]).

Experience shows that (a) adequate damping in system response and (b) reduced response error (including steady-state error and dynamic error) are desirable for good closed-loop directional control. Hence, it seems reasonable to assume that characteristics between second- and third-order followers are preferable for most directional control tasks.

Therefore, it is assumed that an experienced driver will train himself to be as close as possible to the ideal preview-follower system in order to meet the following requirements relating to Equations (2.5)-(2.7):

$$\begin{cases} a_o = a_1 = a_2 = 1 \\ 0 \leq a_3 \leq 1 \\ a_4 = a_5 = a_6 = \text{---} = 0 \end{cases} \tag{2.9}$$

Condition (2.9) is called the "ideal follower condition."

For a single-point preview control system, according to the ideal follower condition, the ideal follower function is

$$F(s) = \left(1 + Ts + \frac{T^2}{2} s^2 + a_3 \frac{T^3}{3!} s^3\right)^{-1} \tag{2.10}$$

and the transfer function of the ideal closed-loop system is

$$\frac{y}{f}(s) = \frac{e^{Ts}}{1 + Ts + \frac{T^2}{2} s^2 + a_3 \frac{T^3}{3!} s^3} \tag{2.11}$$

where a_3 is chosen between 0 and 1.

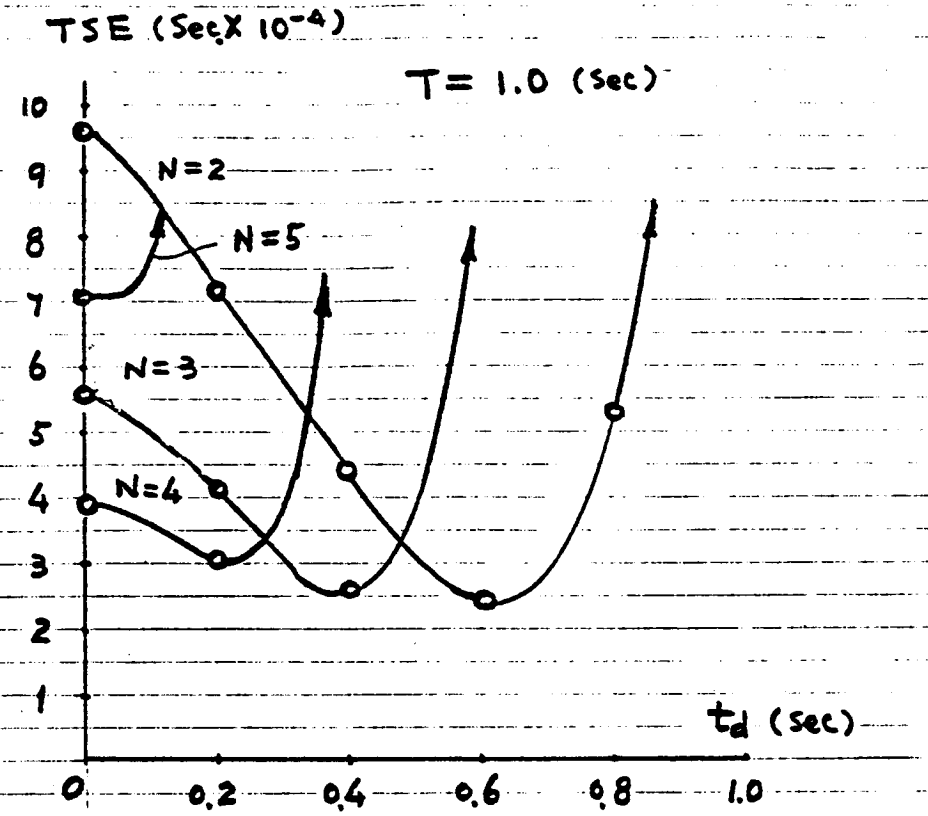


Figure 2.3(a). The total square error of ideal followers of different order vs. delay time for look-ahead time T = 1.0 sec.

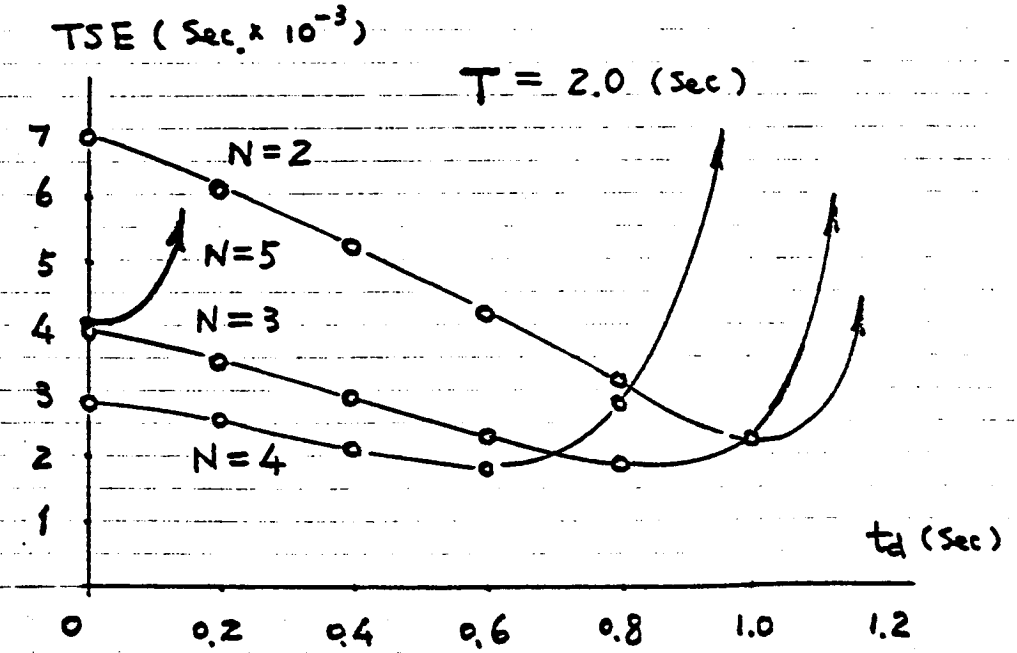


Figure 2.3(b). The total square error of ideal followers of different order vs. delay time for look-ahead time T = 2.0 sec.

For a generalized system with preview function

$$P(s) = (P_o + P_1 s + P_2 s^2 + P_3 s^3 + ---) ,$$

the ideal follower function is

$$F(s) = (P_o + P_1 s + P_2 s^2 + a_3 P_3 s^3)^{-1} \tag{2.12}$$

and the transfer function of the ideal closed-loop system is

$$\frac{y}{f}(s) = P(s) \cdot F(s) = \frac{P(s)}{P_o + P_1 s + P_2 s^2 + a_3 P_3 s^3} \tag{2.13}$$

where, according to the ideal follower condition (2.9), a_3 is chosen between 0 and 1, implying that the system has characteristics between second- and third-order followers.

Note that the ideal follower function is always synthesized as an inverse of the first few terms (low-order terms) of the preview function. The higher-order terms, which represent high-frequency responses, are neglected and the low-frequency responses are emphasized. This is reasonable because, in fact, the motor vehicle system is always a low-pass filter. It is not only impossible, but also not necessary to follow the high-frequency components of the path input.

It is shown in Reference [8] that the ideal follower concept is analogous to an acceleration strategy concept. The optimal $\ddot{y}$ strategy is comparable to the second-order optimal follower and the optimal $\dddot{y}$ strategy is comparable to the third-order optimal follower.

3. MODELING THE CLOSED-LOOP SYSTEM INCLUDING DRIVER AND VEHICLE

The concept of the ideal preview-follower has been described in the previous section. Based on this concept, the issue of how to construct a preview-follower model including both driver and vehicle will be discussed here.

Before illustrating the method for modeling a "real" closed-loop system, it is helpful to see how an "ideal" preview-follower system can be constructed. In general, as was mentioned before, an ideal preview-follower system is illustrated as in Figure (3.1)

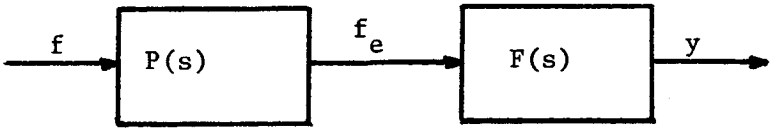


Figure 3.1. The basic configuration of a preview control system.

where $$P(s) = \frac{f_e(s)}{f(s)} = P_o + P_1 s + P_2 s^2 + P_3 s^3 + \text{---} \qquad (3.1)$$

$$F(s) = \frac{y(s)}{f_e(s)} = P_o + P_1 s + P_2 s^2 + a_3 P_3 s^3 \qquad (3.2)$$

where P_o, P_1, P_2 --- are constants dependent on the preview function

and a_3 is a constant between 0 and 1.

Since $\dot{y}(s) = sy(s)$ and $\ddot{y}(s) = s^2 y(s)$, Equation (3.2) can be written as

$$f_e(s) = P_o y(s) + P_1 \dot{y}(s) + (P_2 + aP_3 s)\ddot{y}(s)$$

or,

$$\ddot{y}(s) = (f_e(s) - P_o y(s) - P_1 \dot{y}(s)) \cdot \frac{1}{P_2 + a_3 P_3 s} \qquad (3.3)$$

Hereafter, we can consider $\ddot{y}$ as an output of a subsystem $1/P_2 + a_3 P_3 s$ resulting from an input, f_e, and the feedback from y and $\dot{y}$, the current state of the vehicle. The block diagram of the ideal preview-follower system can be illustrated as shown in Figure 3.2.

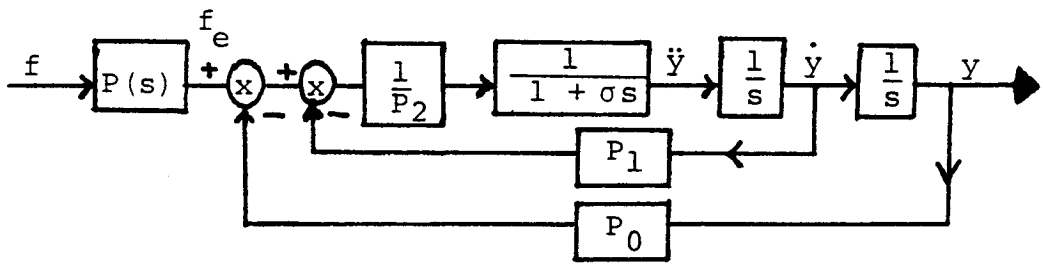


Figure 3.2. A block diagram of the ideal preview-follower system.

In Figure 3.2, $$\sigma = a_3 P_3 / P_2 \qquad (3.4)$$

Observe that, for an ideal preview-follower system, the coefficients of the feedback signals from the vehicle's current state, $\begin{Bmatrix} y \\ \dot{y} \end{Bmatrix}$, are just equal to the corresponding coefficients in a Taylor's Series expansion of the preview function; and the steady-state gain of the $\ddot{y}$ subsystem is equal to $1/P_2$, the inverse of the second-order coefficient in P(s).

In a real driver/vehicle system, $\ddot{y}$ is an output of the subsystem $\ddot{y}/\delta_{sw}(s)$ (the vehicle's transfer function) with a steering-wheel angle input, δ_{sw}, which is a result of the driver's estimation/correction, $c(s)$, according to the path input, f_e, and the current state feedback.

Assuming that there is always a delay, t_d, in the driver's response and a time constant, t_n, in handling the steering wheel, the block diagram of the preview-follower system would be expressed as in Figure 3.3.

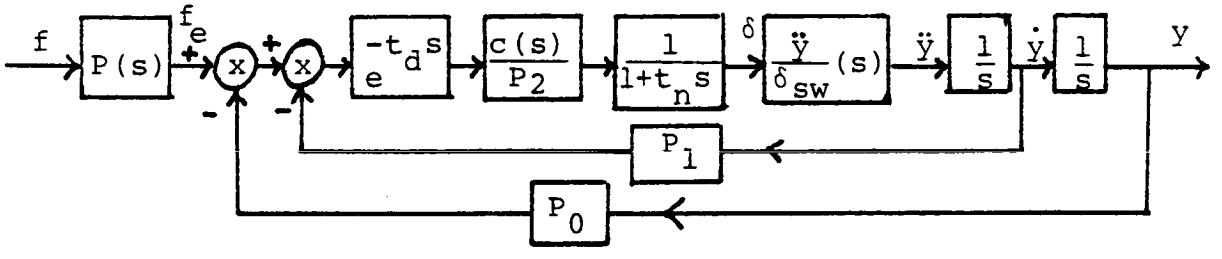


Figure 3.3. A block diagram of a real closed-loop system.

Comparing Figure 3.3 with Figure 3.2, it is clear that the driver's estimation/correction function, $c(s)$, could be chosen to meet the following condition:

$$e^{-t_d s} c(s)^* \frac{1}{1 + t_n s} \cdot \frac{\ddot{y}}{\delta_{sw}}(s) = \frac{1}{1 + \sigma s}$$

or

$$c(s)^* = \frac{1 + t_n s}{1 + \sigma s} e^{t_d s} / \frac{\ddot{y}}{\delta_{sw}}(s) \qquad (3.5)$$

where, generally, $\ddot{y}/\delta_{sw}(s)$ can be expressed in the following form

$$\frac{\ddot{y}}{\delta_{sw}}(s) = Gay \frac{1 + T_{y1} s + T_{y2} s^2 + \ldots + T_{ym} s^m}{1 + T_1 s + T_2 s^2 + \ldots + T_n s^n} \qquad (3.6)$$

where Gay is the steady-state gain

$T_{yi} (i = 1,2,\ldots m)$, $T_j (j = 1,2,\ldots n)$ are constants.

However, it is probably impossible to meet condition (3.5) for the following reasons:

a) The driver is not able to forecast the vehicle state, $y(t+t_d)$ and $\dot{y}(t+t_d)$, as exactly implied by the operator, $e^{t_d s}$.

b) The inverse transfer function, $1/\frac{\ddot{y}}{\delta_{sw}}(s)$, is too difficult for a driver to perform.

Expanding Equation (3.5) by using Taylor's series,

$$\left.\begin{aligned} c(s)^* &= c_o^* + c_1^* s + c_2^* s^2 + \text{---} \\ &\text{where} \\ c_o^* &= \frac{1}{Gay} \\ c_1^* &= \frac{1}{Gay}(T_1 + t_n + t_d - T_{y1} - \sigma) \\ &\text{- - - - -} \end{aligned}\right\} \qquad (3.7)$$

It is conceivable that for a learning driver, who frequently gets experience from operating at steady state or quasi-steady state, the predominant estimation is

$$c(0) = \lim_{s\to 0} c(s)^* = \frac{1}{Gay} \qquad (3.8)$$

For a more experienced driver, who is able to detect the derivative of the path and feedback input, the estimation/correction function could have the following form:

$$\left.\begin{aligned} c(s) &= c_o^* + C_1^* s = \frac{1}{Gay}(1 + T_c s) \\ &\text{where} \\ T_c &= T_1 + t_n + t_d - T_{y1} - \sigma \end{aligned}\right\} \qquad (3.9)$$

and σ is determined by Equation (3.4).

Since the high frequency contents in the δ_{sw} command have little effect on cornering response and they are difficult for a driver to perform, it is assumed that drivers tend to neglect the higher-order terms in Equation (3.7) and prefer to train themselves to have an estimation/correction ability as expressed in Equation (3.9), which is equal to the truncated form of Equation (3.5) or Equation (3.7).

Hereinafter, two important types of preview control are discussed as applications of the method described above.

1) Single-Point Preview Control. As mentioned before in the case of single-point preview control, at the moment, t, the driver receives the effective path input $f_e(t) = f(t+T)$ instead of $f(t)$. The preview function is

$$P(s) = \frac{f_e(s)}{f(s)} = e^{Ts} \tag{3.10}$$

Its Taylor expansion is

$$P(s) = P_o + P_1 s + P_2 s^2 + P_3 s^3 + \text{---}$$

where
$$\begin{aligned} P_o &= 1 \\ P_1 &= T \\ P_2 &= \frac{T^2}{2} \\ P_3 &= \frac{T^3}{3!} \end{aligned} \tag{3.11}$$

According to Equation (3.4),

$$\sigma = a_3 P_3 / P_2 = a_3 T/3 \tag{3.12}$$

According to Equation (3.9),

$$c(s) = \frac{1}{Gay}(1 + T_c s)$$

$$T_c = T_1 + t_h + t_d - T_{y1} - a_3 T/3 \tag{3.13}$$

where a_3 is chosen between 0 and 1. For $a_3 = 0$, one obtains a second-order optimal follower system for single-point preview control. Figures 3.4(a) and (b) are examples of simulation results for lane-change maneuvers with different look-ahead times (T), where

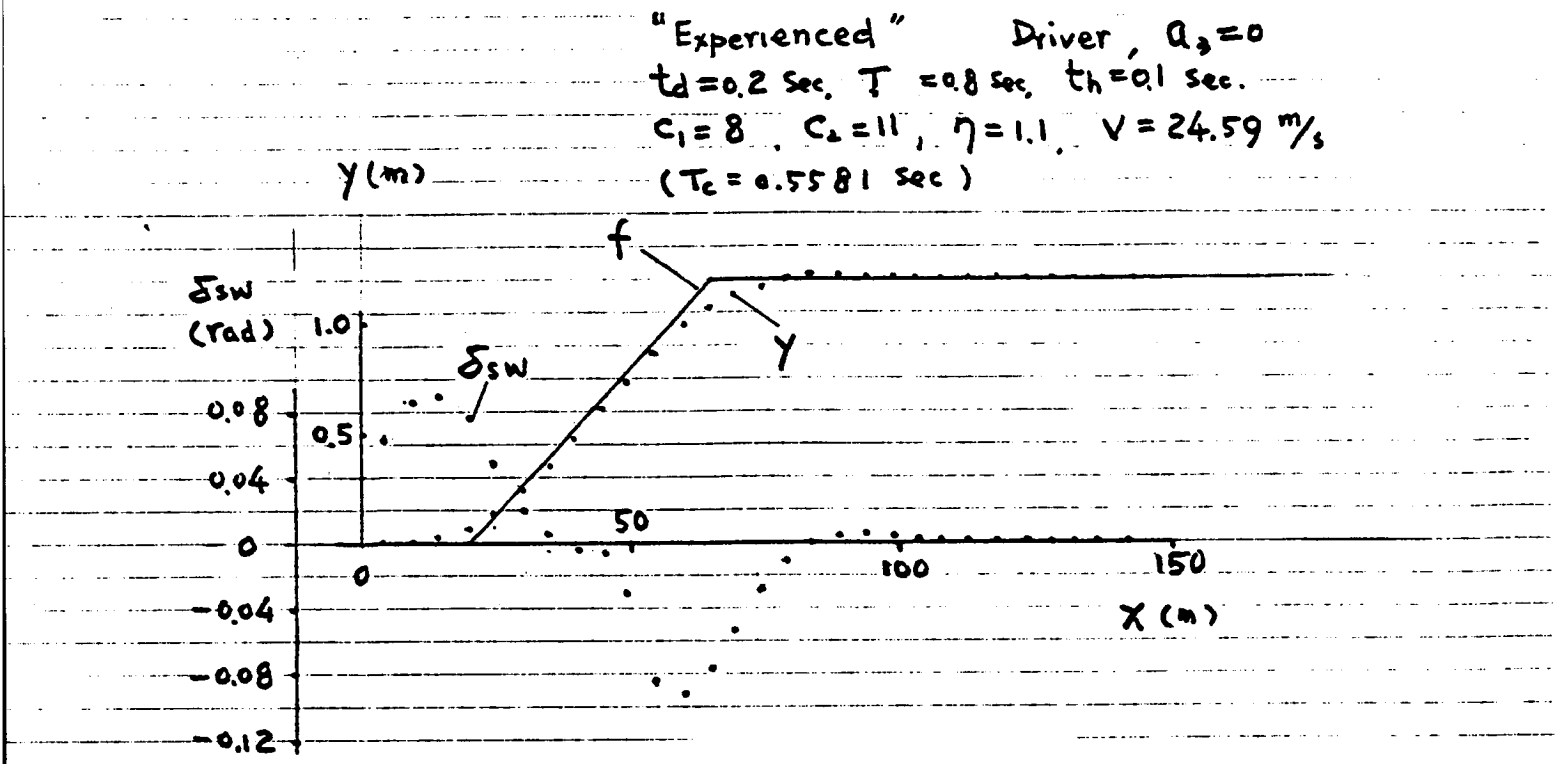


Figure 3.4(a). Simulation results of a second-order optimal follower system with look-ahead time T = 0.8 sec.

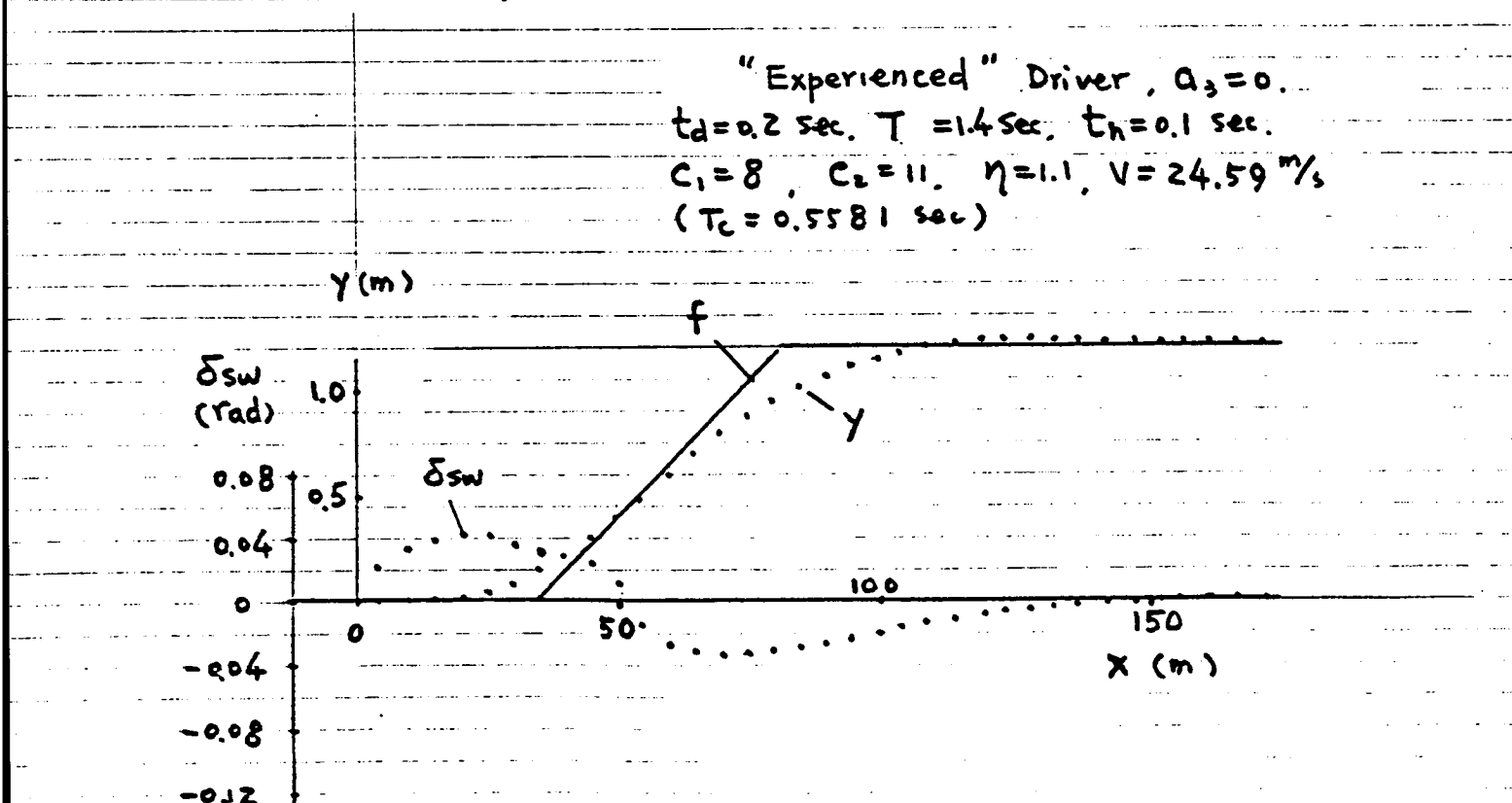


Figure 3.4(b). Simulation results of a second-order optimal follower system with look-ahead time T = 1.4 sec.

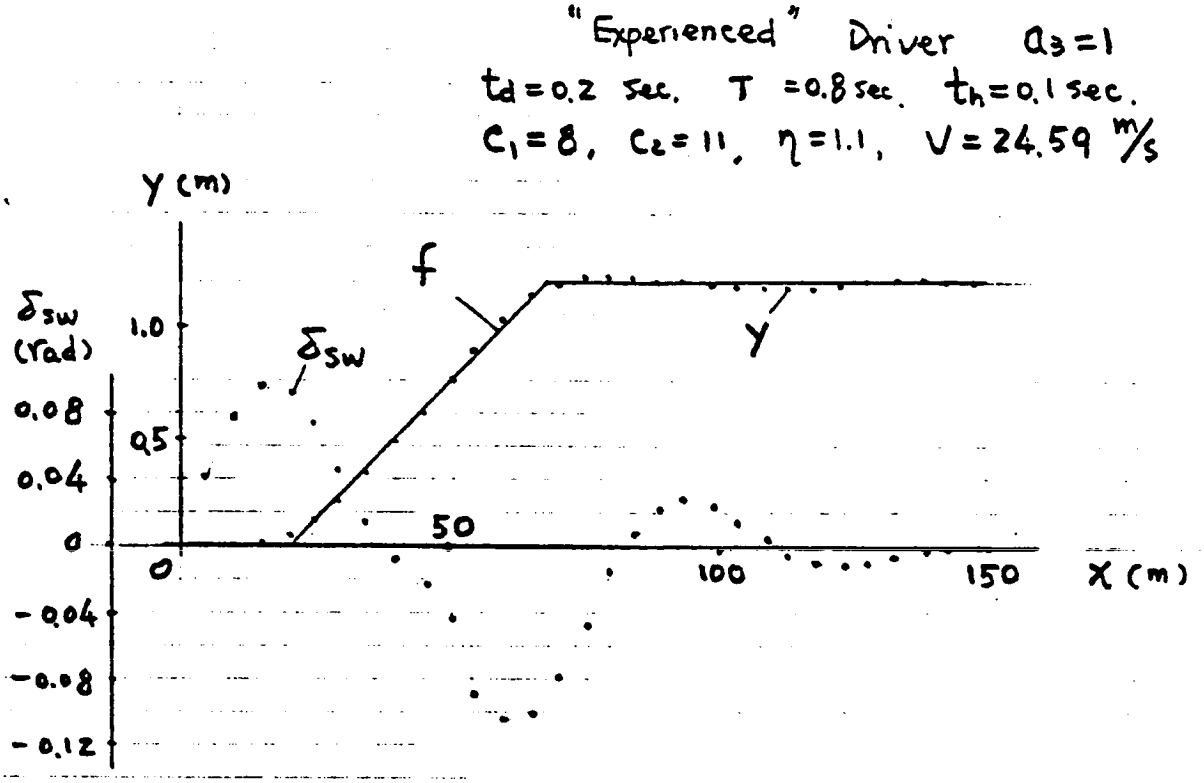


Figure 3.5(a). Simulation results of a third-order optimal follower system with look-ahead time T = 0.8 sec.

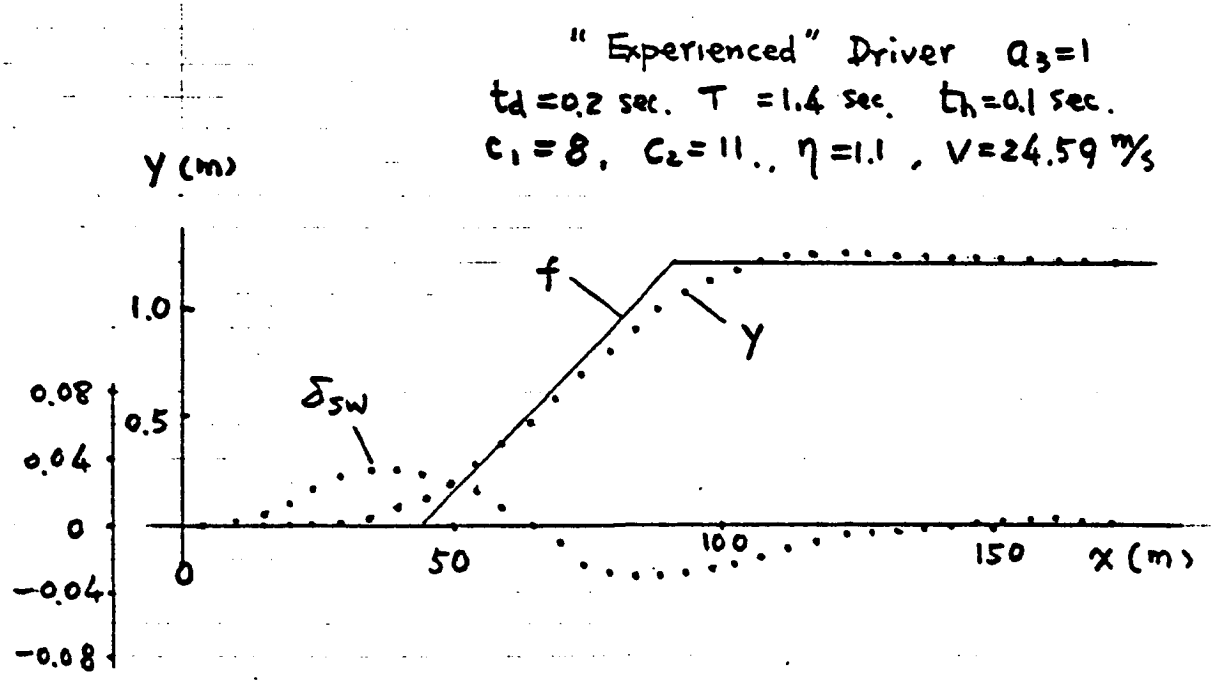


Figure 3.5(b). Simulation results of a third-order optimal follower system with look-ahead time T = 1.4 sec.

$C_1 = \frac{\overline{C_{\alpha 1}}}{W_1}$, $C_2 = \frac{\overline{C_{\alpha 2}}}{W_2}$, cornering coefficients of the the tires on the front and rear axles, respectively.

$\overline{C_{\alpha 1}}$, $\overline{C_{\alpha 2}}$ - equivalent cornering stiffnesses of front and rear axles, respectively.

W_1, W_2 - Normal load on front and rear axles, respectively.

$\eta = \frac{\rho^2}{ab}$ - The dynamic index.

ρ - radius of gyration

a - distance from c.g. to front axle

b - distance from c.g. to rear axle.

For $a_3 = 1$, one obtains a third-order optimal follower system for single-point preview control. Figures 3.5(a) and (b) are examples of simulation results for lane-change maneuvers with different look-ahead times.

It is seen from Figures 3.4 and 3.5 that both the systems with $a_3 = 0$ and $a_3 = 1$ have excellent tracking performance. However, the second-order follower has shorter settling time, while the third-order follower has less maximum error and smoother steering motion.

2) Rectangular Window Preview Control. Generally, the driver may view not a single point, but an interval of path from $x + L_1$ to $x + L_2$ (Fig. 3.6). Assuming that all points in this interval are of equal importance, then the effective path input is

$$f_e(t) = \frac{1}{T_{P2} - T_{P1}} \int_{T_{P1}}^{T_{P2}} f(t+\tau)d\tau \qquad (3.14)$$

where $T_{P1} = \frac{L_1}{V}$, $T_{P2} = \frac{L_2}{V}$

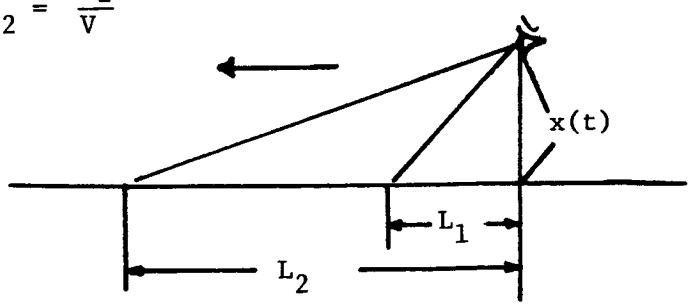


Figure 3.6. A general preview form of a driver.

Since the Laplace transform of $f_e(t)$ is

$$f_e(s) = \frac{f(s)}{T_{P2} - T_{P1}} \left[e^{T_{P2}s} - e^{T_{P1}s} \right] \tag{3.15}$$

therefore, the preview function is

$$P(s) = \frac{f_e(s)}{f(s)} = \frac{e^{T_{P2}s} - e^{T_{P1}s}}{T_{P2} - T_{P1}} \tag{3.16}$$

Expanding Equation (3.16) by using a Taylor's series,

$$P(s) = P_o + P_1 s + P_2 s^2 + P_3 s^3 + \text{---}$$

where

$$\begin{aligned} P_o &= 1 \\ P_1 &= \frac{T_{P1} + T_{P2}}{2} \\ P_2 &= \frac{T_{P2}^3 - T_{P1}^3}{3!(T_{P2} - T_{P1})} \\ P_3 &= \frac{T_{P2}^4 - T_{P1}^4}{4!(T_{P2} - T_{P1})} \\ &- - - - - \end{aligned} \tag{3.17}$$

According to Equations (3.4) and (3.9),

$$\sigma = a_3 P_3 / P_2 = \frac{a_3}{4} \frac{T_{P2}^4 - T_{P1}^4}{T_{P2}^3 - T_{P1}^3} \tag{3.18}$$

$$c(s) = \frac{1}{Gay} (1 + T_c s)$$

$$T_c = T_1 + t_h + t_d - T_{y1} - \frac{a_3}{4} \frac{T_{P2}^4 - T_{P1}^4}{T_{P2}^3 - T_{P1}^3} \tag{3.19}$$

Figures 3.7(a) and (b) show simulation results for a lane-change maneuver for the third-order follower with a rectangular window, where $T_{Pr} = T_{P1} + T_{P2}/2$ and $T_w = T_{P2} - T_{P1}$. The vehicle parameters and T_{Pr} are the same as in Figure 3.4 and Figure 3.5. It is seen that the

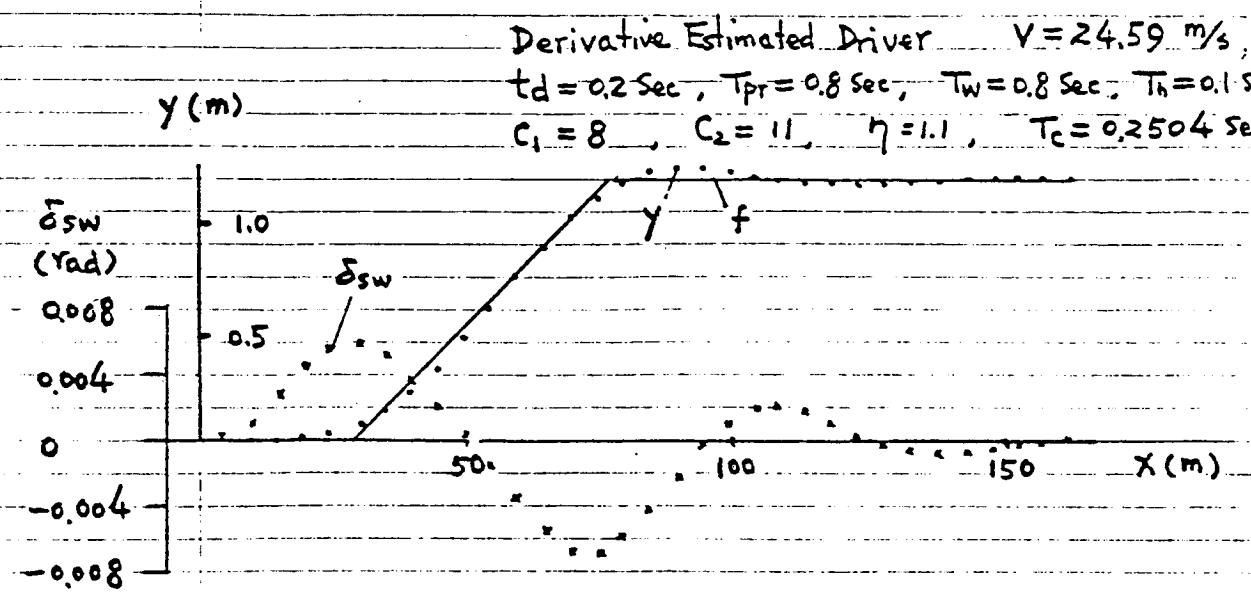


Figure 3.7(a). Simulation results of a third-order optimal follower with rectangular window T_{pr} = 0.8 sec, T_w = 0.8 sec.

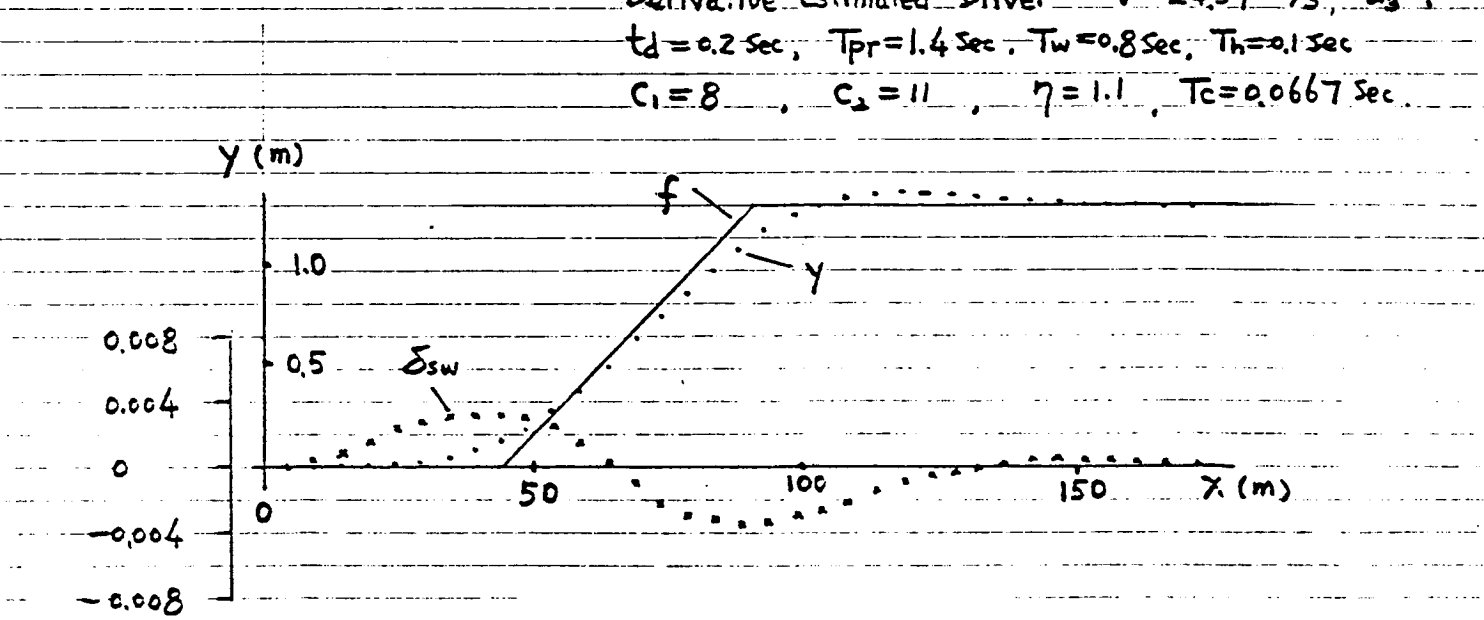


Figure 3.7(b). Simulation results of the system with T_{pr} = 1.4 sec., T_w = 0.8 sec.

results from the application of the two different preview control modes are very similar to each other. However, when using the rectangular window preview model, the motion of the steering wheel is a little smoother and the tracking error is a little larger than that obtained using the single-point preview model.

4. COMPARISON OF SIMULATION RESULTS WITH TEST DATA

In this section test data obtained from a truck performing a double-lane-change maneuver are compared with simulation results using the single-point preview follower model (SPPF) and the rectangular window preview model (RWPF). In both simulations, a linear two-degree-of-freedom vehicle and an "experienced" driver are used.

The following parameter measurements describe the test vehicle:

total vehicle mass	$m = 22,6795$ kg (50,000 lb)
wheelbase	$L = 5.3085$ m (209 in)
front wheel location	$a = 4.0345$ m (158.84 in)
rear wheel location	$b = 1.274$ m (50.16 in)
yaw inertia*	$I = 159,372$ m·N·s^2 (117,531 ft·lb·s^2)
cornering stiffness = front axle**	$C_1 = 210,694.66$ N/rad (47,400 lb/rad)
rear axle	$C_2 = 1,096,957.27$ N/rad (246,772 lb/rad)
vehicle velocity	$V = 20.21$ m/s (66.3 ft/s)

*estimated value; **equivalent, including compliance

The optimal preview time can be selected according to a dominant eigenvalue analysis (see Reference [8]) and for this simulation, a preview time, $T_{Pr} = 1.2$ sec, is chosen for the truck traveling at the velocity of 20.21 m/s (66.3 ft/s).

The simulation results for the lateral position of the truck performing a double-lane-change test are shown in Figure 4.1 (the single-point preview model) and Figure 4.2 (the rectangular window preview model). To see the influence of the coefficient, a_3, the simulation results for $a_3 = 0$, 0.7, and 1 are shown in these figures. It is seen from these figures that both the second-order follower model ($a_3 = 0$) and the third-order follower model ($a_3 = 1$) give good tracking performance. However, the second-order follower shows a little larger error at an earlier stage, but a smaller error at a later stage than the third-order follower. The weak damping in the third-order follower is due to the large "vehicle delay" of the heavy truck.

Figure 4.3 shows test results for the same truck. It is seen that the simulation prediction of $a_3 = 0.7$ is very close to the test data. For both the test data and the simulation results, the maximum position

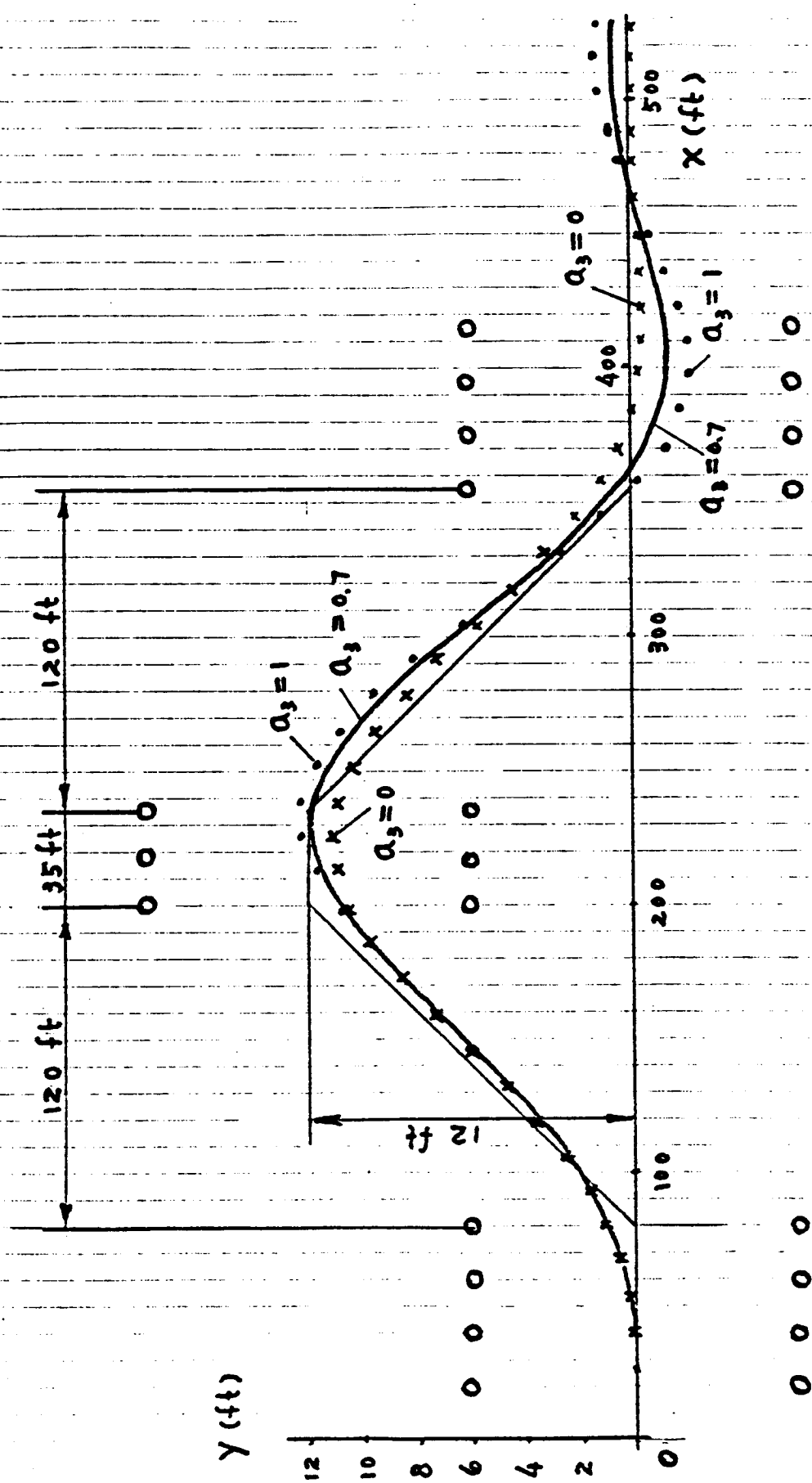


Figure 4.1. Simulation prediction based on the SPPF model for a double-lane-change maneuver.

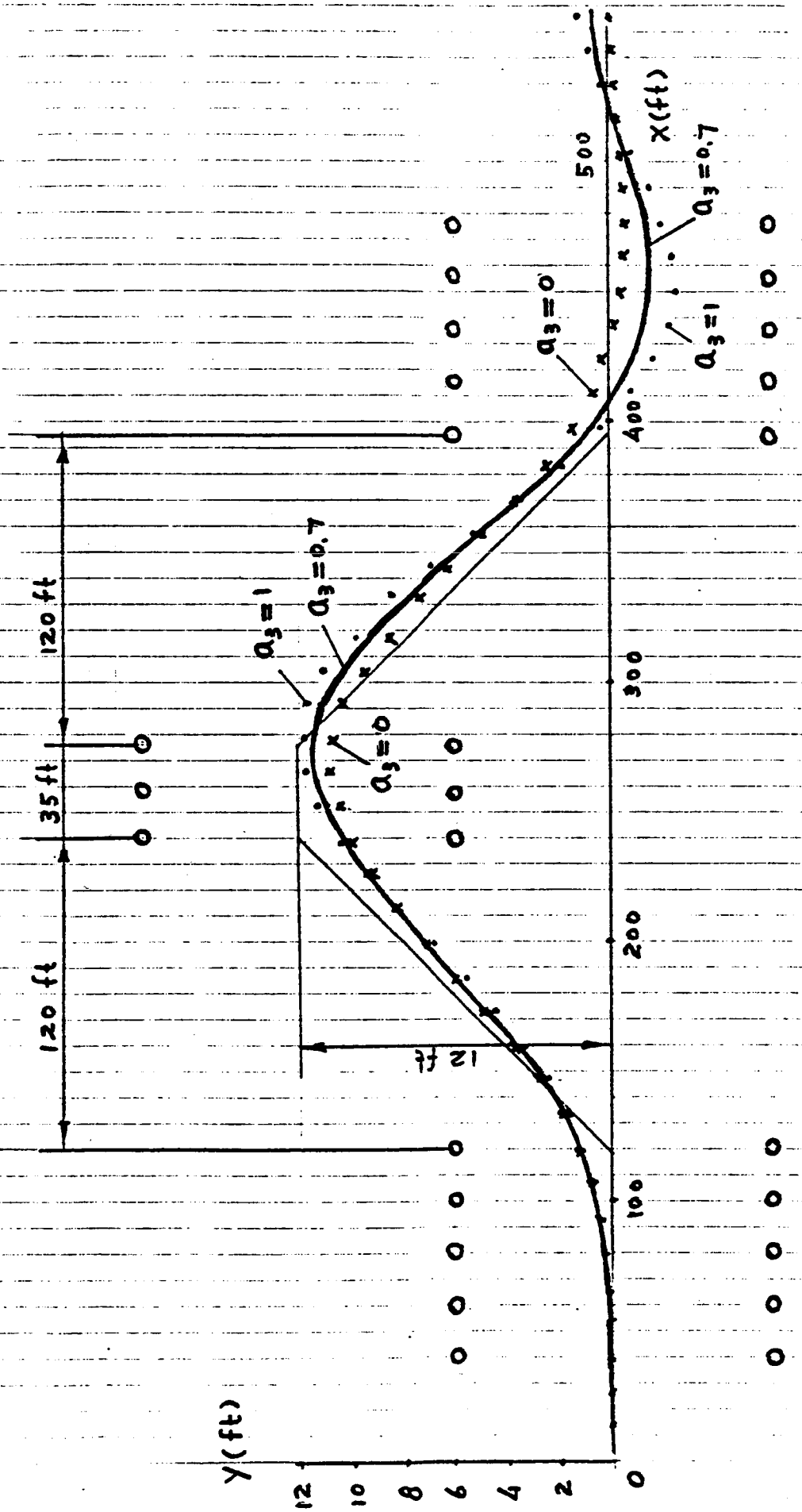


Figure 4.2. Simulation prediction based on the RWP model.

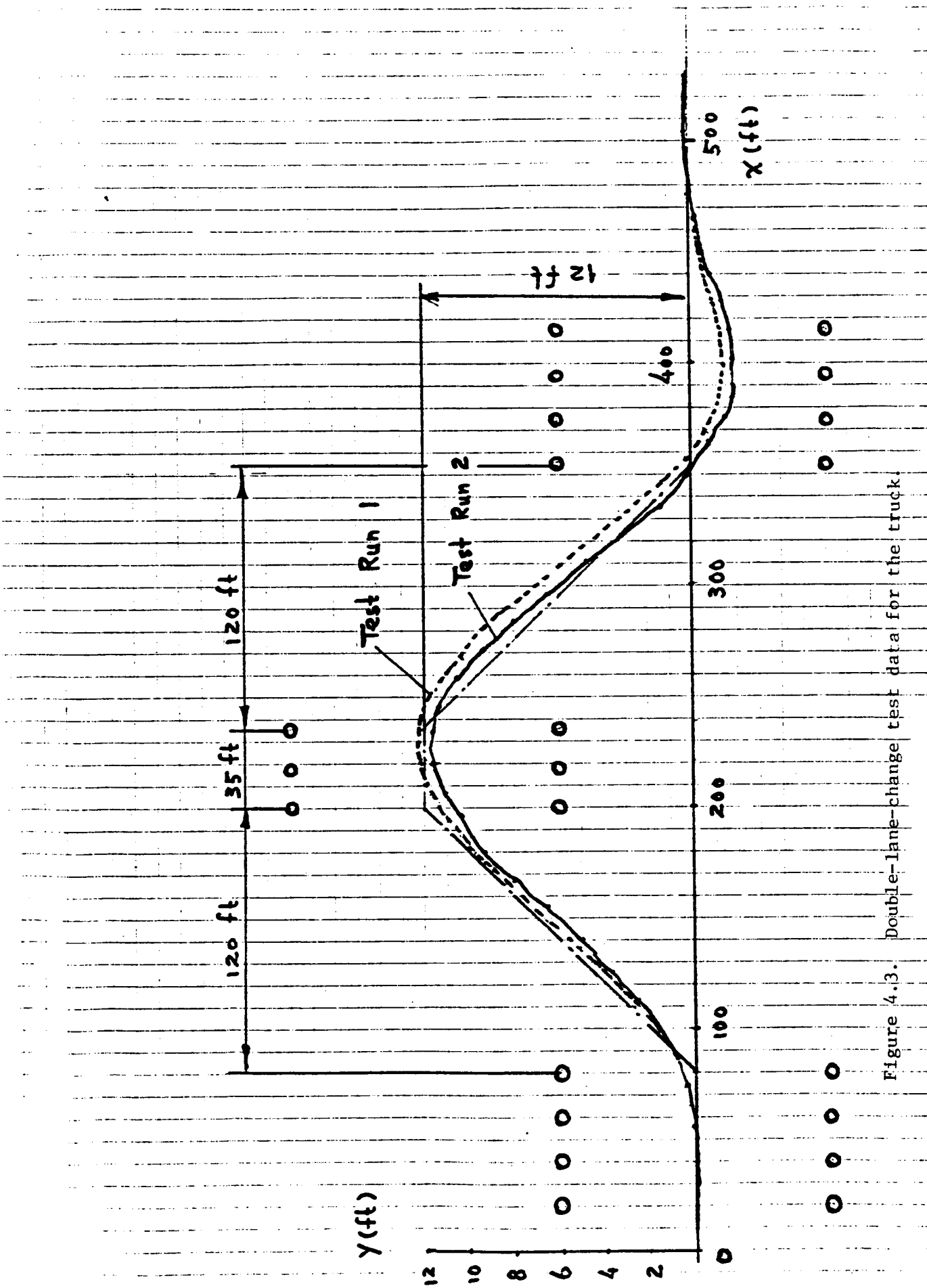


Figure 4.3. Double-lane-change test data for the truck.

errors are almost of the same magnitude (close to 1.5 ft for a 12-ft double-lane-change maneuver) showing that the test results are in good agreement with the computer simulation predictions.

The differences between the single-point preview model and the rectangular window preview model are not very significant even for $T_W = T_{Pr}$. Furthermore, if the window width, T_W, is shortened, the difference between these two models will be even less. (When $T_W = 0$, the results for the rectangular window preview model will be exactly the same as the results for the single-point preview model.)

5. CONCLUDING REMARKS

This paper has attempted to explain the reasoning underlying the assumptions that provide the basis for a preview-follower model. Although the preview-follower concept may seem complicated upon first reading, its virtue lies in the simplicity with which parameters for modeling driver behavior may be determined.

REFERENCES

1. McRuer, D.T. and Weir, D.H. "Theory of Manual Vehicular Control." Ergonomics, 1969.

2. McRuer, D.T., et al. "New Results in Driver Steering Control Models." Human Factors, 1977.

3. Reddy, R.N. and Ellis, J.R. "Contribution to the Simulation of a Driver-Vehicle-Road System." SAE Paper No. 810513, 1981.

4. MacAdam, C.C. "An Optimal Preview Control for Linear Systems." Journal of Dynamic Systems, Measurement and Control, September 1980.

5. MacAdam, C.C. "Application of an Optimal Preview Control for Simulation of Closed-Loop Automobile Driving." IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-11, No. 6, June 1981.

6. Kondo, M. and Ajimine, A. "Driver's Sight Point and Dynamics of the Driver-Vehicle System Related to It." SAE Paper No. 680104, 1968.

7. Maeda, T., et al. "Performance of Driver-Vehicle System in Emergency Avoidance." SAE Paper No. 770130, 1977.

8. Guo, K. "A Study of Methods for Modeling Closed-Loop Vehicle Directional Control." UMTRI Report (to be published in 1983).

ON HUMAN DECISION MAKING AND INFORMATION PROCESSING LIMITATIONS IN DYNAMIC ENVIRONMENTS*

by

Daniel Serfaty, Eric P. Soulsby, and David L. Kleinman

CYBERLAB
Dept. of Electrical Engineering
and Computer Science
University of Connecticut
Storrs, CT 06268

ABSTRACT

The Dynamic Decision Model (DDM) is a normative-descriptive model that has shown to provide an excellent representation of human information processing and decision making in a dynamic multi-task environment. In the present effort, a sensitivity study was performed on the DDM in an attempt to explore the nature by which various task attributes affect human performance. In particular, changes in parameters such as the number of concurrent tasks, task velocity, processing time and task value, were investigated in conjunction with existing notions concerning human operator workload.

Results indicated a general agreement with existing workload/performance theories and some inherent human information processing limits were identified. Extensive experimental studies were performed and quantitatively confirmed to a large extent the DDM analytical predictions of human performance sensitivities. Interesting changes of human operative behavior and strategy were observed under high workload conditions. A Subjective Workload Assessment Technique (SWAT) provided a complementary tool of analysis for the data generated by the various experiments.

* Research supported in part under US Air Force Contract No. F33615-81-K-0510.

I INTRODUCTION

I.1 Background and Motivation of the research

Previous research at the University of Connecticut CYBERLAB was aimed at modeling single human decisionmaking processes in multiple task environments. The experimental part of this effort focused in developing a canonical decision paradigm through which the decisionmaker's performance and behavior was studied as task parameters were changed [1-3]. This paradigm was motivated, in large part, by the C^3 problem of target selection. In this situation, targets of various type move across the display scopes of the human operator, vying for his attention. Each target has a different threat value and processing time requirement. The human, therefore, is faced with the problem of sequencing tasks dynamically so as to maximize the performance of the system.

Fig. 1 shows the fundamental decision loop that is considered in our approach. The human information and decision process involves 1) whether to process a task or gather more information (i.e. monitor); and 2) which one of N tasks to act upon (N is time varying), in order to maximize the system performance. The decision loop is dynamic in nature. As time evolves,

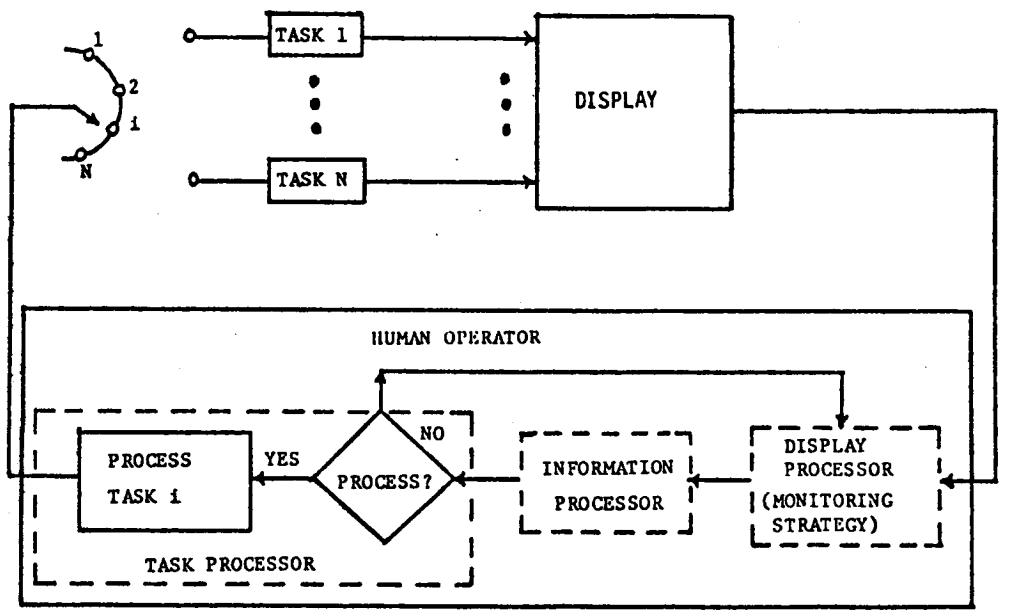


Fig. 1. MULTI-TASK DYNAMIC MONITORING/DECISION LOOP

tasks of different value, duration (processing time) and opportunity window demand the operator's attention, while others depart. The opportunity windows shrink with time as the tasks approach their deadlines.

In the present paper, a joint analytical - experimental sensitivity study was performed to investigate how changes in the information processing and decisionmaking load affects the decisionmaker's performance and operative behavior. The objectives of this effort were multifold:

(1) To analyze the effect of variations in the workload imposed on the operator in regard to his performance and operational behavior

(2) To identify and understand limits on human information processing and action selection, especially in the case of very high workload, and attendant decreasing performance.

(3) To try to find appropriate measures of a priori objective task workload and to compare them to previously known measures.

(4) To validate the Dynamic Decision Model (DDM) predictions of human performance across a wider range of task parameters. Specifically, to do a model-data comparison while varying experimental parameters such as the number of task channels, the velocity of the tasks, the task values, and the processing time range.

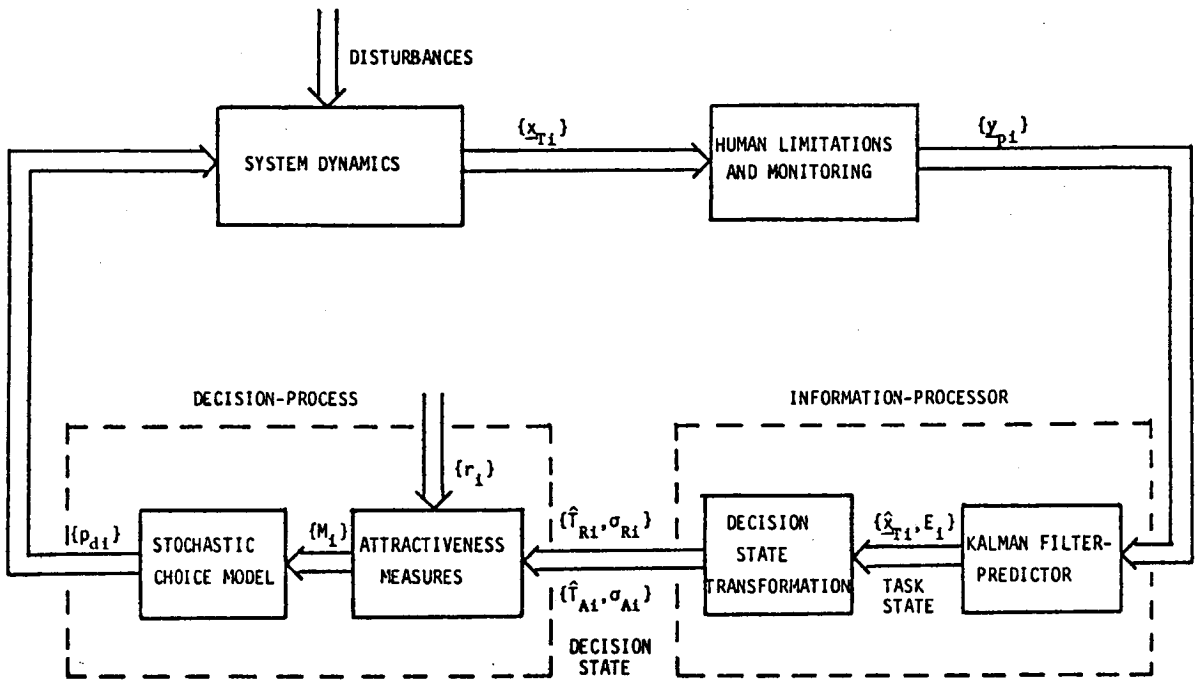
(5) To compare experimental results with subjective measures of workload. The technique used here is the SWAT (Subjective Workload Assessment Technique) [6].

The paper is organized as follows; in the remainder of this section a brief description of the Dynamic Decision Model (DDM) and the experimental paradigm used to validate it will be given. Part II will describe the simulation and experiment done for the sensitivity studies after which Part III will describe some of the main results in terms of model predictions and model data comparisons. In Part IV a study of a proposed apriori measure of absolute workload in terms of the ratio of average Time required/Time available will be presented. Part V discusses a comparison between experimental results and subjective workload prediction using SWAT. Comments on the implication of the results obtained on workload, human strategy changes and the DDM's prediction as well as considerations leading to future research directions will conclude the paper.

I.2 Review of the Dynamic Decision Model

Our main analytical tool will be the Dynamic Decision Model (DDM) previously developed by Patipatti, Kleinman and Ephrath [1-3]. This normative-descriptive model contains several interesting features. First, the analytic framework of DDM is based on optimal control, estimation, and semi-Markov decision process theories. Thus, this approach provides a general methodology for analyzing dynamic decisionmaking under uncertainty. Second the model introduces the important concepts called the "task" state and "decision" state. The task state is the detailed description of the internal variables associated with each task (position, velocity, type, etc.). The decision state, which is a functional transformation of the task state is the minimum number of variables that provides different information for making decisions (time available, time required to complete each task). Third, the DDM employs the widely - validated Kalman filter theory in an information processing submodel to provide the conditional mean and covariance of the decision

state. Fourth, and perhaps most important, the DDM explicitly incorporates human limitations in the information processing as well as in the decision stages. These include the reaction time delays, randomness (i.e. observation noise), limited combinatorial capabilities and randomness in decisions. The modeling of human randomness in decisions has been done via Luce's choice axiom, which provides a convenient mechanism for modelling "bounded rationality" behavior of the human decisionmaker.



DYNAMIC DECISION MODEL OF HUMAN TASK SELECTION PERFORMANCE

Fig. 2

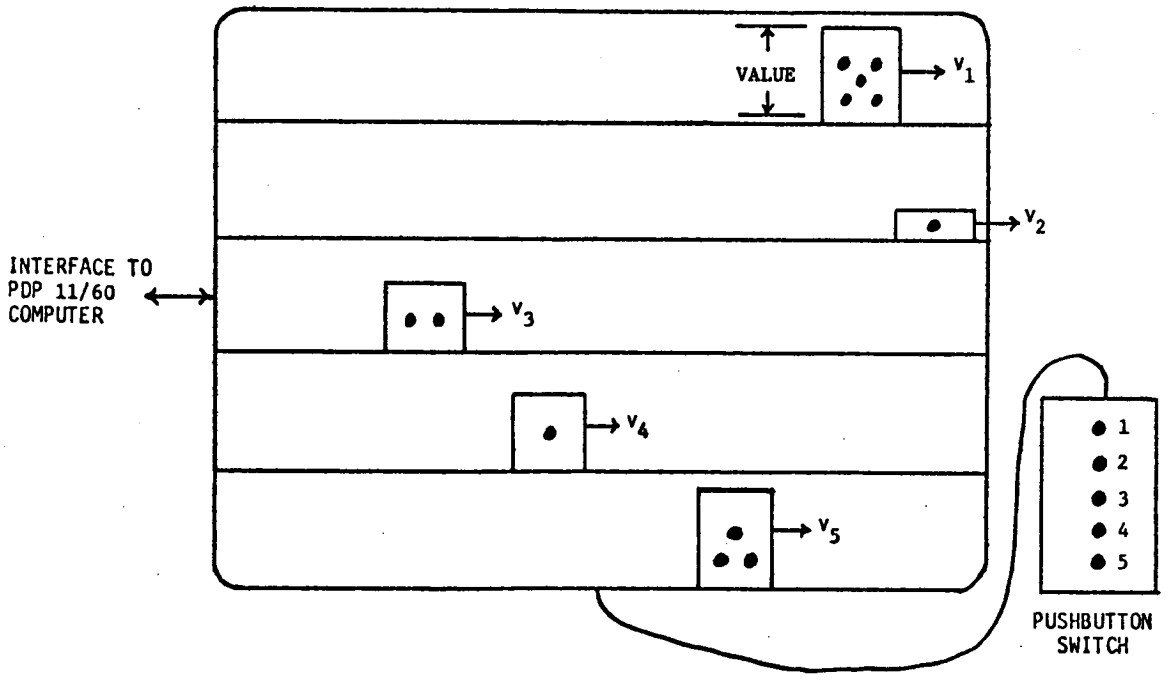
A block diagram of the DDM is shown in Fig. 2. Each of the N tasks in the opportunity window is represented by a dynamic system acted on by disturbances to account for the nonstationarities in task characteristics. The human's perceived outputs $\{\underline{y}_{pi}\}$ are delayed, noisy versions of the task states $\{\underline{x}_{Ti}\}$ and are contingent upon the monitoring process. The perceived outputs are processed to produce the best linear unbiased estimates of the task states $\{\hat{\underline{x}}_{Ti}\}$ and their associated covariances $\{E_i\}$ via a Kalman filter-predictor submodel previously used in the Optimal Control Model (OCM) methodology. The statistics of the task states $\{x_{Ti}, E_i\}$ are, in turn, used to determine the first and second order statistics of the decision state $\{T_{Ri}, \sigma_{Ri}\}$ time required and $\{\hat{T}_{Ai}, \sigma_{Ai}\}$ time available. The statistics of the decision states along with the task values, $r_i(t)$, are combined to determine the attractiveness measure, $M_i(t)$, of each task in the opportunity window. Subsequently, the measures are used to generate the probability $P_{di}(t)$ of acting on each of the N tasks and the probability $P_{do}(t)$ of not acting on any task (or the monitoring probability, $P_{dm}(t)$).

The DDM is capable of predicting various performance measures, such as: the total reward earned, the percentage of total possible reward earned, the number of tasks processed, the total amount of time spent acting on tasks, etc. [1-3].

I.3 The Experimental Paradigm

A simple, yet realistic, computer controlled experimental set-up was considered as indicated in Fig. 3. In the experiments, the subjects observe a CRT screen on which multiple, concomitant tasks are represented by moving rectangular bars. The bars appear at the left edge of the screen and move at different velocities to the right, disappearing upon reaching the right edge. Thus, the screen width represents an "opportunity window".

The height of each bar corresponds to the reward (value) of the task. The amount of time required to process a task (in seconds) is represented by the number of dots displayed on a bar. A task is processed by the subject when he pushes the appropriate push-button as shown in Fig. 3. By processing a task successfully, the subject is credited with the corresponding reward (r_i), and the completed task is eliminated from the screen. An attempt by the decisionmaker to act on a task that can not possibly be completed (i.e. the time required is greater than the time available) constitutes an error. No partial credit is given.



EXPERIMENTAL PARADIGM FOR TASK SELECTION

Fig. 3

II THE SENSITIVITY PROGRAM

MODEL SIMULATION AND EXPERIMENTS

II.1 Simulation Method for the Sensitivity Study on the DDM

On the basis of existing notions concerning the effect of various task parameters on the human operator's performance and subsequent workload [4], four task attributes were considered. They include:

A) different task velocities,
B) different task values,
C) different task processing times,
D) varying the number of task lines (channels) to be monitored.

As indicated in Fig. 3, the task velocity is inversely related to the time available in which to successfully engage and process a task. More specifically, the time available to process task i at time t is given by

$$T_{ai}(t) = \frac{L-\ell_i(t)}{V_i(t)}$$

where

L is the length of the opportunity window (1024 screen units) - 1 inch = 85 s.u.

ℓ_i is the position of the task from the left edge

V_i is the velocity of the task

Hence, it was postulated that by increasing the average task velocity a decrease in performance (as measured by the number of tasks completed and the total reward earned) would be obtained. A similar relationship between performance and the average processing time of the tasks was expected, since as more time was required to process a task fewer tasks could be completed prior to leaving the opportunity window.

By increasing the number of distinct task values or the number of task channels (lines) which must be monitored simultaneously, a decrease in performance was expected due to the increased demands placed on the information processing portion of the model in conjunction with increasing uncertainty (i.e. conflicts) among decision alternatives.

II.2 The Experimental Program

To parallel most of the sensitivity studies performed on the DDM by model simulation, the experimental paradigm described in I.3 was used. For each experimental setting, 3 different subjects, all graduate students in the Department of Electrical Engineering and Computer Science at the University of Connecticut, were tested on 2 different experiments (random arrivals) yielding 6 values for each average data point obtained. The

duration of each experiment was 90 sec and the subject had immediate visual feedback on his overall performance. The subjects were trained on various experimental conditions; however data was taken only after a certain level of subjective confidence was felt by the subjects.

At the end of each sequence of experiments, a short interview was designed to try to determine the various subjective strategies used. Overall, it was interesting to remark that although each of the subjects used different strategies to maximize their rewards, the intra and inter subject differences in performance were minor. Therefore, the experimental results presented in the next part will describe only the mean performance and not the inter subject variations.

III RESULTS AND MODEL - DATA COMPARISONS

Only the most significant results will be presented here. A complete presentation of the DDM sensitivity study can be found in [5].

III.1 The Sensitivity to task velocity

With regard to task velocity three separate DDM computer simulations were performed:

1) all tasks had the same velocity which was varied from 25 screen units (su) per second to 200 su/sec.

2) all tasks had the same mean velocity with uniform distribution range of ±25 su/sec, the mean varied from 25 su/sec. to 175 su/sec.

3) all tasks had the same mean velocity of 100 su/sec with uniform distribution ranges of 0, ±25, ±50, ±75, ±100.

Results from part (1) are in general agreement with our hypothesis concerning the relationship between velocity and the time available to process a task. As expected, a sharp decrease in the percentage reward earned was observed as the velocity of the tasks was increased. It does appear, however, that there is a tendency for the model to reach a "saturation point" as indicated by the number of tasks (20±2) completed and in the total expected reward earned during the simulation duration (90 sec).

In part (2) of the velocity study, the mean velocity was increased in a manner similar to part (1) but with some uncertainty in this value. Results were expected to show a general agreement with those obtained in part (1), which they did: The HO performance was affected very little by the small uncertainty (±25%) on the task velocity.

Finally, part (3) attempted to investigate the effect of various amounts of uncertainty on the velocity of each task as perceived by the information processing portion of the model. A mean velocity of 100 su/sec was chosen for all the tasks with the interval over which the velocity was uniformaly distributed (e.g. the velocity standard deviation) being increased. Results indicated that the model was able to successfully overcome the perceived uncertainty in task velocity and develop a strategy to improve performance, although this may have been a by-product of having more than enough time in which to process the tasks.

The experiments done for the sensitivity study on task velocity parallels the simulation study (2) where all the tasks have the same mean velocity (with a small uncertainty of ±25 su/sec) which was varied from "slow" to "fast". Here again the general hypothesis was that the performance of the subjects (in terms of percentage of reward earned) will degrade as the velocity of the tasks was increased due to the corresponding decrease in the time available to process a task (i.e. the faster the task, the sooner it reaches the end of the opportunity window). Fig. 4 presents model/data comparison in terms of the percentage of reward earned and the total reward earned. Model predictions and experimental results were in excellent agreement up to the "nominal" situation of a mean velocity of 75 su/sec (a task of moderate velocity). Beyond that point, the human subject performed better than the model, reflecting

a process of adjustment by the human to the increasing workload demands (an ability the model doesn't have at present).

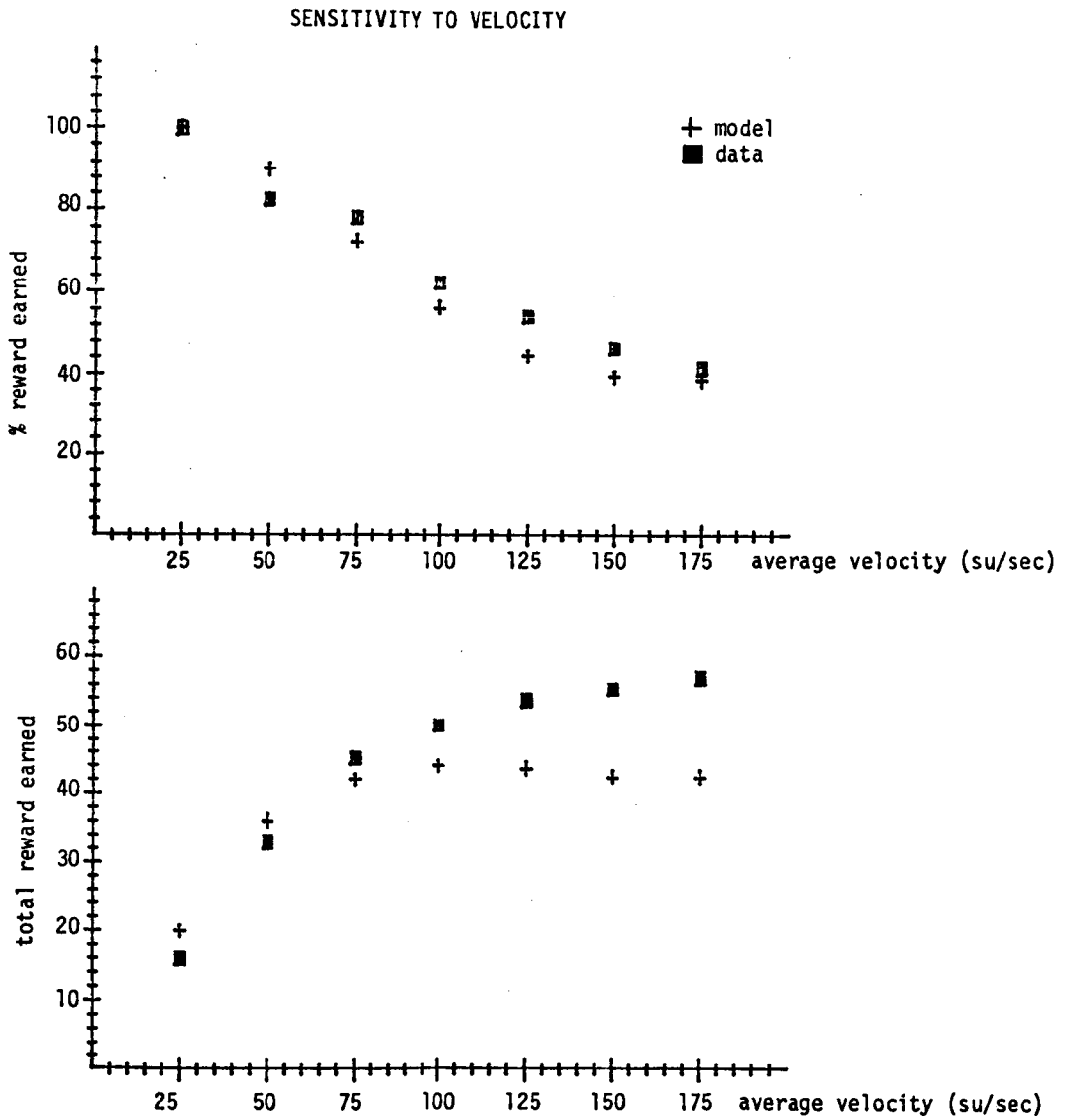


Fig. 4

In summarizing the velocity study, it seems apparent that the most influential aspect of velocity on performance manifests itself in the resulting time available/time required trade-off that must be evaluated in task selection. Clearly the information processing aspect of uncertainty in the perceived velocity played a lesser role in determining the resulting performance. This result is in agreement with existing notions regarding available time to process a task and the performance achieved, and forms a central theme of the workload theories which subscribe to the notion that the rate at which tasks must be processed is indicative of the workload imposed on the operator. In high workload situations, the divergence in performance between model predictions and human subjects seems to signal the fact that the human develops a strategy that adapts to high workload demands (high task speed, short inter arrival times).

III.2 The Sensitivity to range of Task Values

In this study it was hypothesized that by increasing the number of possible reward values of the tasks a resulting increase in the uncertainty of selection due to conflicts would appear. Results of this part of the model simulation study showed consistent performance, as measured by the total percentage tasks completed, across the range of task values. It seems apparent that the model was able to successfully discriminate tasks of low value from those of high value and adopt a strategy reflecting this segregation.

As a way to verify the assumption that the model was able to discriminate high valued tasks from low valued tasks, a comparison of the actual reward earned and the mean value of the tasks multiplied by the number of tasks earned was performed (see Fig. 5). Clearly, here we see that the model was able to segregate the more desirable (high valued) tasks to be processed.

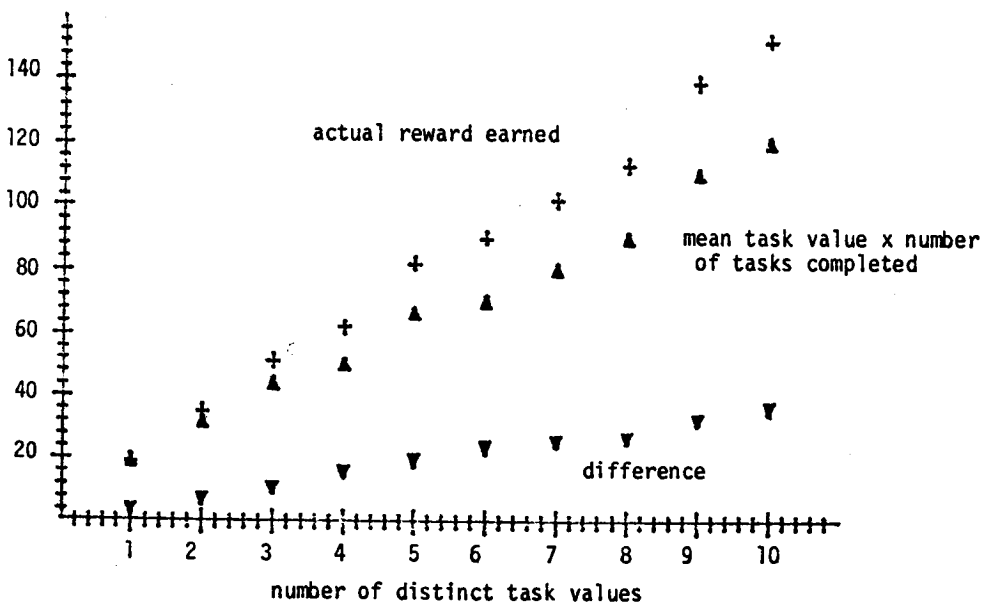


Fig. 5

With regard to overall performance, a normalized reward (actual reward earned/average task value) was computed and plotted in (Fig. 6). Here again the consistent performance achieved by the model is exhibited.

Summarizing the study on the relationship between the performance achieved and the number of distinct task values, it appears that the model is able to successfully segregate more desirable tasks to be processed from lesser ones. By doing so, consistent performance is achieved. Hence it appears that the range of possible values produces a strategy which is able to compensate for any conflicts in choice such that a "saturation level" in performance is obtained.

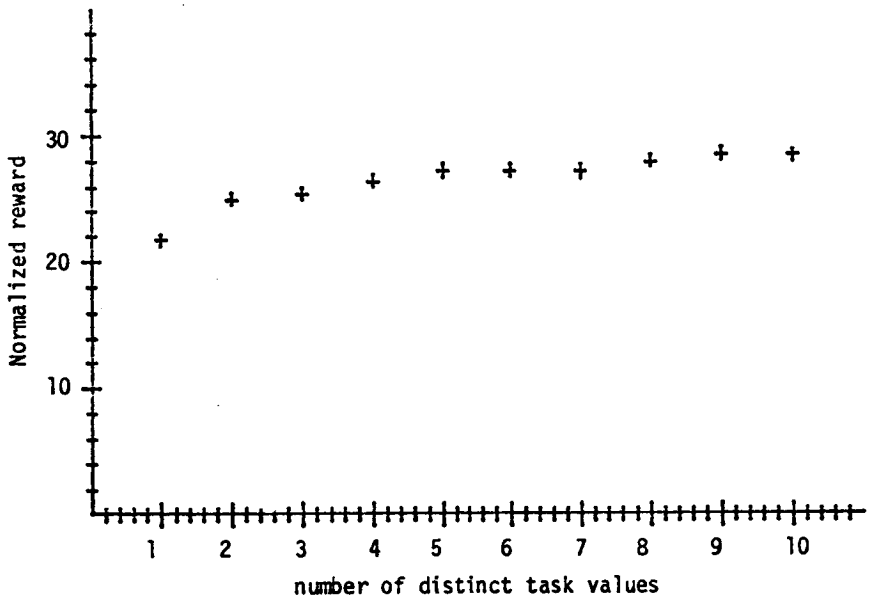


Fig. 6

It appears that a heuristic measure of an a priori absolute attractivity of a task i is a ratio γ_i

$$\gamma_i = \frac{\text{Value of task i}}{\text{Time required to process task i}}$$

which induces a multi-level threshold strategy:

1. for $\gamma_i > \gamma'$ "Try always to process task i"
2. for $\gamma_i < \gamma''$ "Never even try to process task i"
3. for $\gamma'' \leq \gamma_i < \gamma'$ "Depends on concurrent tasks"

For example: for the simulation with 10 different task values and 5 different processing times the values $\gamma' \simeq 6.2$ and $\gamma'' = 4.3$ have been found.
Note: No experiment was performed in the study of sensitivity to the # of task values.

III.3 Sensitivity to the range of task processing time

In this study, the range of possible seconds needed to process a task was varied from all tasks having the same processing time of 1 second to that of the tasks having anywhere from 1 to 8 seconds processing time. The velocity of the tasks was chosen to be a nominal value (100 screen units/sec) corresponding to a moderately paced task with a small standard deviation about this value. Here again performance was expected to decrease as a result of the time available/time required trade-off that affects the decisionmaker's ability to process tasks prior to their deadlines. Indeed both model and experimental results confirm this hypothesis (Fig. 7).

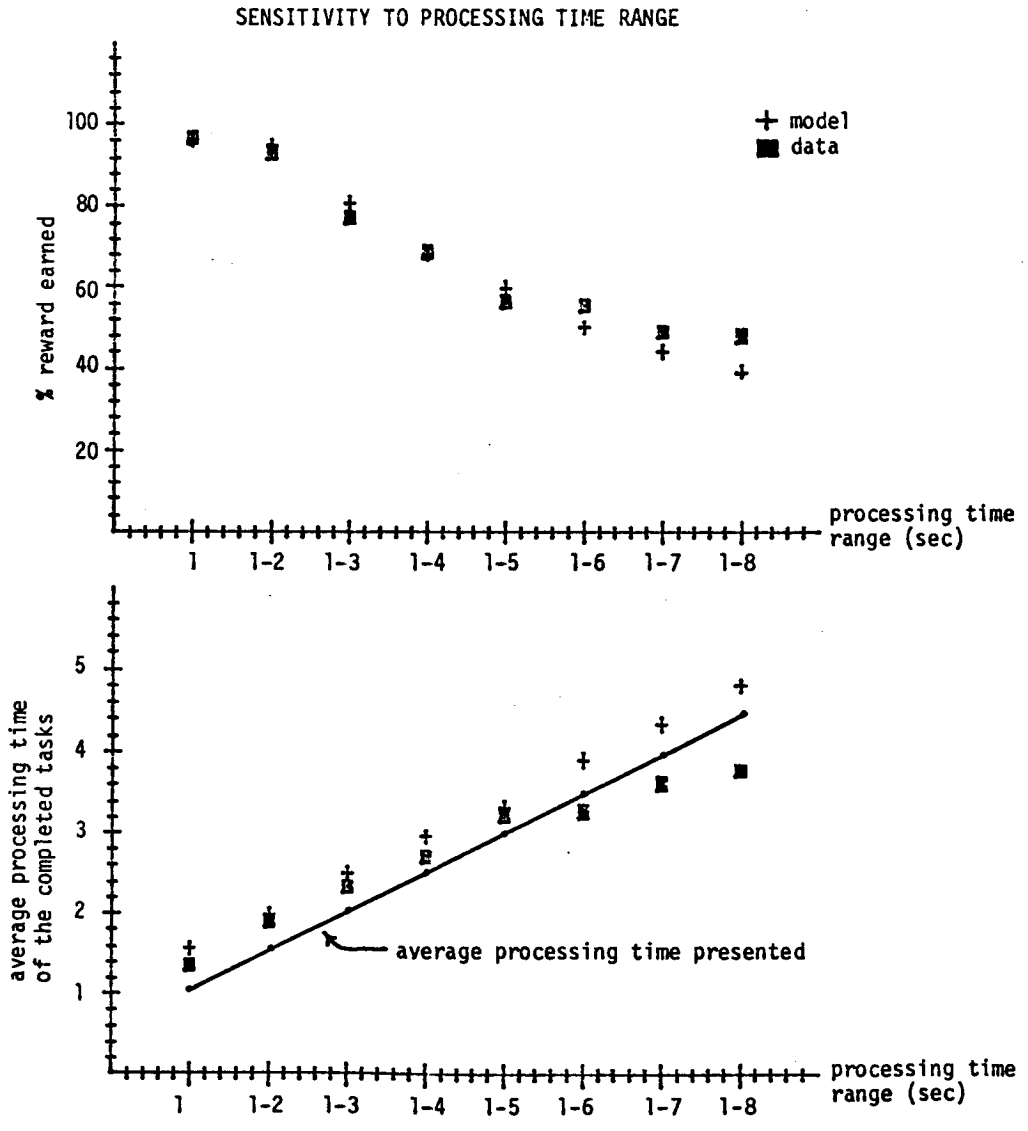


Fig. 7

As can be seen from the plots, the model and human are in good agreement up until the nominal case of tasks which had a processing time range of 1-5 sec. After this nominal situation the human processed tasks with processing times which were below the average value of the tasks presented to him. This change in the type of tasks processed by the human reflects a change in strategy that accounts for the better performance achieved as compared with the model. In other words, the human adapts a more efficient behavior as the information load (range of possible processing times) increases. He discriminates among the tasks that are favorable (small processing time) from those deemed unfavorable (large processing time hence smaller time available to process task prior to its deadline).

III.4 Sensitivity to the number of task channels

In this study, it was hypothesized that by increasing the number of task lines to be monitored simultaneously a decrease in performance would result due to increased demands on the information processing and action selection capabilities of the human. Fig. 8 shows the performance achieved by both the model and the human. As expected the percentage of reward earned decreases as more task lines are presented to the decisionmaker. The human again outperforms the model as the number of task lines increased beyond the nominal condition of five lines. This appears to be due to the fact that the human adjusts his behavior in an attempt to keep his performance constant when the number of tasks to be monitored simultaneously approached his short term memory capacity (5-7 lines).

The plot of the number of tasks completed by both the model and the human as a function of the number of task lines shows the so-called inverted-U shape hypothesized to exist under conditions of high workload. The experimental data, however, indicates a leveling off in the number of tasks completed by the human; achieved by adjusting his behavior to account for the increase in the number of task lines, (i.e. maintaining a subjective level of performance).

Also indicated in Fig. 8 is that the human is outperforming the model in terms of the probability of committing an error by the model is increasing as a function of the number of lines monitored, whereas the human maintains a relatively constant low error level. This fact can explain the constant human performance phenomenon exhibited earlier, in the sense that the human, by trying to process less information as the workload increases, is able to commit fewer errors and therefore to achieve a higher performance than the model prediction.

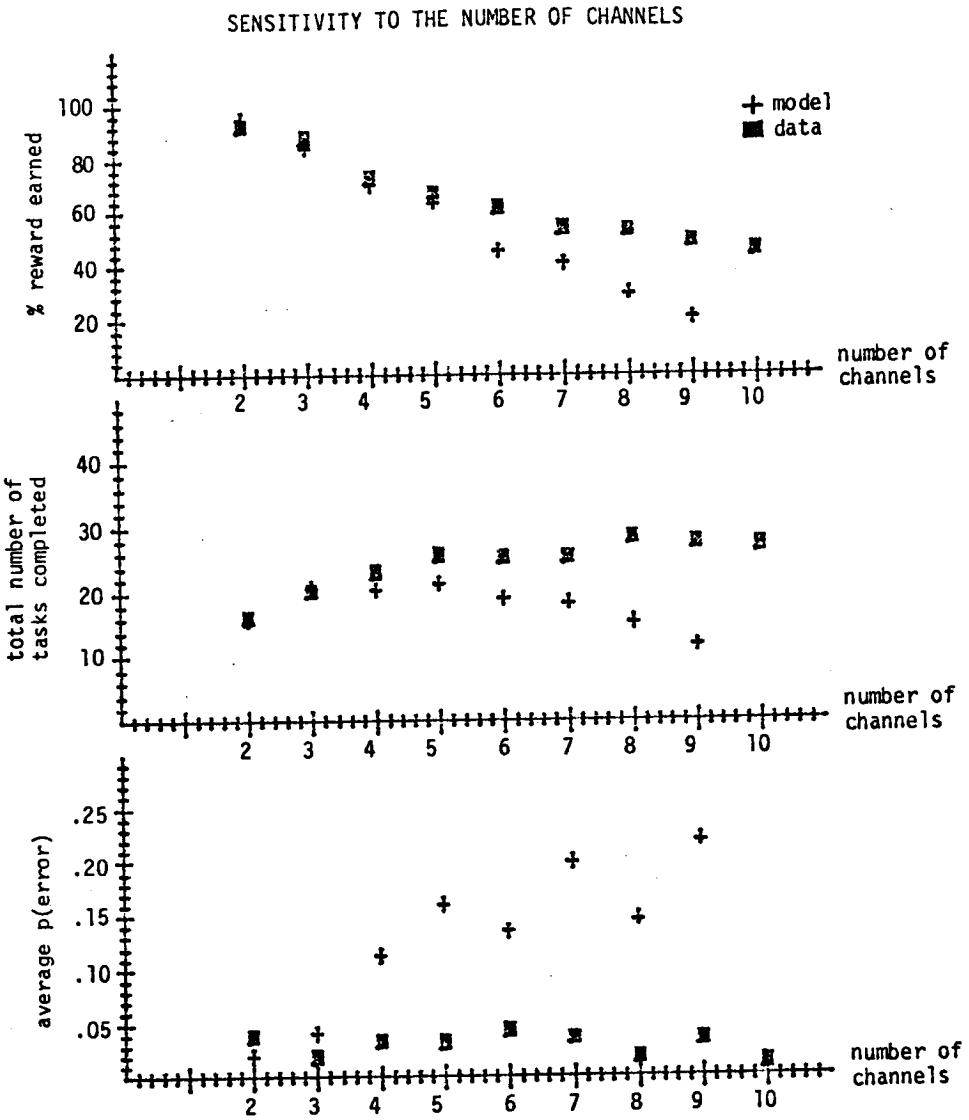

SENSITIVITY TO THE NUMBER OF CHANNELS
% reward earned
100
80
60
40
20
model
data
number of channels
2 3 4 5 6 7 8 9 10
total number of tasks completed
40
30
20
10
number of channels
2 3 4 5 6 7 8 9 10
average p(error)
.25
.20
.15
.10
.05
number of channels
2 3 4 5 6 7 8 9 10

Fig. 8

IV A HYPOTHETICAL A PRIORI MEASURE OF WORKLOAD: THE ρ RATIO

As a result of the sensitivity program described before, in particular the study of sensitivities to velocity and processing time range, it was hypothetized that the ratio ρ

$$\rho = \frac{\overline{T_r}}{\overline{T_a}} = \frac{\text{Average time required}}{\text{Average time available}}$$

is a suitable measure of a priori workload exhibited by a specific dynamic task in the current experimental paradigm. It can be easily shown that the factor ρ is given by the following formula:

$$\rho = \frac{n+1}{2n} \cdot \alpha (1 - \beta^2)$$

where

- n is the maximum number of unit time (dots) per task
- $\alpha = \frac{\overline{V}}{V_{max}}$ is normalized velocity
- $\overline{V}$ is the average nominal velocity
- $V_{max} = \frac{L}{n \times \delta}$ is a screen window limitation
  - L: length of the opportunity window
  - δ: time value of a unit time (in seconds)
- $\beta = \frac{\Delta V}{\overline{V}}$ represents the uncertainty on the velocity (uniform distribution)

Fig. 9 depicts a map of theoretical curves of equal ρ for different combinations of α and β. The hypothesis that the average factor $\rho = \overline{T_r}/\overline{T_a}$ correlates with workload and affects degradation of performance is verified by model simulation as well as by experimental results also shown in Fig. 9 for the case n=5, α = 1/2 (nominal). It can be seen that an increase in ρ leads to a decrease in performance.

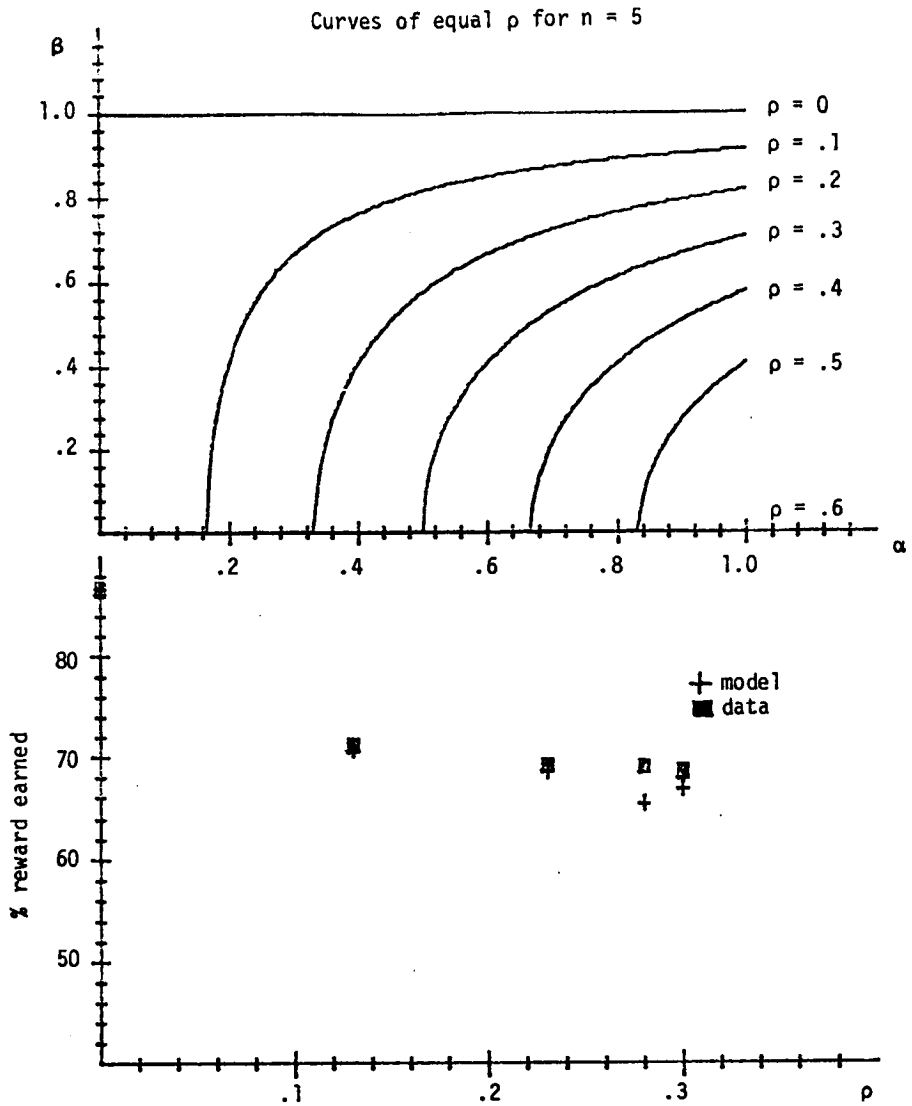


Fig. 9

A sensitivity study was performed on the model and also experimentally to check the assumption that for constant ρ the performance of the decision-maker remains constant, although the range of dynamics parameters such as velocity and dot time value varied in a wide range. Fig. 10 confirms this hypothesis: the experimental data shows that the human develops a strategy to maintain his performance relatively constant when the average a priori workload (ρ factor) is kept constant. Note that when the value of the dots is < .25 sec, the performance degrades, probably due to the fact that we are very close to human time delay limitations. The model, however, does not achieve a constant performance level, but rather exhibits a degradation in

in performance. This degradation yields an increase in the average error probability unlike the relatively small and constant average error probability observed in the experimental data.

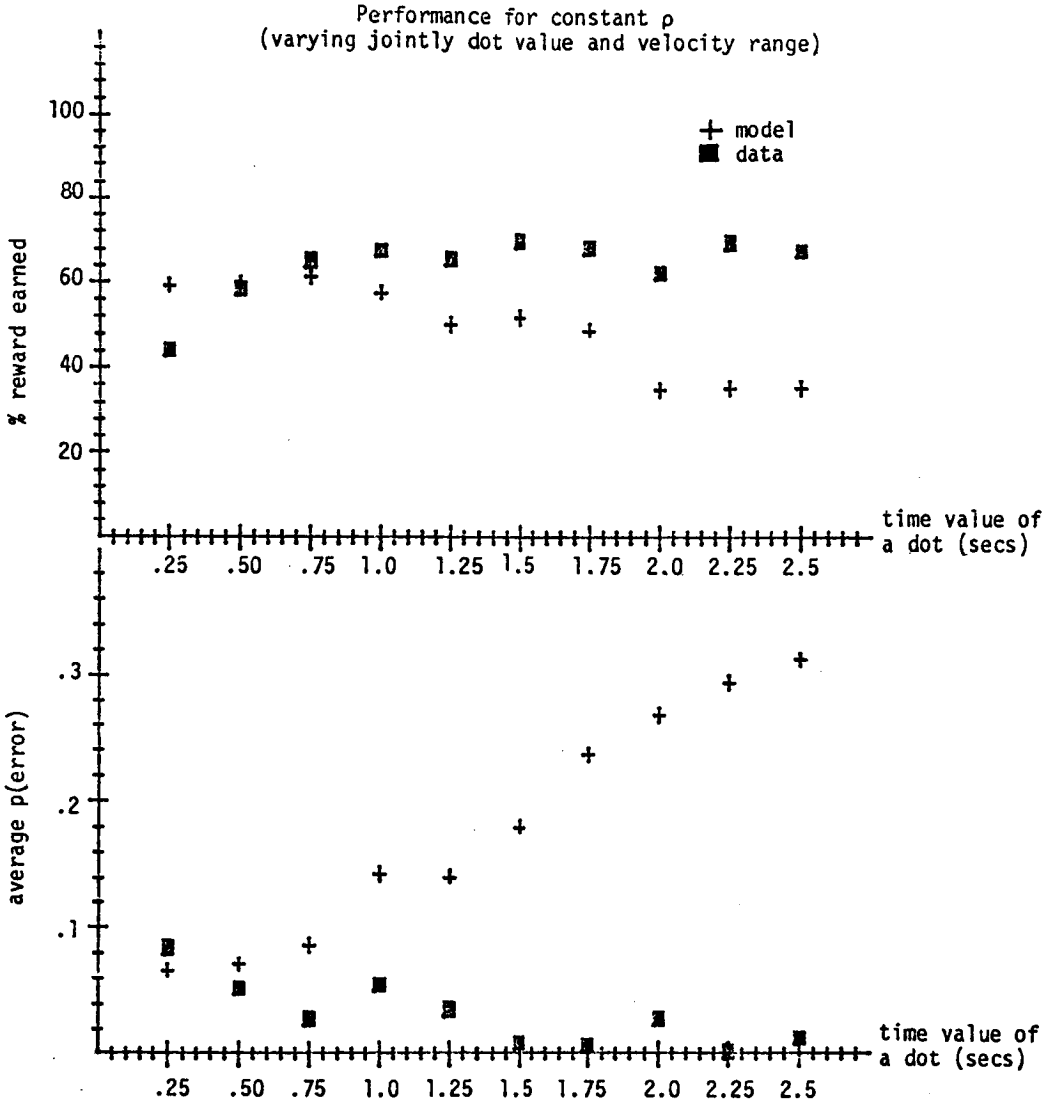


Fig. 10

In conclusion the experimental results allow us to assume that the ρ factor can represent an appropriate apriori measure of human workload in the current dynamic environment.

V SUBJECTIVE WORKLOAD AND EFFECTIVENESS

In an effort to compare the level of workload imposed upon the human with that perceived by the human, rankings of subjective workload were calibrated with subject's rank ordering of workload along the three dimensions of Time Load, Mental Effort Load and Stress Load. The methodology used was the Subjective Workload Assessment Technique (SWAT) well documented in [6]. The experimental setting used was the paradigm presented in part I and the method of controlling imposed workload was by increasing the number of task channels from 2 to 10, all other parameters (velocities, task values and processing time ranges) remaining fixed. As expected, the average workload perceived by the subjects increased linearly with the number of task lines to be monitored [Fig. 11].

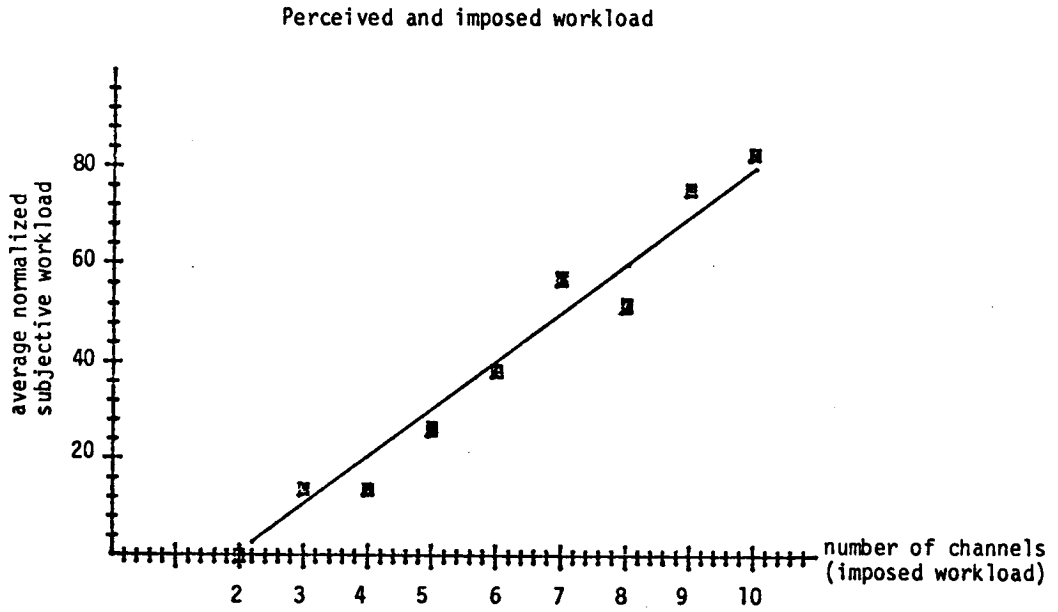


Fig. 11

Corresponding to this increase in perceived workload, a change in human strategy under high workload conditions was noticed in conjunction with the performance level attained (recall Fig. 8). This change of strategy is depicted in Fig. 12, where an effectiveness factor ε given by

$$\varepsilon = \left[\frac{\text{average value of tasks processed}}{\text{average value of tasks presented}} - 1 \right] \times 100$$

is plotted as a function of subjective workload for the same experiment.

Essentially this effectiveness factor indicates that the human adapts a more discriminating strategy (acting on tasks of average task value higher than the average of those presented to him) when he is confronted with increasing information processing demands. This discriminating behavior accounts for the ability of the human to outperform the DDM as discussed earlier in adapting a strategy for high workload conditions. [Fig. 12] shows clearly the 'jump' of operative behavior in terms of average effectiveness between low and high workload conditions.

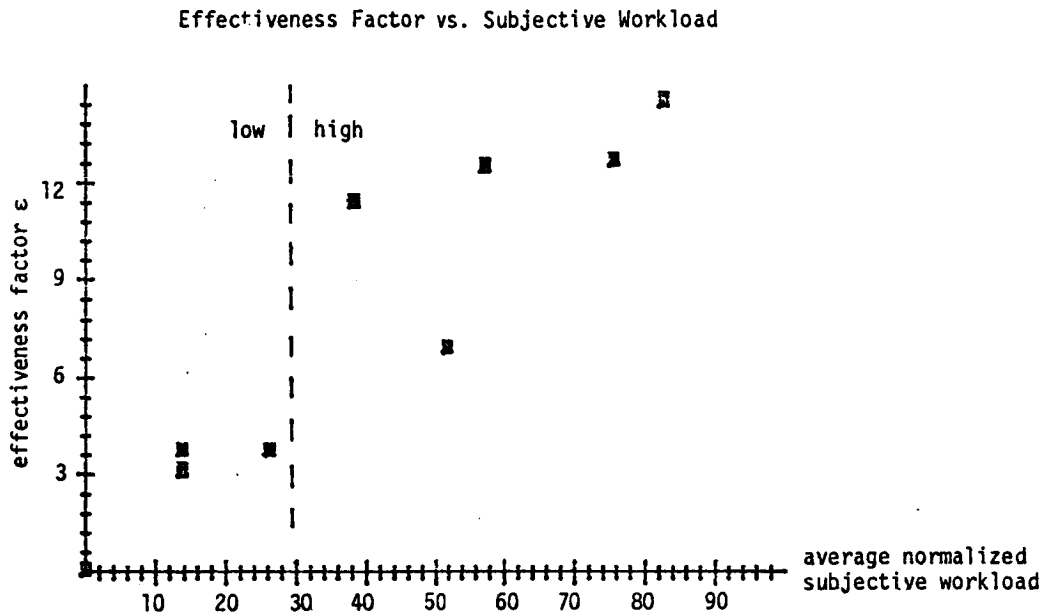


Fig. 12

VI SUMMARY AND CONCLUSIONS

In this joint experimental-analytical study, performance obtained by the Dynamic Decision Model (DDM) and human subjects in the experimental paradigm of Fig. 3 was examined as a function of four task attributes which included different task velocities, different task values, different task processing times, and varying the number of task lines to be monitored. Results of the study showed a general agreement with existing hypotheses regarding human performance as a function of the four attributes mentioned. The primary factor affecting performance was that of task velocity as indicated in the time available/time required trade-off of traditional operator workload research. In addition, it was noted that the human operator was able to successfully discriminate various undesirable tasks from desirable ones as a means to keeping performance consistent. This may be due to a somewhat conservative threshold - strategy employed by the decisionmaker, but in general it is felt to be consistent with one's intuition.

The results of the model simulation study also indicate that a decrease in performance can be obtained not only from increased information processing demands, but also from situations in which action selection uncertainty (i.e. conflicts) exist. This was most clearly represented in the study involving the increased number of task lines to be monitored simultaneously. In that study it has been shown that beyond a certain number (~7) of task lines (channels), the operator performance decreases sharply due to the increasing conflict in action selection and limited information processing capability.

However, in the experiments, the subjects showed different performance in the high workload condition. Instead of decreasing, the subject's performance was maintained at a quasi constant level due to adaptation in operative behavior in order to (1) process less information (probably around Short Term Memory capacity), (2) commit less errors and (3) maintain a subjective level of performance across high workload environments. In diagram form the mechanism by which decisions are made by the Human appears to be that of the following:

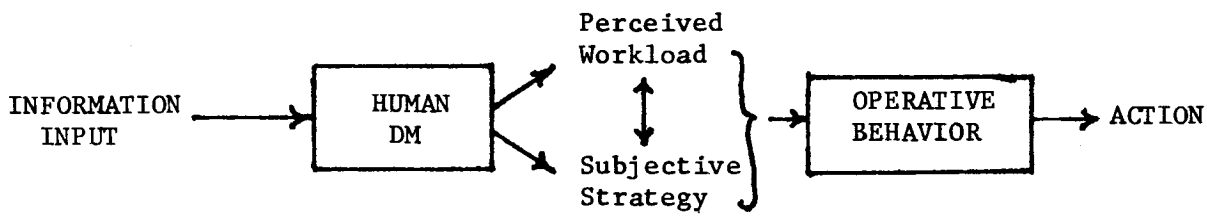


Fig. 13 Anatomy of a decision

Another contribution of the present effort is the proposition of an objective measure of a priori workload of a dynamic task in terms of the factor ρ. This measure ($\rho = \bar{T}_r/\bar{T}_a$) has been found to be adequate in the existing experimental paradigm, although more research will have to be done to improve the measure in different dynamic environments. Subjective workload was found to correlate very well with objective workload imposed by an experiment and an interesting find was that humans tend to be more efficient in reasonably high workload conditions.

REFERENCES

1. Pattipati, K. R., Kleinman, D. L. and Ephrath, A. R. "A Dynamic Decision Model of Human Task Selection Performance" IEEE Trans. on SMC, Vol. SMC-13 (2), March/April 1983.

2. Pattipati, K. R., Kleinman, D. L. and Ephrath, A. R., "A Dynamic Decision Model of Human Task Selection Performance", University of Connecticut Tech. Report, Department of EECS, TR 82-1, March 1982.

3. Pattipati, K. R., "Dynamic Decision Making in Multi Task Environments: Theory and Experimental Results", Ph.D. Thesis, The University of Connecticut, 1980.

4. Moray, N. ed. Mental Workload: Its Theory and Measurement, New York, Plenum Press, 1979.

5. Soulsby, E., Serfaty, D. and Kleinman, D. L., "A Sensitivity Study on the Dynamic Decision Model" to be published as a University of Connecticut Technical Report.

6. Reid, G. B., Shingledecker, C. A., and Eggemeier, F. T., "Application of Conjoint Measurement to Workload Scale Development". In Proc. of 1981 Human Factors Society Annual Meeting, Oct. 1981 (a), pp. 522-526.

CROSS MODALITY SCALING: MASS AND MASS MOMENT OF INERTIA, A FOLLOW-UP

Margaret M. Clarke
Oak Ridge Associated Universities
P. O. Box 117
Oak Ridge, Tennessee 37830

Stephen J. Handel
University of Tennesse
Knoxville, Tennessee 37916

John V. Draper
Oak Ridge Associated Universities
P. O. Box 117
Oak Ridge, Tennessee 37830

Nineteenth Annual Conference on Manual Control
Massachusetts Institute of Technology
Cambridge, Massachusetts

May 24, 1983

INTRODUCTION

Mass, the resistance to linear acceleration, and mass moment of inertia ("inertia"), the resistance to angular acceleration are fundamental properties of objects, and we frequently rely on our perception of these properties in handling and controlling objects in our environment. However, in many modern control systems where the mechanical link between the physical control device and the controlled object is replaced by an electric actuator, direct perception of mass and inertia is impossible and must be deliberately put back in the system. To do this efficiently it may be necessary to know both the extent to which mass and inertia are useful inputs to the operator and also the relative effects on the operator of changes in these two properties, since trade-off decisions might have to be made to input one at the expense of the other.

In general, little research has been done on human perception of inertia, especially when one compares the effort to the quantity of psychophysical research on perception of qualities of other sensory modalities. A brief summary of the highlights of the literature on the perception of inertia follows.

This report was prepared under contract number DE-AC05-760R00033 between the U.S. Department of Energy and Oak Ridge Associated Universities.

Knowles and Sheridan (1966) studied how well subjects could discriminate changes in friction and inertia of rotary controls and what effect these changes had on the subject's control preferences. These researchers found greater sensitivity to changes in inertia than for changes in friction and found that addition of a small amount of inertia produced large increases in preference ratings. Large increases in inertia did not produce proportional increases in preference ratings, however.

Pew and Chananie (1969) investigated sensitivity to inertia and viscous damping, which is resistance to motion that is proportional to acceleration. These authors found sensitivity to inertia increased with inertia, i.e., at higher inertia levels subjects were more sensitive to changes in inertia. Addition of damping reduced sensitivity to changes in inertia.

Kreifeldt and Chuang, (1979a, 1979b) used weights and inertias typical of toothbrushes and tennis raquets to show that sensitivity to weight is about ten times greater than sensitivity to inertia. However, subjects were more influenced by inertia than weight when they judged object similarity and their preference of objects. Clarke, Handel, and Kreifeldt (1982) investigated the perceptual scaling of inertia covering the range of many simple hand-held objects (0.15 through 1.15 inch pound second^2) and reported a linear relationship between physical inertia and perceived inertia. Their data suggested that the sensation of inertia increased less rapidly than the sensation of weight when both weight and inertia are physically increased. Zahner and Kaminaka (1982) also found a linear relationship between perceived inertia (which they called "effort") and the physical property. In addition, they suggested that there are wide individual differences in human sensitivity to relative changes in these properties.

The study described in this paper continues the earlier work of Clarke et al. (1982) and involves the perceptual equation (cross modality scaling) of mass and inertia. Cross modality scaling involves subjects inducing a relationship between two different sensory stimuli. The pioneering work of S. S. Stevens (1975) showed that the perceived magnitude of many sensory stimuli is a power function of physical intensity (for example perceived heaviness is a power function of weight). The relationship may be

expressed thus: The logarithm of the perceived magnitude is a linear function of the logarithm of the physical intensity, and the slope of the function is the exponent of the power function. If the power function holds for two different sensory stimuli, in the present case for inertia and weight, then the function relating the logarithm of perceived heaviness to the logarithm of perceived inertia should also be linear. The ideal way to induce this function would be to ask subjects to choose the inertia stimulus that had the same "intensity of feeling" for them as a given weight (or vice versa); that is, subjects would be able to choose pairs of stimuli (inertia paired with weight) that had equivalent perceptual intensity. A series of such equivalent pairs derived at different stimulus intensity levels would allow a plot of the logarithm of weight versus the logarithm of inertia. The function describing the weight to inertia relationship should be linear, and the slope should be the ratio of power functions for weight and inertia taken separately.

The nature of the physical stimuli involved in this study made it impossible for the subjects to adjust weight (or inertia) until it matched the other dimension. Instead, subjects were presented with one inertia and one weight at each trial and were asked to give the ratio of inertial resistance to heaviness or vice versa. The details of the experiment are described below.

METHOD

APPARATUS

1. Inertia

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 with 12 holes drilled symmetrically at different positions along its length was glued at the end of the shaft. The hole positions were selected so that the inertia increased by equal ratios (about 1.5 times) when a pair of 2-lb weights was moved from 2 holes equidistant from the center point of the 24" length to the adjacent holes. The 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, as shown in Table 1.

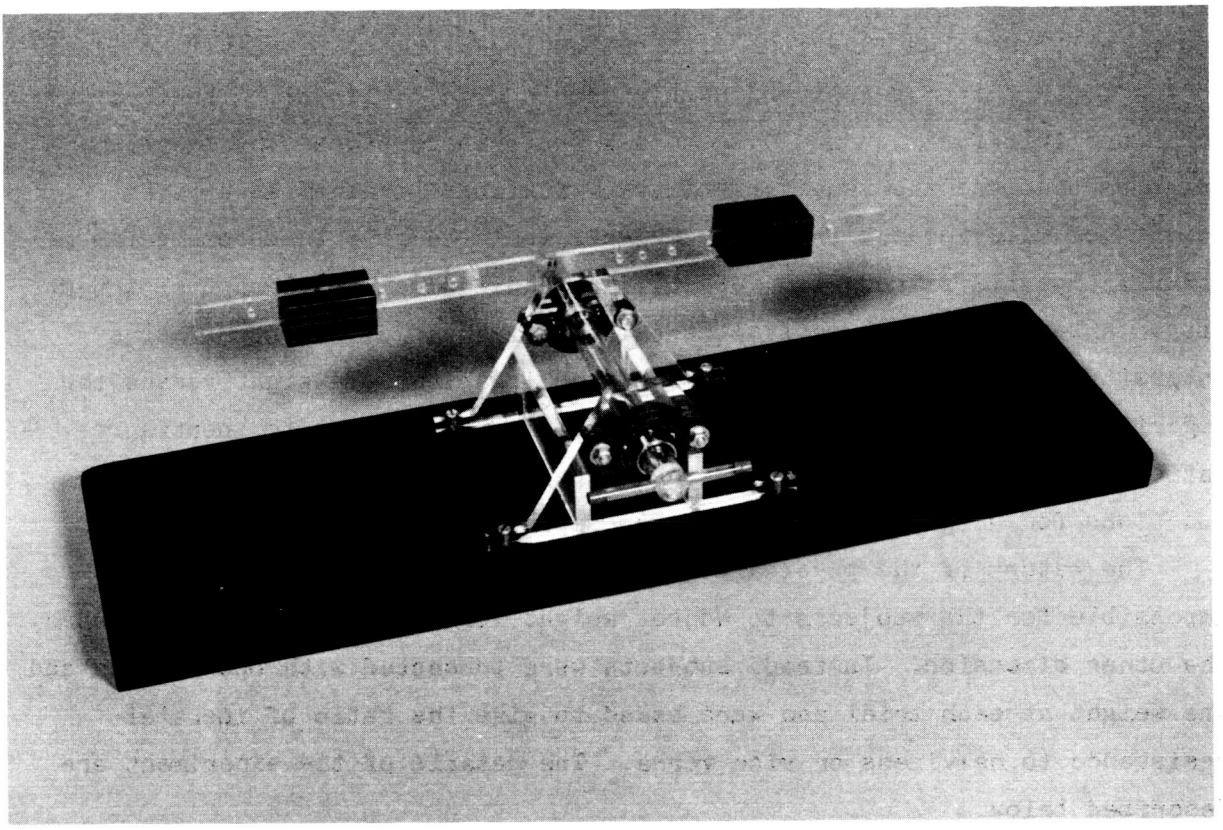


Figure 1. Mass Moment of Inertia Apparatus

Table 1.
Inertia Values of Different 2-lb. Weight Positions

Weights in holes number	Inertia (inch lb second^2)
1	0.15
2	0.20
3	0.29
4	0.43
5	0.70
6	1.15

2. Weights

A set of twelve weights was constructed by filling plastic bottles with lead shot embedded in wax. The bottles were painted white and appeared identical. The weight in each bottle is shown in Table 2.

Table 2. Weights of Twelve Bottles (lb, oz)

Bottle	Weight	Bottle	Weight
1	104	7	795
2	235	8	915
3	340	9	1032
4	455	10	1151
5	570	11	1255
6	690	12	1365

EXPERIMENTAL DESIGN

A preliminary trial conducted by the authors showed that the heaviest weights need not be compared to the smallest inertias. Each moment was, therefore, presented to subjects along with a band of weights: for example, moment 1 (2-lb weights in hole 1) was presented with weights 1 through 7, moment 2 with weights 2 through 8, up to moment 6 with weights 6 through 12. Each subject made 42 judgments (7 weights x 6 inertias). Presentation of inertia-weight combinations was random, and whether weight or inertia was presented first was alternated. Subjects 1,3,5, etc. were asked to give the ratio of weight/inertia; subjects 2,4,6, etc., gave the ratio inertia/weight.

SUBJECTS

Subjects were 13 volunteers (7 male, 6 female) either employees, family members, or friends of employees of Oak Ridge Associated Universities.

INSTRUCTIONS TO SUBJECTS

Subjects sat so they could turn the inertia handle comfortably with their preferred hand, maintaining whichever hand and grip they chose. Subjects were asked to turn the handle as long as they wished from side to

side and to concentrate on the feeling of "top heaviness" or "wanting to keep going" when they changed the direction of rotation. The part of the apparatus manipulated to change inertia was hidden by a screen to avoid giving the subject visual cues. Subjects were then asked to let go of the handle and to hold the weighted bottle in the same hand for as long as they wished and to concentrate on the feeling of weight. They were then asked to state the ratio they felt best described the relationship between the sensation of weight and the sensation of inertia. Subjects were allowed to practice the task until they felt comfortable with it.

DATA ANALYSIS

Our general procedure was to use regression analysis to determine the weight which was perceived as equal in magnitude to each inertia. Then weight-inertia pairs were plotted and their functional relationship determined.

The first step was to break subjects into four groups based on the form of the ratio they were asked to give (weight to moment of inertia, moment of inertia to weight) and sex (male, female). Next, the geometric mean of the judgments for each of the 7 weight x 6 inertia combinations was calculated. The weight judged equivalent to a given moment of inertia was determined by regression analysis with the logarithm of weights as the independent (x-axis) variable and the logarithm of the geometric means of judgments as the dependent (y-axis) variable.

Six regressions were performed, one for each moment of inertia. The weight considered equivalent to a moment of inertia was the exponent of the point on the x-axis associated with the point on the y-axis at which the ratio of judgments was equal to one, i.e., the point at which the logarithm of the ratio of judgments was equal to zero. Finally, the logarithms of the equivalent weights were plotted with the logarithms of the moments of inertia and a regression line was calculated.

RESULTS

Preliminary analysis indicated that judgments of weight to inertia and judgments of inertia to weight were equivalent. Moreover, judgments were equivalent at all levels of moment of inertia. There were small differences

between males and females, which will not be discussed here.

The relationship between the logarithm of perceived heaviness and the logarithm of weight was linear with a slope of 1.04 across all levels of moment of inertia (Figure 2). Thus, a 10-fold increase in weight results in a 10.4-fold increase in perceived heaviness. This is somewhat at variance with Stevens's (1975) results. He found exponents (slopes) of about 1.5, indicating that a 10-fold increase in weight results in a 15-fold increase in perceived heaviness. The reason for this difference is unclear--perhaps the pairing of weights and inertias somehow affects the exponent. The relationship between the logarithm of perceived inertia and the

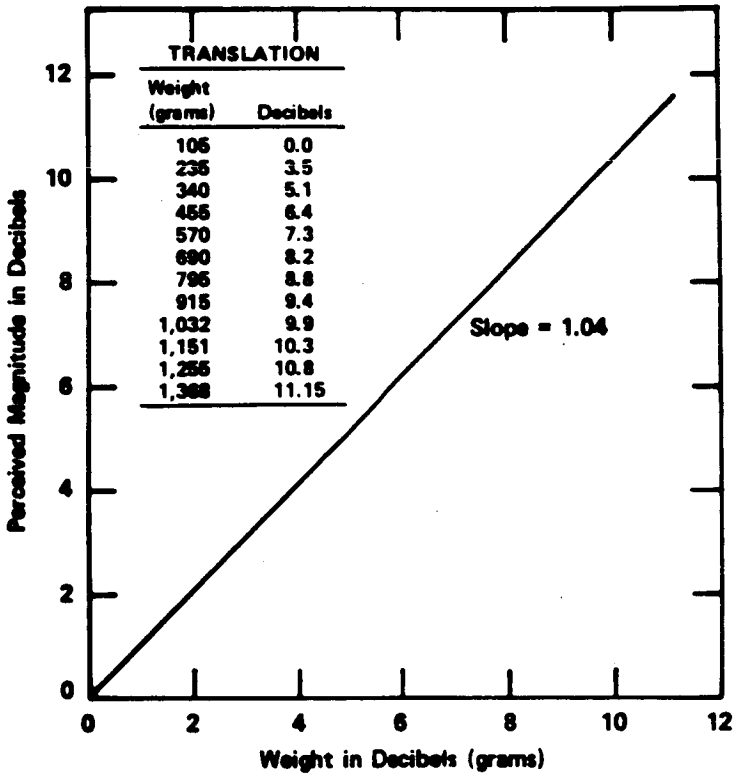


Figure 2. Relationship of perceived heaviness to weight.

logarithm of the physical inertia was also linear across all conditions with an exponent of 0.85 (Figure 3). Thus a 10-fold increase in inertia results in an 8.5 increase in perceived inertia. This is similar to our previously reported work (Clarke, et al. 1982).

The relationships between perceptual magnitude and physical intensity are power functions for both weight and inertia. Therefore, it was possible to define a power function between weight and inertia or a linear function between log weight and log inertia. The weight judged equivalent to each inertia was derived by regression analysis. Figure 4 shows the plot of weight-inertia pairs. In Table 3, the equivalent weights/inertias

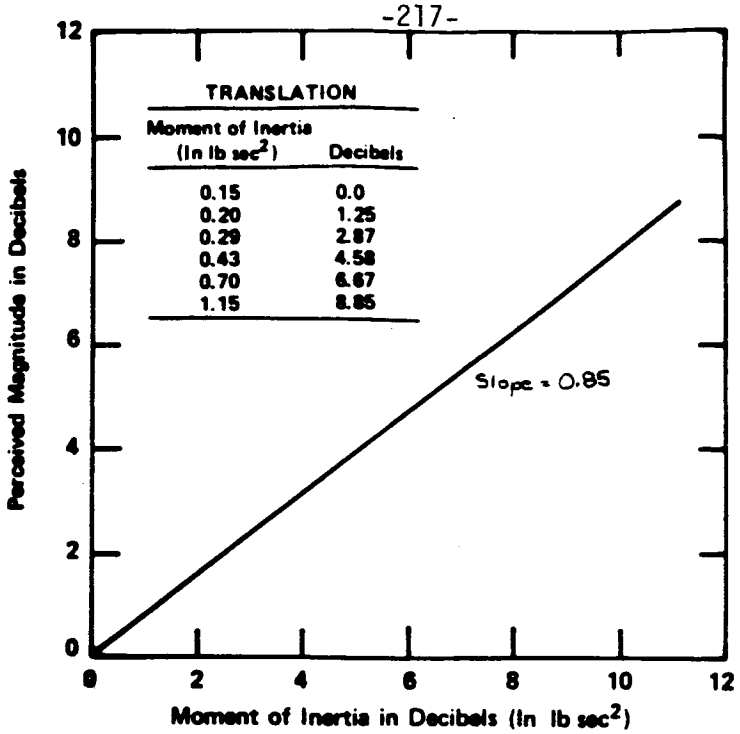


Figure 3. Relationship of perceived inertia to inertia.

are listed separately for males and females. The slope of the line for all subjects is 0.74 indicating that a 10-fold increase in inertia is perceptually equivalent to a 7.4 increase in weight. It is interesting to

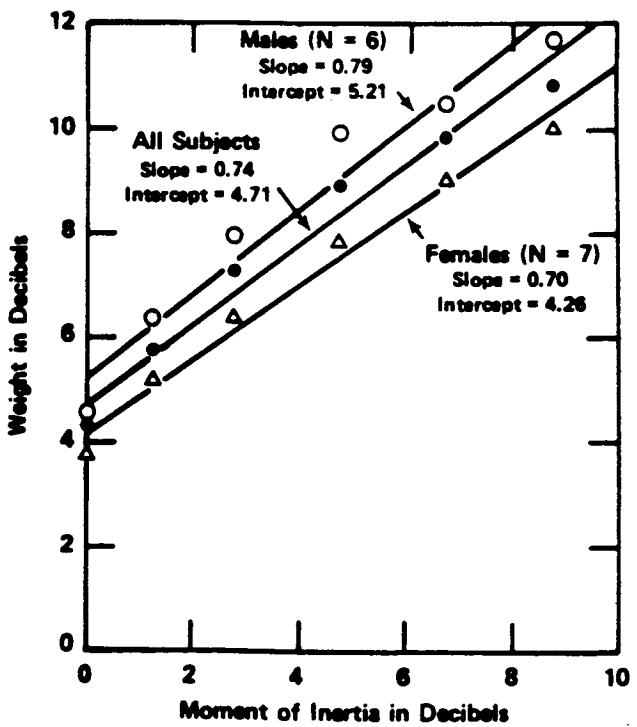


Figure 4. Relationship of weight to perceived inertia

note that this obtained slope (0.74) is less than the predicted value, which is the exponent for inertia divided by the exponent for weight or 0.85/1.04 = 0.82.

Table 3. Weights judged equivalent to the moments of inertia by males and females

MOI:		0.15	0.20	0.29	0.43	0.70	1.15
WEIGHT:	Males	0.57	0.76	1.00	1.45	1.83	2.30
	Females	0.64	1.00	1.42	2.25	2.52	3.33

The data analysis also suggest that females (exponent 0.70) are less sensitive to changes in inertia than males (exponent 0.79). The weight perceived as equivalent to a given inertia by males was heavier than that perceived as equivalent to the same inertia by females. This confirms an earlier study (Kreifeldt and Chuang 1979) showing that females have a greater Weber ratio for inertia than males.

DISCUSSION

The regression line in Figure 4 and the weight/inertia equivalents in Table 3 allow extrapolation to other values. We can make some interesting comparisons using the weights and inertias for common household objects calculated by Kriefeldt and Morse (1983). Lifting a one-pound weight would probably be perceived somewhat more intensely than rotating a blow drier horizontally (0.19 inch lb second^2), and somewhat less intensely than swinging a claw hammer (0.33 inch pounds second^2). Holding a 2.5-pound weight would be equivalent to rotating a cross end saw (0.61 inch lb second^2).

The various exponents already discussed suggest that people are less sensitive to changes in inertia than they are to changes in weight. For example, a 7.4-fold increase in weight has as its equivalent a 10 fold increase in inertia. Perhaps operators of manual control systems find changes in inertia less noticeable and therefore less fatiguing than changes in weight. As a result, accurate weight feedback may be more important than inertia feedback in control systems where an electric actuator replaces the mechanical link between the physical control device and the controlled object.

REFERENCES

Clarke, M.M., Handel, S.J., and Kreifeldt, J. G. Cross modality scaling: Mass moment and moment of inertia. Proc. 18th Annual Conference on Manual Control, 1982, in press.

Knowles, W. B. and Sheridan, T. B., The 'feel' of rotary controls: Friction and inertia, Human Factors, 1966, 8, 209-215.

Kreifeldt, J. G., and Chuang, M.C. Moment of inertia, psychophysical study of an overlooked sensation. Science, 1979(a), 206, 588-590.

Kreifeldt, J. G. and Chuang, M.C. Multidimensional studies of a tool "feel" sensation - moment of inertia. Proc. of the 23rd meeting of the Human Factors Society, 1979(b), 97-100.

Kreifeldt, J. G. and Morse, R. Moment of inertia of hand held objects and their handling properties. Proc. 19th Annual Conference on Manual Control, 1983, in press.

Pew, R. and Chananie, J. D. Subjective scaling of springs, shock absorbers, and fly wheels. Proc. 5th Annual NASA-University Conference on Manual Control, 1969, NASA SP-215, 471.

Stevens, S. S. Psychophysics. New York: John Wiley & Sons, 1975.

Zahner, C. M. and Kaminaka, M. S. Can a tool user sense "moment of inertia"? A human factors engineering study. 1982 Winter Meeting, American Society of Agricultural Engineers, paper 82-1613.

MOMENT OF INERTIA OF HAND HELD OBJECTS AND THEIR HANDLING PROPERTIES

J. G. Kreifeldt Dept. of Engineering Design
R. Morse Tufts University
Medford, MA

19th Annual Conference on Manual Control
May 23-25, 1983
MIT - Cambridge, MA

ABSTRACT

The moment of inertia (I) of an object defines its resistance to rotational acceleration. As such it is an important determiner of the "feel" of the object. Thus this property is important in areas such as hand controller design, consumer product design, remote manipulation and robotics. Because there is little experience based intuition of I and lack of data, the moment of inertias for approximately 80 common hand held objects were measured and are given herein. In addition, the predictive power of 3 simple models for analytic determination of I from mass and length measures are compared. A linear regression equation proved to be most predictive. This equation which accounts for approximately 75% of the initial I variance provides a simple first estimate of I based on easily obtained measures.

A summary of present knowledge of moment of inertia results in human factors is also given.

INTRODUCTION AND PREVIOUS WORK

The moment of inertia (I) of a body defines its resistance to rotational acceleration, just as the mass (M) defines its resistance to linear acceleration. The two defining equations are:

$$I = \frac{T}{\alpha}$$

T = Torque
α = angular acceleration
I = Moment of inertia

$$M = \frac{F}{a}$$

F = linear force
a = linear acceleration
M = mass

However, unlike mass the moment of inertia depends about the location and orientation of the rotational axis with respect to the body and thus there is no single value for I. It is common to specify I referenced to the center of gravity for a "hand hold" point on the object with the rotation occuring in a plane that makes sense for the object. For example, a baseball bat is not usually spun (like a lathe) and so I about an axis colinear with the long axis of the bat may have less use than the I value in the plane of swing of the bat.

The computation and/or measurement of I is fairly complex although available in standard sources[1, 2]. Basically I is a dynamic property while

M in our gravitational field is a static one as shown by measurement methods. Perceptually, weight and balance (CG) of an object held in the hand may be perceived without moving the object but sensing its I value requires rotation. The influence of moment of inertia on (human)manipulation performance and as a determiner of the object's "feel" continues to receive attention.

The units of measurement for I remain a continuing source of confusion as different investigators use different meanings for the ounce, pound, gram etc. that appears in the unit. Generally all investigators will mean (e.g.) oz weight or force. The moment of inertia may be reported as (e.g.) ounce (weight) - in - sec^2 in an attempt to conform to I defined as Torque (inch-ounce) divided by rotational acceleration (radian/sec^2). Or I may be reported as (e.g.) ounce - in^2 in a halfhearted attempt to conform to I defined by mass x $(distance)^2$ only this time simply using weight units for mass units.

A conversion table for convenience follows:

GIVEN:	MULTIPLY BY	TO GET
$lb_{(weight)} - in^2$	$\frac{1}{(32.2)(12)}$	$lb - in - sec^2$
$lb_{(weight)} - in^2$	$\frac{16}{(32.2)(12)}$	$oz - in - sec^2$

Weber fractions for I have been obtained from several sources. Knowles & Sheridan[3] obtained Weber fractions on the order of 10 to 20 percent for finger manipulation of rotary knobs with I values 0.00730 to 0.1139 in-oz-sec^2. Kreifeldt and Chuang[4] report Weber fractions of about 20 to 30 percent for baton-like objects with I values from 0.0011 oz-in-sec^2 to 11.03 oz-in-sec^2. These Weber fractions were obtained for an I value range of ten thousand to one and for objects requiring somewhat different manipulation but are in reasonable agreement suggesting that human sensitivity to I as measured by the Weber ratio is on the order of 20%. That is, two values of I would have to differ by 20% before the difference could begin to be reliably detected. It is common to separate values by 3-5 times the Weber ratio for reliable and consistent differentiation. Thus two values of I would have to differ by 60% or more for reliable differentiation. By comparison, the Weber ratio for lifted weight is about 3% indicating that a human's ability to perceive a weight difference is 5-10 times better than the ability to perceive I differences.

This should not be confused, however, with the statement that the human is more (or less) sensitive to weight than moment of inertia. In fact, Kreifeldt and Chuang[5], using multidimensional scaling techniques, report that for particular values of weight and moment of inertia, subjects perceptually weighted I considerably more than W when determining similarity of "feel" between two baton-like objects. Physically, it is not possible to say whether or not the objects had more I than W since the two quantities are incommensurable.

Weber fractions provide design information of the "tolerance" nature (e.g. how much can a value be changed before the change is noticeable). However, it is also important to know how much must a value be changed to

make it seem (e.g.) twice as great. Standard psychophysical scaling techniques such as magnitude estimation [6] have been applied to the problem of a subject's ability to scale moment of inertia. This technique assumes that sensation(S) can be related to the physical intensity of the stimulus (I) through a power function:

$$S = kI^{\beta}$$

Where β is the experimentally determined exponent governing the rate of growth of sensation with growth of the stimulus intensity. Pew and Chananie[7] in a subjective scaling study of moment of inertia of a side arm control stick obtained exponent values of 1.27 and 1.82 (in a repeat experiment) over an I range of 14.13 to 35 in-oz-sec^2 suggesting that the sensation of I grows faster than the physical value itself. As an example, if the exponent were 1.8, then an object having a moment of inertia twice that of another would feel as though the difference were greater by a factor of 3.48 ($=2^{1.8}$). Pew and Chananie also found that the exponent was affected by the friction and viscous damping of the side arm controller. Clarke, Handel and Kreifeldt[8] obtained a value of 1.06 over an I range from 2.40 to 18.40 in-oz-sec^2 for a low friction rotary device grasped by a tee-shaped handle and rotated by a supination-pronation motion of the forearm. An exponent of unity suggests that doubling the I value doubles the sensation.

Handel[9] also obtained β exponents using the same device but with the arm held in different positions and either manipulating the device directly or via a force feedback remote manipulator. In this case Handel obtained a β value of 0.83 for direct manipulation and 1.09 for remote manipulation suggesting that a remote manipulator has some intervening effect on the sensation. Handel also found that moment of inertia with the wrist flexed was judged greater (67% of the time) than with the wrist straight or with the shoulder extended.

Zahner and Kaminaka[10] obtained subject(s) ratings of the relative effort needed to swing baton-like "tools" of different I values from 0.99 to 9.81 in-oz-sec^2 (at 1.33 lb). Subject effort was observed to grow as the square root of I. Unfortunately the CG also varied so it is not strictly possible to ascribe the change as related to I only, CG only or some combination.

Knowles and Sheridan[3] also demonstrated that subjects preferred a small amount of inertia to none for rotary controls while Kreifeldt and Chuang[5] found that subjects preferred the feeling of small values of mass and moment of inertia over large values in simulated tennis racquets.

While one may have an intuitive notion of how much an object "weighs" through common experience and its effect in terms of lifting though maximum force tables for humans, less experience is available for I and torque. Literature[12] suggests that a male can exert a maximum torque on the order of 10 ft.-lb. (about 2000 in-oz). In as much as we do not exist in a uniform rotational acceleration field, if we assume a rotational acceleration of 10 rad/sec^2, then the maximum I value would be 200-in-oz-sec^2. Or conversely, if an object has an I value of 100 in-oz-sec^2, then the rotational acceleration maximum would be about 20 radian/sec^2.

As a summary of knowledge to date.

(1) Weber ratios for moment of inertia are on the order of 20% (± 10%)

(2) exponent values (β) are on the order of 1.2 (±.40) and may depend on other properties of the body

(3) humans can discern moment of inertia as distinct from mass although they may be unaware of this

(4) perceived effort in swinging a "tool" increased with the square root of its moment of inertia

(5) humans have a definite preference for select values of moment of inertia of manipulated objects.

Purposes of the Study

The purposes of this study follow from the knowledge acquired regarding moment of inertia ranging from sensitivity to preference and the need to supply practical information about this design parameter.

(1) The moment of inertias for a number of man-made hand held objects have been determined and compared in order to identify any implicit trends in I that might exist.

(2) Simple calculation models for I based on easily measured parameter values have been tested against the measured I values. Analytical calculation of I for a realistic object is impractical if possible at all, and measurement of I is complex and not easily made for large objects.

MOMENTS OF INERTIAS OF COMMON OBJECTS

Apparatus

The moment of inertia of a body can only be calculated geometrically for an exceptionally simple object. This approach is impractical for any realistic object such as a hand drill.

Measurement techniques exist which depend upon determining the object's period of oscillation when actuated in a pendulum mode. For this study, a torsional pendulum was used to determine the I value(s) for an object because of its flexibility. Since the object need only be affixed to the pendulum plate, objects of widely varying shapes and orientations could be measured without modification to the apparatus.

The torsional pendulum consisted of a disk of 1/8" foamcor, 11.66" in diameter and weighing .11 lbs. The CG of the plate was calculated and marked on its surface and 3 equal angle diameter lines were drawn through the CG point to the edge of the plate. Three 74.9" lengths of wire were attached to the

edge of the plate by cotter pins and spaced 120 degrees apart. The plate was suspended by its wires from a wooden support so that the wires were perpendicular to the foamcor plate. Each wire was adjusted for length by a turnbuckle on the top of the wooden support. Two 7" x 1/4" foamcor strips were attached perpendicular to the edge of the plate and 180° apart so that they hung to 1/2" above the floor. A piece of fiberboard with a 11.66" diameter circle marked in 5 degree graduations rested on the floor directly below the perdulum plate. (See Figure 1)

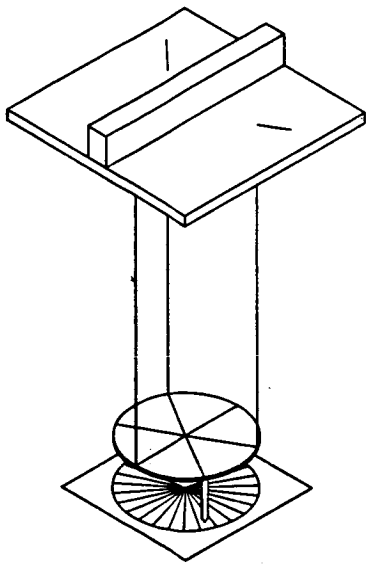


Figure 1 - Torsional Pendulum Constructed to Measure Moment of Inertia.

If the plate is rotated through a small angle, released and allowed to oscillate, its moment of inertia may be calculated from[2]:

$$I_o = \frac{T^2}{4\pi^2}\left(w\frac{r^2}{l}\right)(326.4)$$

where: I_o = moment of inertia of the plate (lb-in^2)

w = weight of the object (lb)

r = radius of the plate (in)

l = length of the suspension wires (in)

T = period of oscillation (sec)

I_o for the pendulum was measured as 2.07 lb-in^2.

If an object is placed on the plate so that its center of gravity is above the CG of the plate and the period of oscillation again measured, the I value for both object and plate is determined. The I value for the object can be found by subtracting that for the plate alone.

Torsional Pendulum Calibration

The torsional pendulum was evaluated for accuracy by comparing measured and calculated moments of inertia for a number of uniform metal rods. The percentage error based on the calculated value was less than 4% except for calculated values of I less than 0.0621 lb-in^2 where the error was less than 11%.

The center of gravity of the body was determined either by finding its balance point on a knife edge or by the plumb line method.

Methodology

The following properties were measured for each object:

(1) length along three principal axes
(2) weight
(3) center of gravity
(4) location of a likely grip point
(5) moment of inertia (lb-in^2)

The principal axes were defined on the body in accordance with the most likely axes of rotation and depended upon the object. Location of the most likely grip point was defined as the distance from one end of the object to the point most likely grasped when the object is manipulated. For example, the hand hold point (H) on a hammer was defined as a distance equal to 1/2 of the 50th% tile male handbreadth from the end of the handle. For some objects (e.g. a knife) H was defined as the center of the handle. Some objects have no obvious grasping point so definition of H was more arbitrary.

Object Measured

A wide variety of man-made, hand held objects were chosen and the moment of inertia measured along one, two or three principal axes of each. The eighty objects selected are listed in Table 1.

For each object the moments of inertia were measured about the CG (I_{CG}) and the "hand hold" point (I_H). In addition, the following measurements were also taken: (See Figure 2.)

(1) L_A - maximum object length along an axis perpendicular to the axis of rotation
(2) L_B - minimum object length along axis perpendicular to axis of rotation and perpendicular to L_A
(3) L_C - object length along axis of rotation
(4) A - approximate rotated area ($L_A \times L_B$)
(5) V - approximate volume of object ($L_A \times L_B \times L_C$)
(6) W - weight of the object.

In the following, I is given in lb-in^2 for convenience of measurement. Dividing by 386.40 will convert to in-lb-sec^2.

TABLE 1 THE EIGHTY OBJECTS SELECTED FOR MEASUREMENT

1 PHILLIPS SCREWDRIVER (Y)
2 CLAWHAMMER (X)
3 CALCULATOR (X)
4 NEEDLENOSE PLIERS (X)
5 UMBRELLA (X)
6 SKI POLE (X)
7 GLASS MUG (Z)
8 COFFEE MUG (Z)
9 MANUAL EGG BEATER (X)
10 CAN 3 TENNIS BALLS (Z)
11 CHOPPING KNIFE (X)
12 11OZ CAN SHAVING CREAM (Z)
13 1 1. BOTTLE GIN (Z)
14 12OZ CAN SODA (Z)
15A ELECTRIC DRILL (Y)
15B ELECTRIC DRILL (X)
16 STAPLE GUN (X)
17A 1 QT SAUCEPAN (Y)
17B 1 QT SAUCEPAN (X)
18A 3 QT SAUCEPAN (Y)
18B 3 QT SAUCEPAN (X)
19 BICYCLE PUMP (X)
20 CHINA DINNER PLATE (Y)
21 SOUP LADLE (Y)
22 SERVING SPOON (Y)
23 BUTTER KNIFE (X)
24 ICE CREAM SCOOP (Y)
25 CORKSCREW (Y)
26A ELECTRIC RAZOR (Y)
26B ELECTRIC RAZOR (X)
27 HAND BROOM (Y)
28 DUSTPAN (Y)
29 18OZ BOTTLE SHAMPOO (Z)
30 FLASHLIGHT (X)
31 12OZ AEROSOL CAN (Z)
32 22OZ GLASS CLEANER (Z)
33 32OZ ICE TEA MIX (Z)
34A BLOW DRIER (Y)
34B BLOW DRIER (X)
35 10" FRYING PAN (Y)
36 2 QT GLASS BOTTLE-EMPTY (Z)
37 2 QT GLASS BOTTLE-FULL (Z)
38A HARDCOVER BOOK 1 (Y)
38B HARDCOVER BOOK 1 (X)
38C HARDCOVER BOOK 1 (Z)
39 SOFTCOVER BOOK 2 (Y)
40 HARDCOVER BOOK 3 (Y)
41A HARDCOVER BOOK 4 (Y)
41B HARDCOVER BOOK 4 (X)
41C HARDCOVER BOOK 4 (Z)
42A SOFTCOVER BOOK 5 (Y)
42B SOFTCOVER BOOK 5 (X)
42C SOFTCOVER BOOK 5 (Z)
43 HARDCOVER BOOK 6 (Y)
44 SOFTCOVER BOOK 7 (Y)
45 FLAT HEAD SCREWDRIVER (X)
46A SOLDERING IRON (X)
46B SOLDERING IRON (Y)
47A ELECTRIC MIXER (Y)
47B ELECTRIC MIXER (X)
47C ELECTRIC MIXER (Z)
48B PAINTBRUSH (Y)
48C PAINTBRUSH (X)
49A PLIERS (Y)
49B PLIERS (X)
50A PUTTY KNIFE (Y)
50B PUTTY KNIFE (X)
51A FILE (Y)
51B FILE (X)
52A VISE GRIP PLIERS 10R (Y)
52B VISE GRIP PLIERS 10R (X)
53A VISE GRIP PLIERS 7R (Y)
53B VISE GRIP PLIERS (X)
54 CHISEL (Y)
55A 3/4-7/8 WRENCH (Y)
55B 3/4-7/8 WRENCH (X)
56A 5/8 WRENCH (Y)
56B 5/8 WRENCH (X)
57A TIN SNIPS (Y)
57B TIN SNIPS (X)
58A 15/16-1 WRENCH (Y)
58B 15/16-1 WRENCH (X)
59A METAL 20" RULER (Y)
59B METAL 20"RULER (X)
60A 1 M METAL RULER (X)
61A STAPLER (Y)
61B STAPLER (X)
62A CAN CAR WAX (X)
62B CAN CAR WAX (Z)
62C CAN CAR WAX (Y)
63 HAND HELD BRUSH (Y)
64 TAPE DISPENSER (Y)
65 CARDBOARD TUBE CAN (Z)
66A CARPENTERS LEVEL (X)
66B CARPENTERS LEVEL (Y)
67A BACK SAW (Y)
67B BACK SAW (X)
68A CROSSCUT SAW (Y)
68B CROSSCUT SAW (X)
69A HACKSAW (Y)
69B HACKSAW (X)
70 T-SQUARE (Y)
71A PORTABLE RADIO (Y)
71B PORTABLE RADIO (Z)
71C PORTABLE RADIO (X)
72A KNIFE (Y)
72B KNIFE (X)
73A SABRE SAW (Y)
73B SABRE SAW (X)
73C SABRE SAW (Z)
74A 16OZ CAN WAX (X)
74B 16OZ CAN WAX (Z)
74C 16OZ CAN WAX (Y)
75 SMALL CLAW HAMMER (X)
76 8OZ PLASTIC TUBE GEL (X)
77A POP RIVET GUN (Y)
77B POP RIVET GUN (X)
78A PENTAX CAMERA (Y)
78B PENTAX CAMERA (X)
78C PENTAX CAMERA (Z)
79A 1.5PT WATER BOTTLE (X)
79B 1.5PT WATER BOTTLE (Z)
79C 1.5PT WATER BOTTLE (Y)
80A 8" ADJUSTABLE WRENCH (Y)
80B 8" ADJUSTABLE WRENCH (X)

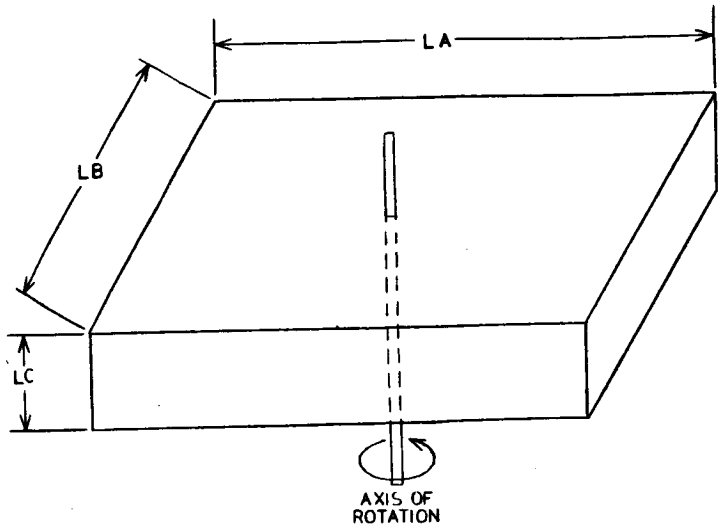


Figure 2 Definition of the Object Dimensions Taken

The mean standard deviation and range for all variables measured are given in Table 3.

Pearson Correlation Coefficients were calculated for I_{CG} and I_H with weight, L_A, L_B, L_C, A and V and are set out in Table 4 together with levels of significance. Notice that all correlation coefficients are highly significant except those for L_C, and L_B (with I_{CG}) and that correlation with L_A (maximum length transverse to the axis of rotation) is the highest. Correlation using the "area" or "volume" as defined is less than that obtained using L_A alone.

In order to develop a linear predictive model for I based on the 6 defined variables, a stepwise multiple regression was performed with the results given in Table 5.

It is clear that only L_A and weight (in that order) contribute significantly to predicting either I_{CG} or I_H for common hand held objects. It is somewhat surprising that weight adds so little (18%) to the predictive power of the regression equations and that the other variables do not add any significant amount. Weight and L_A were negatively and only slightly correlated ($r = -.1268$, $p = .079$) or essentially uncorrelated.

Predictive Models for I

Three simple models for predicting I_{CG} from simply made measurements of the object are compared.

(1) Linear regression model
(2) object geometrically modeled as a uniform long slender rod
(3) object geometrically modeled as a uniform rectangular slab.

TABLE 2 VALUES MEASURED FOR EACH OBJECT IN TABLE 1

OBJECT	ICG ($lb.-in.^2$)	IH ($lb.-in.^2$)	WEIGHT (lb.)	LA (in.)	LB (in.)	LC (in.)	AREA ($in.^2$)	VOLUME ($in.^3$)
1	2.50	3.90	0.30	10.30	1.10	1.10	11.30	12.50
2	19.90	126.50	1.50	13.00	5.40	1.30	70.20	91.30
3	0.50	0.50	0.20	5.50	0.50	3.00	2.80	8.30
4	1.60	3.90	0.40	8.10	2.00	0.50	16.20	8.10
5	46.10	268.40	0.80	27.80	1.80	1.80	50.00	90.10
6	162.90	423.40	0.50	53.50	1.00	0.80	53.50	42.80
7	4.00	26.30	1.10	5.50	3.20	4.50	17.60	79.20
8	1.70	17.80	0.70	3.40	3.40	4.50	11.60	52.00
9	8.20	33.10	0.90	11.90	2.90	2.50	34.50	86.30
10	4.00	4.00	0.60	8.10	2.90	2.90	23.50	68.10
11	8.30	13.30	0.50	15.40	2.00	0.80	30.80	24.60
12	2.30	2.30	0.90	6.40	2.50	2.50	16.00	40.00
13	23.30	23.30	3.20	11.60	3.30	3.30	38.30	126.30
14	1.10	1.10	0.90	4.80	2.50	2.50	12.00	30.00
15A	32.40	52.00	3.90	9.30	2.50	6.90	23.30	160.40
15B	26.30	45.80	3.90	9.30	6.90	2.50	64.20	160.40
16	15.60	27.10	2.30	7.10	5.30	1.20	37.60	45.20
17A	3.40	12.60	0.40	10.90	6.20	2.90	67.60	196.00
17B	2.90	14.10	0.40	10.90	2.90	6.20	31.60	196.00
18A	27.30	109.80	1.60	15.10	9.50	4.40	143.50	631.20
18B	21.50	122.00	1.60	15.10	4.40	9.50	66.40	631.20
19	12.20	38.70	0.40	16.90	1.20	1.20	20.30	24.30
20	15.00	32.20	1.40	9.10	9.10	0 90	82.80	74.50
21	3.20	9.70	0.30	11.30	3.10	3.20	35.00	112.10
22	1.70	6.30	0.20	11.80	2.40	1.90	28.30	53.80
23	1.10	1.10	0.20	9.30	0.80	0.20	7.40	1.50
24	0.70	1.80	0.10	7.90	2.30	1.30	18.20	23.60
25	0.50	0.60	0.20	5.10	0.30	3.10	1.50	4.60
26A	1.60	2.30	0.70	4.80	2.90	1.70	13.90	23.70
26B	1.40	2.20	0.70	4.80	1.70	2.90	8.20	23.70
27	7.90	18.70	0.60	13.40	3.50	1.70	46.90	79.70
28	14.70	44.70	0.90	12.50	11.80	4.90	147.50	722.80
29	3.60	3.60	1.20	6.30	3.00	3.00	18.90	56.70
30	3.20	3.20	0.70	8.40	2.50	2.30	21.00	48.30
31	4.40	10.10	0.90	8.50	2.40	2.40	20.40	49.00
32	9.80	57.00	1.60	10.50	2.30	4.20	24.20	101.40
33	6.00	6.00	1.60	5.60	4.00	4.00	22.40	89.60
34A	13.00	71.30	1.40	9.10	3.50	8.20	31.90	261.20
34B	12.80	29.70	1.40	9.10	8.20	3.50	74.60	261.20
35	31.90	122.30	1.40	16.70	10.10	2.20	168.70	371.10
36	19.20	63.40	1.70	9.50	4.80	4.80	45.60	218.90
37	30.10	187.50	6.10	9.50	4.80	4.80	45.60	218.90
38A	14.60	28.80	1.50	9.30	6.30	1.30	58.60	76.20
38B	10.30	10.30	1.50	9.30	1.30	6.30	12.10	76.20
38C	4.90	19.10	1.50	6.30	1.30	9.30	8.20	76.20
39	45.80	96.70	2.80	11.00	8.50	1.50	93.50	140.30
40	70.00	152.00	5.40	10.30	7.80	2.50	80.30	200.90
41A	10.80	21.60	1.30	8.50	5.80	1.10	49.30	54.20
41B	7.70	7.70	1.30	8.50	1.10	5.80	9.30	54.20
41C	3.60	14.50	1.30	5.80	1.10	8.50	6.40	54.20
42A	5.90	8.90	0.70	8.00	5.20	1.00	41.60	41.60
42B	4.40	4.40	0.70	8.00	1.00	5.20	8.00	41.60
42C	2.10	5.00	0.70	5.20	1.00	8.00	5.20	41.60
43	40.10	91.10	3.30	9.50	7.80	1.90	74.10	140.80
44	21.30	45.40	1.30	11.00	8.50	0.60	93.50	56.10
45	0.60	0.70	0.20	6.50	1.10	1.10	7.10	7.90
46A	12.20	39.50	3.00	7.40	6.20	2.20	45.90	100.90
46B	9.30	85.30	3.00	7.40	2.20	6.20	16.30	100.90
47A	6.90	33.60	1.90	7.80	2.80	4.90	21.80	107.00
47B	7.40	21.80	1.90	7.80	4.90	2.80	38.20	107.00
47C	2.80	17.20	1.90	4.90	2.80	7.80	13.70	107.00
48B	1.50	3.60	0.30	9.40	3.10	0.70	29.10	20.40
48C	1.30	3.40	0.30	9.40	0.70	3.10	6.60	20.40

Table 2 (Continued)

OBJECT	ICG ($lb.-in.^2$)	IH ($lb.-in.^2$)	WEIGHT (lb.)	LA (in.)	LB (in.)	LC (in.)	AREA ($in.^2$)	VOLUME ($in.^3$)
49A	3.30	6.80	0.60	8.10	0.50	1.90	4.10	7.70
49B	3.40	6.90	0.60	8.10	1.90	0.50	15.40	7.70
50A	0.90	1.30	0.20	7.60	1.30	0.70	9.90	6.90
50B	0.90	1.30	0.20	7.60	0.70	1.30	5.30	6.90
51A	4.10	12.80	0.50	12.70	0.90	0.20	11.40	2.30
51B	4.10	12.80	0.50	12.70	0.20	0.90	2.50	2.30
52A	6.60	12.80	1.20	8.90	0.80	2.50	7.10	17.80
52B	7.30	13.50	1.20	8.90	2.50	0.80	22.30	17.80
53A	2.90	4.90	0.80	7.30	0.60	2.20	4.40	9.60
53B	3.10	5.00	0.80	7.30	2.20	0.60	16.10	9.60
54	0.80	1.40	0.20	7.10	0.80	0.80	5.70	4.50
55A	5.00	6.20	0.60	9.50	0.30	1.80	2.80	5.10
55B	5.20	6.30	0.60	9.50	1.80	0.30	17.10	5.10
56A	0.90	1.20	0.20	7.00	0.30	1.30	2.10	2.70
56B	0.90	1.20	0.20	7.00	1.30	0.30	9.10	2.70
57A	6.00	14.90	0.80	10.00	0.90	3.80	9.00	34.20
57B	6.60	15.50	0.80	10.00	3.80	0.90	38.00	34.20
58A	21.80	43.40	1.00	15.50	0.60	1.40	9.30	13.00
58B	21.80	43.40	1.00	15.50	1.40	0.60	21.70	13.00
59A	6.70	6.70	0.20	20.00	1.20	0.10	24.00	2.40
59B	6.80	6.80	0.20	20.00	0.10	1.20	2.00	2.40
60A	41.10	41.10	0.30	39.60	0.10	1.20	4.00	4.80
61A	0.80	0.80	0.30	5.50	1.30	1.80	7.10	12.90
61B	0.80	0.80	0.30	5.50	1.80	1.30	9.90	12.90
62A	1.50	1.50	0.60	5.30	1.50	2.80	8.00	22.30
62B	1.80	1.80	0.60	5.30	2.80	1.50	14.80	22.30
62C	0.70	0.70	0.60	2.80	1.50	5.30	4.20	22.30
63	1.40	3.20	0.20	10.00	2.90	1.30	29.00	37.70
64	8.60	8.60	2.30	6.30	2.50	2.70	15.80	42.50
65	2.30	2.30	0.40	8.00	2.40	2.40	19.20	46.10
66A	34.80	34.80	1.00	18.00	2.20	1.00	39.60	39.60
66B	35.30	35.30	1.00	18.00	1.00	2.20	18.00	39.60
67A	21.30	81.90	0.90	18.50	0.80	5.50	14.80	81.40
67B	23.20	83.80	0.90	18.50	5.50	0.80	101.80	81.40
68A	99.40	234.00	1.80	29.50	0.90	6.20	26.50	164.60
68B	4.80	239.40	1.80	29.50	6.20	0.90	182.90	164.60
69A	25.30	63.60	0.80	18.00	1.20	4.40	21.60	95.00
69B	26.00	64.30	0.80	18.00	4.40	1.20	79.20	95.00
70	29.70	29.70	0.40	26.30	12.30	0.50	323.50	161.70
71A	19.60	58.00	2.70	8.80	2.10	4.80	18.50	88.70
71B	26.50	64.90	2.70	8.80	4.80	2.10	42.20	88.70
71C	8.10	46.50	2.70	4.80	2.10	8.80	10.10	88.70
72A	1.40	2.70	0.10	11.40	0.40	1.00	4.60	4.60
72B	1.40	2.70	0.10	11.40	1.00	0.40	11.40	4.60
73A	14.40	42.70	3.40	7.00	2.40	6.00	16.80	100.80
73B	23.90	52.20	3.40	7.00	6.00	2.40	42.00	100.80
73C	12.00	40.30	3.40	6.00	2.40	7.00	14.40	100.80
74A	4.60	4.60	1.30	5.80	1.80	3.70	10.40	38.60
74B	5.80	5.80	1.30	5.80	3.70	1.80	21.50	38.60
74C	2.10	2.10	1.30	3.70	1.80	5.80	6.70	38.60
75	3.50	17.00	0.30	10.50	3.10	0.70	32.50	22.80
76	2.10	2.10	0.50	7.90	3.00	1.90	23.70	45.00
77A	5.30	16.80	0.80	8.90	0.80	4.50	7.10	32.00
77B	6.40	17.90	0.80	8.90	4.50	0.80	40.00	32.00
78A	5.60	5.60	2.10	5.50	3.80	3.70	20.90	77.30
78B	5.50	5.50	2.10	3.80	3.70	5.50	14.10	77.30
78C	5.40	5.40	2.10	5.50	3.70	3.80	20.30	77.30
79A	7.40	7.40	2.10	8.20	2.50	3.70	20.50	75.90
79B	6.70	6.70	2.10	8.20	3.70	2.50	30.30	75.90
79C	0.90	0.90	2.10	3.70	2.50	8.20	9.30	75.90
80A	2.80	8.30	0.50	8.30	0.50	2.10	4.10	8.70
80B	2.80	8.30	0.50	8.30	2.10	0.50	17.40	8.70

TABLE 3 STATISTICS FOR THE 80 OBJECTS

VARIABLE	DIMENSION	MEAN	S.D.	RANGE
I_{CG}	lb - in^2	12.36	19.77	0.50-162.90
I_H	lb - in^2	34.65	59.48	0.50-423.40
Weight	lb	19.52	17.28	0.10-6.10
L_A	inch	10.44	6.88	3.40-53.50
L_B	inch	3.00	2.49	0.20-12.30
L_C	inch	2.87	2.27	0.20-9.50
Area	$(inch)^2$	32.18	41.82	1.50-323.50
Volume	$(inch)^3$	79.07	111.70	1.50-722.80

TABLE 4 PEARSON CORRELATION COEFFICIENTS, (r) SIGNIFICANCE LEVEL AND CONFIDENCE INTERVALS FOR I_{CG}, I_H

	I_{CG}			I_H		
VARIABLE	r	p	95% INTERVAL	r	p	95% INTERVAL
Weight	.3167	0.000	.15 - .47	.3415	0.000	.18-.49
L_A	.7402	0.000	.65 - .81	.7156	0.000	.62-.79
L_B	.2054*	0.11	.03 - .37	.2372	0.004	.06-.40
L_C	-.0028*	0.49	-.18 - .17	.0691*	0.220	-.11-.24
Area	.3111	0.000	.14 - .46	.3956	0.000	.24-.55
Volume	.2432	0.003	.07 - .40	.3617	0.000	.20-.51

*Coefficient may be taken as 0.000.

TABLE 5 STEPWISE MULTIPLE REGRESSION RESULTS

VARIABLE	I_{CG}			I_H		
	R	R^2	ΔR^2	R	R^2	ΔR^2
L_A	.7402	.5478	----	.7156	.5121	----
Weight	.8480	.7191	.1713	.8378	.7020	.1889
L_B	.8480	.7191	.0000	.8456	.7151	.0131
L_C	.8483	.7196	.0005	.8469	.7172	.0021
Volume	.8490	.7208	.0012	.8473	.7180	.0008
Area	.8686	.7550	.0338	.8473	.7180	.0000

(a) The linear regression models for I_{CG} and I_H have been determined to be:

$$I_{CG} = 2.28L_A + 7.67 \text{ Weight} - 20.76 \qquad (R^2 = .720)$$

$$I_H = 6.67L_A + 24.29 \text{ Weight} - 64.51 \qquad (R^2 = .70)$$

L_A = inch

Weight = lb

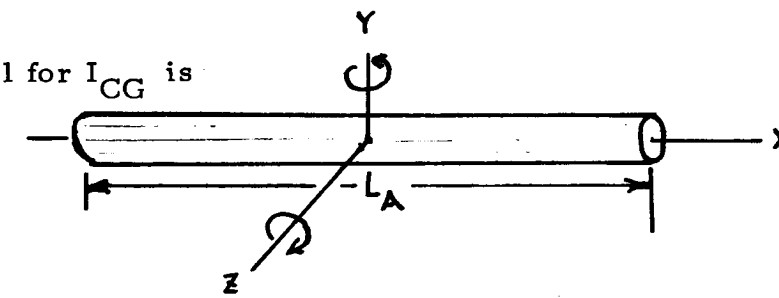
I = lb - in^2

In using these equations, any value predicted to be less than zero should automatically be set to zero.

(b) The long slender rod model for I_{CG} is

$$I_{CG} = \frac{1}{12} W \cdot L_A^2$$

W = lb
L_A = in
I = lb - in^2



(c) The rectangular slab model for I_{CG} is

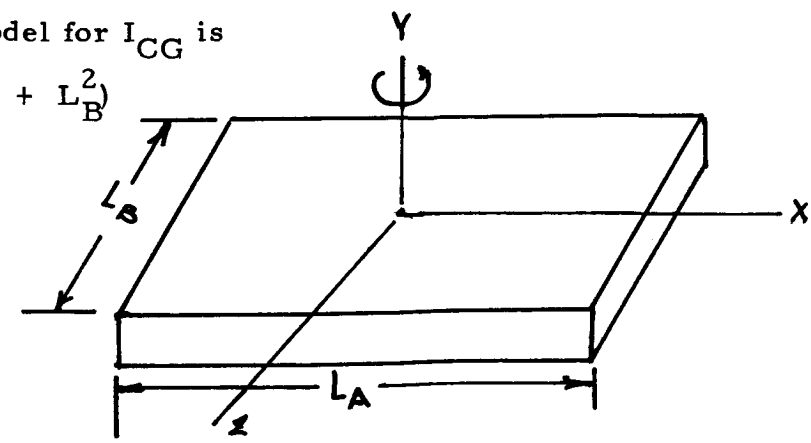
$$I_{CG} = \frac{1}{12} W (L_A^2 + L_B^2)$$

L_A = inch

L_B = inch

W = lbs

I = lb-in^2



A comparison of the three models is given in TABLE 6 in terms of the multiple R^2 values. Only I_{CG} predictions are compared and any negative values from the linear model are set equal to zero.

TABLE 6 COMPARISON OF THREE PREDICTIVE MODELS FOR I_{CG}

MODEL	EQUATION ($lb - in^2$)	R^2
LINEAR	$I_{CG} = 2.28L_A + 7.67W - 20.76$	.72
ROD	$I_{CG} = \frac{1}{12} W \cdot L_A^2$	.634
SLAB	$I_{CG} = \frac{1}{12} W(L_A^2 + L_B^2)$	.653

Thus, the linear predictive equation which does not resemble any of the geometric models is a better predictor of I_{CG} than modeling the object as either a long slender rod or as a slab.

DISCUSSION & CONCLUSIONS

The regression analysis approach provides a fairly good first estimate for the value of I_{CG} or I_H of common hand held objects. The maximum length of the object transverse to the axis of rotation accounts for the largest amount of variance followed by the weight. This latter fact is reasonable when it is remembered that I and W are theoretically independent in the sense that W can be varied independently of I by readjusting the geometry of the object. On the other hand, if the geometry of an object is fixed, then W and I become partially dependent. Across all 80 objects measured, there is a correlation of about 0.3 (10% of the variance) between W and I indicating a weak relationship between the two basic parameters of a body needed to explain its motion by Newtonian physics. I is more strongly related to a geometric property (L_A) of the body than to its weight which is reasonable.

The data presented in TABLE 2 as well as the prediction equation should add to our intuitive and practical understanding of direct or remote manipulation of hand held objects. The data and equation should also supply information useful in robotics where torques encountered in object rotation must be considered.

REFERENCES

1. F. Beer, E. Johnson, Jr., Mechanics for Engineers - Statics McGraw-Hill Book Co., Inc. N.Y. 1956

2. Boothroyd, C. Poli, Applied Engineering Mechanics. New York Marcel Dekker Inc., 1980

3. W. B. Knowles, T. B. Sheridan, "The Feel" of Rotary Controls: Friction and Inertia," Human Factors, June 1966.

4. J. G. Kreifeldt, M. C. Chuang, "Moment of Inertia: Psychophysical Study of an Overlooked Sensation," SCIENCE 2 Nov. 1979.

5. J. G. Kreifeldt, M. C. Chuang, "Multidimensional Studies of a Tool "Feel" Sensation-Moment of Inertia." Proc. 23rd Annual Meeting of the Human Factors Society, 1979.

6. S. S. Stevens, Psychophysics, John Wiley & Sons, Inc. N.Y. 1975.

7. R. Pew, J. D. Chananie, "Subjective Scaling of Springs, Shock Absorbers, and Fly Wheels." Proc. 5th Annual NASA-University Conference on Manual Control, NASA SP-215, 1969.

8. M. Clarke, S. Handel, J. Kreifeldt, "Cross Modality Scaling: Mass and Mass Moment of Inertia." Proc. 18th Annual Conference on Manual Control, (in press).

9. S. Handel Personal Communication.

10. C. Zahner, M. Kaminaka, "Can a Tool User Sense "Moment of Inertia?" A Human Factors Engineering Study." Proc. 1982 Winter Meeting, American Society of Agricultural Engineers. Dec. 14=17, 1982, paper 82-1613.

11. H. VanCott, R. Kinkade, (eds.), Human Engineering Guide to Equipment Design. Superintendent of Documents, U.S.G.P.O., Washington, D.C. 1972.

BIBLIOGRAPHY

1. B. Woodruss, H. Helson, "Torque: A New Dimension in Tactile-kinesthetic Sensitivity, " American Journal Psychol., 78, 1965.

2. R. Drillis, D. Schneck, H. Gage, "The Theory of Striking Tools," Human Factors, 5, 1963.

Hands Coordination in Data Entry with a Two-Hand Chord Typewriter

Daniel Gopher, Walter Koenig
Cognitive Psychophysiology Laboratory
Department of Psychology
University of Illinois

ABSTRACT

This paper describes the results of an experiment conducted to investigate the process of acquisition and operation of a data entry skill based upon a newly designed two-hand chord keyboard. This keyboard represents an effort to identify effective alternatives to the existing typewriter. It is consisted of two separate 5-key panels (one for each hand), and characters on each panel are entered by typing chords composed of one to five fingers. Each panel is capable of producing the full dictionary of characters, and hence can be considered to be an independent typewriter. Three important questions raised by this design are (a) the best coding principle represent identical letters on the left and right panels, (b) the duration of the initial acquisition period, rate of progress, and final levels of performance on this system, (c) the nature of coordination between hands in simultaneous chord production.

The paper reviews the results of an experiment conducted to examine these questions. Three groups of subjects were trained for 10 hours in typing single or pairs of letters with the two hands. In one group, hands symmetry was employed to assign chords to letters. In another group, coding was based on spatial congruence. Finally, the two principles were combined for a third group, who practiced the task in a vertical posture. Experimental results show that subjects reached an average data entry rate of 130-150 characters per minute after 10 hours of training. Representation of letter codes by spatial patterns was much superior to hand symmetry. Unification of the two principles, in a vertical posture created yet another step improvement. The paper examines the significance of these results to the theory of motor control, and their implications to the design of data entry devices.

The last decade witnessed a renewed interest in a systematic investigation of typing as a model task representing the more general class of transcription skills such as piano playing, writing, speaking, or data entry with computer keyboards (e.g, Norman & Rumelhart, 1982; Gentner, 1982, Cooper, 1983; Schaffer, 1978; Logan, 1982). The reasons for this interest are both theoretical and practical. The study of typing behavior appears to provide an attractive opportunity to examine theories and concepts of skill acquisition. Moreover, typing is a highly complex perceptual motor skill requiring the coordinated performance of two very different basic mechanisms, namely, semantic processors and motor controllers. These theoretical questions can be investigated in the context of typing with considerable ease especially as highly experienced subjects (typists) are easy to obtain. In addition, advanced laboratory facilities, computer technology, and fast photography enable a fine graded examination of task behavior.

The practical value of this research is evident given the rapid proliferation of computer systems in all areas of our daily activities. Wandering through computer shops around the United States, one is still amazed by the number and age of children interacting with the various demonstration units at any single shop. These shops have clearly taken the place of the traditional toy shops as the major place of interest for these youngsters. Every one of these units includes a data entry keyboard in one form or another. Human-machine interfaces with engineering systems, at all levels, are not lagging far behind. Most of the newly designed systems include a data entry keyboard and some reference to a background computer.

Naturally, typing on the standard keyboard using commercial type units has received the greatest measure of attention. While the merit of these studies is evident, we see a need to augment and compare this research with the study of typing skills using radically different devices. We further argue that if such a research is not conducted, our understanding of the basic faculties of transcription skills may be biased by the specific characteristics of the standard keyboard design. These biases may also constrain our efforts to develop better solutions to the persistent problem of communicating with computers via a data entry keyboard.

This claim is based on the assumption that the structure of every skill is strongly influenced by the physical characteristics and the operational constraints of the environment in which it is developed. Typing skills reflect, therefore, the coping strategies of the human organism with the design characteristics of the standard typewriter. Most of the existing data entry keyboards follow the "Q-W-E-R-T-Y" arrangement suggested by Christopher L. Sholes in 1873. Thus, the beginning of 1983 marks the end of the first decade, of the second century, in which the Sholes' design dominates the typing field. The main features of this design from a human performance viewpoint are:

a) Every letter and character is entered by a separate key (although some keys may have more than one function), leading to a large size keyboard with several rows of keys.

b) Each finger is responsible for several keys (letters).

c) Each hand and finger is responsible for an exclusive set of characters.

d) Typing of most words requires considerable hand and fingers travel within the coordinates of the keyboard.

e) Single printing head forces serial sequencing of the final outputs, even when other elements allow parallel entries.

These design characteristics were determined by the technological constraints of the era in which the typewriter was invented--mid 19th century, the golden age of mechanical engineering. Despite the fact that over the years all the hardware components of the typing machine has changed, the keyboard interface remained unchanged. Compare, for example, the above design features with the single hand typewriter described by Gopher and Eilam (1979), or with the two-hand chord typewriter which is the system under study in the present research (see Figure 1).

Insert Figure 1 About Here

This keyboard consists of two separate 5-key panels (one for each hand). Letters on each panel are entered by typing a single chord composed of 1-5 keys pressed together. Each panel can produce the full alphabet, and hence can be cosidered to represent an independent typewriter.

The main features of this system are:

a) There are only five keys for each hand, and each finger always rests on its corresponding key.

b) No travel of hand and fingers is required, and no exertion of large muscles.

c) Letters are identified by a combination of keys, and entries are produced by typing chords. Hence, both memory and response requirements have been changed.

d) Each hand can produce the full dictionary of characters. Thus, in principle, operate independently.

e) The system enables parallel entry and display.

It is quite obvious that these design characteristics constitute a radical change in the task requirements, while still residing within the same general problem area of transcription skills. Two main questions arise in introducing such a system:

a) What are the similarities and differences between the dimensions of skill developed on the chord keyboard and those acquired on the standard typewriter? In other words, what are the elements of the skill that reflect the inherent properties of the human processing system and, therefore, would remain uncanged on the two systems. And what are those that are determined by the specific constraints

of the operational environment and should, therefore, be changed considerably.

b) Can we facilitate training and improve overall typing performance with the new degrees of freedom given to us by modern technology?

However, before these two issues can be studied, an initial question has to be answered. What is the best principle to represent the code for the same letter on the two panels? For example, if the letter "A" is represented on the right hand panel by a chord composed of the thumb and index fingers (first and second keys counted from left to right); what will be the appropriate code for the letter "A" on the left hand panel? Two possible principles come immediately to mind:

a) Hand Symmetry--according to this principle, the same letter is represented on the right and left panels, by the symmetrical fingers of the hands (for example, the letter A is represented by the thumb and index fingers of each hand).

b) Spatial Congruence--according to this principle letters are represented by their pattern of key arrangement of the two panels (for example, "A" will now be entered by pressing the first and second keys on the right and left hand panels, with keys counted from left to right). Thus, while a Hand Symmetry code is based upon a body reference point and proprioceptive information. Spatial Congruence is based upon an external arrangement and spatial information (the two coding principles are demonstrated in Figure 2, together with their combination in the vertical position, which is discussed below). It is important to note that these coding principles constitute two fundamentally different representation formats of action plans in long term memory. Formats that are the heart of an ongoing debate in current motor theory (e.g., Kelso & Wallace, 1978; Gopher, in press). Note also that the two principles create mirror images of each other on the two panels, either in terms of spatial patterns or operating fingers.

Insert Figure 2 About Here

The study of the two principles is of a high theoretical value may be of lesser practical merit, because there is one spatial position in which the two are united.

c) Combined in Vertical Position--if the two chord panels are tilted vertically (or upright from their flat horizontal position) the conflict between the two representation principls is resolved, and the two unite (see Figure 2). The question is, of course, whether such a unification provides any advantage over the better of the two other arrangements. Note that placing the keyboard panels at an upright position not only resolves the representation conflict, but also constitute a better ergonomic posture for the hands (Kromer, 1972).

The experiment reported in this paper is devoted to the study of the effects of using the three principles, on the coordination between hands, and the acquisition of a data entry skill.

METHOD

Task:

In the initial phase subjects were presented with charts describing the letter codes relevant to their respective conditions (i.e., hand symmetry, spatial or combined, see Figure 2). They were required to memorize the 26 letter codes of each hand. During this phase, they could press various letter patterns on the keyboard and the respective letter would appear on the display.

Following this stage and throughout the seven experimental meetings subjects were required to respond as quickly as possible to letters presented on the display with the appropriate codes. The letters were presented singly, on the left or right side of the display, or dually. Subjects were required to respond to stimuli presented on the left side of the screen with the left hand, the right side of the screen with the right hand and to dual presentation with both hands.

Trials were organized into mixed or dual blocks. In mixed blocks the subject was presented with an equal number of single (left or right) and dual letter trials. In these blocks the full 26 letter alphabet was used. In dual blocks the subject was presented with dual stimuli (left and right) every trial. However, only a limited eight letter set was used. The eight letters selected for use suggest some interesting compatibility and conflict issues, that would be elaborated on in the result section. The trial format was the same for both mixed and dual letter blocks. There was a warning signal 800 msec before the appearance of a letter. It informed the subject on the nature of the coming trial (single or dual), and if single, whether the letter would be presented on the right or left side. There was then a 300 msec presentation of the letter stimulus, followed by another 1700 msec response interval, and feedback was given to the subject on his performance.

Apparatus

The experiment was governed by a PDP 11/40 computer system. Letters were presented on a plasma panel display. Responses were entered through the two-hand chord keyboard (Figure 1) and recorded directly onto magnetic tape. The keyboards were positioned horizontally for the hand symmetry and vertically for the combined condition. Hand rests were added in the combined condition. The keyboards were situated between the subject and the plasma board display.

The letters presented were 1/2 cm tall and 1/4 cm wide. In the dual presentations the letters were 4.2 cm apart. The line of sight distance was approximately 60 cm. The visual angle for the single stimulus was approximately .45 of a degree The visual angle between the stimulus was approximately 3.75 degrees. In addition to behavioral measures, electrophysiological measures were taken during meetings 3-7. The electrode sites were Fz, Cz, Pz, C3, and C4. EOG measures were taken to correct for horizontal and vertical eye movements.

Procedure

The subjects participated in 7, 1.5 hour training sessions. The first session was dedicated to code acquisition and initial familiarization. Meetings 2-7 followed the same format: 8 mixed blocks of 52 trials followed by 2 dual blocks of 64 trials. In a given meeting there was a total of 416 mixed trials and 128 dual trials resulting in a total of 544 trials per meeting. In addition to the mixed and dual blocks, meetings 2, 4, and 7 also included a control task. This task involved the presentation of the same letter for 32 trials. There were two blocks; one to the left hand and one to which the right hand responded to singly. This control task was used to estimate simple reaction time without the cognitive processing of letter codes.

Subject performance was motivated with a bonus system. In this system both speed and accuracy were emphasized. Subjects received $1.50 bonus for every 10% improvement in their response time. However, a bonus was awarded only if the percentage of errors per block of trials did not exceed 5%.

Subjects

Subjects were 12 right-handed male college students with English as their native language. They were payed for their participation at a rate of $3.50 pe hour, and received additional bonuses based on their performance. Subjects were assigned randomly to the three experimental groups.

RESULTS

In the present paper we limit our description of results to the main performance data obtained for the three experimental groups.

Initial Memorization

The initial phase under all conditions was memorization of the 52 letter chords on the two hands. Subjects were given a complete freedom to use their own methods to study and memorize the alphabet.

The general finding of this stage was that subjects in all groups could memorize the codes within 35-40 minutes, to a criteria that they could produce any letter upon request. There are some indications that subjects in the vertical-combind group have mastered the codes even faster. At the end of 40 minutes subjects could, therefore, be presented with letters on the display, to which they reacted in a mode similar to touch typing on a regular keyboard. It should be recognized that because each of the fingers always rests on its home key, once the codes have been comitted to memory, no further visual supervision is required.

Acquisition

Figures 3 and 4 depict the learning curves of single and dual letter entries for one subject who practiced under the vertical-combined condition. These curves are typical to the learning functions of the majority of subjects, and show a continuous improvement throughout the meetings, with no asymptote yet encountered. From the results presented for this subject in Figure 3, and from the average results presented in Figure 5 it can be observed that at the 7th meeting, single letter entries reach the level of 550 to 660 msec with a small but reliable advantage to letters entered by the left hand. Such an advantage was observed in the performance of

---

Insert Figures 3 & 4 About Here

---

subjects in all groups. Its average size was 65 msec ($t(8) = 5.76$ $p < .001$).

When a pair of letters was presented, the left hand letter was usually entered first. This outcome can be clearly observed in Figure 4 and is typical to the performance of subjects in all experimental groups. The two inbound curves in this figure represent the average entry time for the left and right hand letters under dual letter presentations. Entry times improved regularly on both hands, and the time to complete a pair reached, for this subject, an average level of 800 msec at the end of the 10 hour training period. A level which compares with a typing rate of 150 characters per minute.

The lower curve in Figure 4 is a replot of the entry times of single letters with the left hand. It demonstrates the existence of a small but consistent delay in the beginning of typing under dual letter presentations. Similar delays were evident for all subjects, but were much more pronounced in the hand symmetry group (see Figure 5). The upper curve in Figure 4 represents the additive values of single letter entries (left and right) at each time point (block of trials). It was computed to examine the development of parallel entry capabilities for dual letter presentations. The difference between this curve and the entry times of the right hand letter (the second in a pair) can then be expressed as a percent of saving due to response overlap in dual letter entries. For the subject whose data is presented in Figure 4, this saving was 26%, at the end of the 7th meeting.

Differences Between the Three Representation Principles

Figures 5 and 6 summarize the main differences in performance between the three experimental groups, at the last experimental meeting. These plots are based on correct responses only. Because subjects received a bonus for improved speed only when errors were less than 5%, the general error percentage in all groups was low (about 3%). Error distributions did not show any reversed trends due to speed accuracy tradeoff, and are therefore omitted from the present discussion.

Figure 5 depicts the results obtained in mixed experimental blocks, in which trials included an equal number of single and dual letter presentations. Four data points are plotted for each group: 1) average time for tapping, 2) single letter entries (left and right entries were averaged for both of these measures), 3) the time to type the first letter in a simultaneous pair, and 4) the time to complete a pair. As can be seen,

---

Insert Figure 5 About Here

---

the three groups did not differ in the speed at which subjects could perform a simple reaction time task (tapping) on the keyboard. In single letter conditions the hand symmetry group appears to be slower than the other two groups, but this difference did not reach statistical significance ($F(2/9) = 1.12$, $p < .37$, $F(2/9) = 2.254$ $p < .16$ for single left and right letters respectively).

Large and reliable differences between groups appeared in dual letter entries. In both the time to begin and the time to complete the entry of a pair, the hand symmetry group was considerably slower than the other two groups ($F(2/9/ = 4.93$, $p < .036$, $F(2/9) = 6.63$, $p < .017$ for beginning and completing a pair respectively). In addition, the vertical-combined group was faster than the spatial group in the time to complete a pair.

The same pattern of results appears in Figure 6 which depicts the results from blocks composed only of dual letters presentations. Aside from including only pairs of letters, these blocks differed from mixed blocks in two additional important features. First, they incorporated only a limited set of 8 letters, while the complete set of 26 letters, in equal probabilities, was presented in mixed blocks. Secondly, these 8 letters included only chord combinations of one to three fingers that were relatively easier, as determined from their response times in single letter trials. The objective of selecting these letters was to provide enough data points to contrast specific combinations of interest. For example, comparison between pairs of the same letter that have the same chord pattern under hand symmetry or spatial congruence conditions, versus those that create a mirror imagery of each other on the two hands and, therefore, represent a conflict situation (see Figure 7). Simple chords were selected to separate the study of a conflict in representation principle from the issue of coordinating difficult motor chords. The results plotted in Figure

---

Insert Figure 6 About Here

---

6 clearly indicate that even with a limited and simplified set, the hand symmetry group was considerably slower than the other two groups, and that the subjects in the vertical-combined condition were still faster to

---

Insert Figure 7 About Here

---

complete a pair entry ($F(2/9) = 3.88$, $p < .06$; $F(2/9) = 8.83$, $p < .008$ for the first and second letter in a pair respectively).

Figure 8 presents the differences between the three groups at the seventh meeting, in their ability to overlap responses in dual letter trials. To recapitulate, this ability was calculated by comparing the time to complete a pair entry, with the time to enter a left and a right letter in succession. In essence, this measure provides an estimate of the developed time-sharing capabilities in the three groups. Here again, the hand symmetry group lagged behind. As the figure shows the other two groups were not only better in relative terms, but also in their absolute levels of time saving, that were considerably larger than those of the hand symmetry group, despite their much shorter response times. The vertical-combined group was again better than the spatial congruence group. The F ratio for

---

Insert Figure 8 About Here

---

the differences in relative saving was 5.06, $p < .034$. The absolute differences failed to reach statistical significance ($F(2/9) = 2.37$, $p < .15$).

The last comparison to be described in the present paper, is between typing of pairs of the same letters, when the codes of these letters have the same spatial and hand symmetry arrangement on the two hands, versus those letters that create a conflict with regard to one of the two representation principles (see Figure 7). To state differently, the issue is, what is more confusing to type the same letters when different fingers have to be operated (spatial congruence), or to enter the same letters when they create a different spatial pattern (hand symmetry). Note that unlike previous comparisons, this analysis calls for a comparison within the performance data of each group. Thus, in the vertical-combined condition, when the same letters are presented to the two hands there always exist both a spatial and a hand symmetry matching. Under hand symmetry, conflicts occur because same letters create different spatial patterns on the keyboard. Under spatial congruence, they require the use of different fingers. Four letters were included in each of these groups (conflicting vs. nonconflicting). Table 1 summarizes the main results of this comparison. It is based on the combined data of subjects performance, in

---

Insert Table 1 About Here

---

meetings 5, 6, 7.

From inspecting this table it is clear that the performance of subjects in the hand symmetry group did not suffer much from the spatial conflicts. Their response times were about the same in the conflicting and nonconflicting groups of letters ($t(7) = .753$, $p > .475$). The same was true for subjects in the combined group, whose performance for the second group of letters was even slightly better ($t(7) = 3.343$, $p < .013$). In contrast, the performance of subjects in the spatial congruence group showed pronounced deterioration due to the requirement to type the same letters with asymmetric fingers ($t(7) = 4.703$, $p < .005$). Note however, that this deterioration should be only interpreted in relative terms, because the absolute levels of performance for the conflicting group were about equal for the hand symmetry and the spatial congruence groups.

DISCUSSION

The significance of the present results should be examined both from an applied and a theoretical viewpoint. Taken together, they showed that: a) subjects were able to perform in a touch typing mode, and memorize the codes for all letters, after a brief period of self teaching, b) that progress in learning was fast and included the development of parallel entry capabilities, c) that in general, representation of codes by spatial patterns was considerably better than coding by hand symmetry, d) that performance with an upright tilted panel was better than with a horizontal panel.

These findings raise the possibility that for many system applications, a chord keyboard of the type described in the present study, may constitute a viable and attractive alternative to the traditional typewriter or data entry keyboard. Touch typing ability and similar entry speeds on a standard keyboard, are the achievement of several months of daily practice (e.g., Hill, Rejall, & Thorndike, 1913; Dvorak, Merrick, Dealey, & Ford, 1936; Gentner, 1982). In light of the fundamental differences between the two typing keyboards in their skill components, one can conclude, that it appears to be easier for humans to commit 52 chords to memory and activate them upon request, than to learn the ways of the hand to a similar number of keys spread out on a typing keyboard. This conclusion is also supported by another line of experiments, with a single hand chord typewriter for the Hebrew language (Gopher & Eilam, 1979; Gopher, in press).

At this stage no claims can be made with regard to top or asymptotic typing speeds. It may very well be that a regular typing skill is harder to acquire, but it yields better performance when proficient. Our present argument is limited to the novice trainee, and the nonprofessional user. However, the results are provocative enough to encourage a comparative study with prolonged training. An important characteristic of the present keyboard from an application perspective, is that its major skill components are so different from those required in standard typing. Some of our subjects were proficient typists, but no interference or facilitation of performance could be associated with this fact. In different words, one can acquire and maintain the two skills side by side with little interference.

One other practical implication of the data that should not be ignored, is the improved performance with an upright keyboard. As has been argued in the introduction section, such an angulr tilt provides both a representation and an ergonomic advantage. The full impact of the latter factor was most likely not revealed in the present study, in which a special effort was made to minimize the effects of fatigue. The total duration of each experimental session was relatively short, trials were discrete, and subjects were given many intermissions.

The degrees of freedom given to us by modern technology enable us to consider such a change in the keyboard orientation with little impact on the total costs of system design and production. Should we wait for the process of natural evolution to develop a human mutant, whose hands are optimized for performance on flat panels? Or shall we take the initiative and change our panels to fit the ergonomics of the human hand? It is true that when visual monitoring is necessary, upright panels are deficient, but with chord

keyboards vision requirements are minimal. Moreover, even on the regular typewriter vision plays a heavy role only until proficiency is acquired. Then, motor factors constitute the major constraints on performance (Norman & Rummelhart, 1982). Can the panels be tilted at this stage to improve performance?

From a theoretical viewpoint, the most important finding was the robust advantage of a memory representation by spatial patterns over a representation by operating fingers, which was most pronounced when the two hands had to be coordinated in simultaneous activation. This advantage lends support to current theoretical views that argue that, at their upper level, complex action plans are represented by the spatial pattern that the corresponding movement creates in the outside world (e.g., Bernstein, 1967; Kelso & Wallace, 1978; Gopher, in press).

Additional support to the importance of the spatial patterns in the memorization and retrieval of motor chords, comes from data obtained by Nachum Fossfeld at the Technion in Israel, using the same two hand chord keyboard for the Hebrew language. Following the training of his subjects in letter typing (in a combined condition only). He required them to vocalize letter names in response to visually displayed chord patterns shown within a schematic drawing of the keyboard on a computer display. The correlation between the verbal response times to identify letters and the response times for actual typing of these letters was 0.73, indicating about 50% of common variance. Furthermore, these verbal response times correlated at a 0.93 level with the same responses made by a control group that did not receive any former training in typing on the chord keyboard. It, therefore, appears that the difficulty of recognizing and retrieving patterns is a main factor in the response variability of motor productions, even in the most advantageous condition (i.e., vertical posture).

There was only one exception in the results to the clear superiority of a representation by spatial patterns. When the same letter had to be typed by both hands, and its code required the operation of asymmetrical fingers, performance in the spatial congruence group degraded. A similar decrement was not observed in the contrasting case, when subjects in the hand symmetry group were presented with same letter pairs that create mirror imagery patterns on the two panels. A possible interpretation for this finding is that subjects find it easier to operationalize the meaning of the property "same" as associated with same fingers, than to relate it to the same spatial patterns. Such an interpretation also corresponds to the informal intuition of the vast majority of the visitors in our laboratory, which favors hand symmetry as a coding principle.

How can we reconcile the obvious inconsistency between this interpretation, and the clear superiority of the spatial representation principle in the majority of cases. It should be recognized that from the total number of 676 possible pair combinations of letters, pairs of the same letters are less than 4%, and conflicting pairs are only 3% of the total number of combinations. Hence, in the majority of trials, subjects were presented with letter pairs that differ both in their fingers and key arrangement. It is in those cases that subjects were better able to perform, based upon a spatial representation principle.

But why would it be more difficult to represent different chord patterns by their corresponding fingers? At this stage we do not have a definite answer to this question. However, we would like to propose a possible direction for such an answer. Suppose that the brain commands the operation of fingers, such that when fingers of one hand are activated, there is also an increase in the activation of the symmetrical fingers of the other hand. We are not aware of neurophysiological work that studied this question, but there is some behavioral research that supports the existence of symmetrical activation (e.g., Rabbit, Vyas, & Fearnly, 1975). It follows that when subjects are instructed to enter two different letters on the two panels, they have to activate different fingers on each hand, and at the same time inhibit the corresponding activation of the symmetrical but irrelevant fingers of the two hands.

For subjects in the spatial congruence condition, this latter requirement is consistent with their general rule of training. They are taught to ignore all hand symmetry considerations and concentrate solely on the spatial patterns created by the chords. In contrast, subjects in the hand symmetry group are subjected to an eternal, peculiar, double bind conflict. In entering different letters they are required, on the one hand, to inhibit the natural symmetrical activations of fingers on the two hands. At the same time they are called to memorize hand symmetry because this is the general representation rule that they draw upon. The only time that this conflict is resolved, is in the event that both hands have to enter the same letter. It is this combination that create the largest problem for subjects in the spatial congruence group. At this stage this interpretation is a post hoc speculation, and should be validated by direct manipulations of the relevant variables.

The last outcome to be discussed is the improved performance that was obtained when the keyboard panels were tilted to an upright position. The exact causes for this improvement cannot be determined from the results of the present experiment. The main question is, whether it should be attributed only a representation factor (because all conflicts between spatial and finger based representations are resolved). Or, whether there is an actual facilitation in the motor coordination of the two hands in the upright posture.

Note that the only difference between the best and the worst group of the present experiment was the angular position of their panels. If it is indeed a representation problem, can subjects in the hand symmetry condition be taught to visualize their panels in an upright position, and then the differences in performance between groups will disappear? If the keyboard panels of subjects in the vertical combined-condition are flattened at the end of their training, would there be an immediate, transient, or enduring change in performance levels? These are two of the questions that have to be answered in future research.

In conclusion, and despite the many uncertainties in the interpretation of the present results, they have proved both the applied and the scientific merit of continuing this venture. A different approach to the development of data entry devices is both feasible and useful. It may also provide convergent evidence on the ways in which our words command our hands.

Acknowledgement

The work reported in this paper was supported by Grant #N00014-83-K0092 from the Naval Office of Research program in personnel and training research. Dr. Henry Halff was the scientific monitor for this grant. The system described in this paper was originally developed for the Hebrew language by the first author together with Nachum Fossfeld at the Technion, Israel. We gratefully acknowledge the programming work, technical help and scientific suggestions of Gabriele Gratton and Art Kramer. We are also indebted to Emanuel Donchin and Michael G. H. Coles for their continuous involvement in this work.

References

Bernstein, N. A. The coordination and regulation of movements. Oxford: Pergammon Press, 1967.

Cooper, W. E. (Ed.) Cognitive aspects of skilled typewriting. New York: Springer Verlag, 1983.

Dvorak, A., Merrick, N. L., Dealy, W. L., & Ford, G. C. Typewriting Behavior. American Book Co. 1936.

Gentner, D. R. Evidence against central model of timing in typing. Journal of Experimental Psychology: Human Perception and Performance, 1982, 8, 793-810.

Gentner, D. R. The development of typewriting skill. CHIP Report 114, University of California, San Diego. La Jolla, CA 92093, September 1982.

Gopher, D. On the contribution of vision based imagery to the acquisition and operation of a transcription skill. In: Prints, W., Sanders, A. (Eds.) Cognition and Motor Processes, Springer-Verlag, in press.

Gopher, D., & Eilam, Z. Development of the letter-shape keyboard: A new approach to the design of data entry devices. Proceedings of the 23rd annual meeting of the human factors, 1979.

Hill, L. B., Rejall, A. E., & Thorndike, E. L. Practice in the case of typewriting. Peolagogical Seminary, 1913, 20, 516-529.

Kelso, S. J. A., & Wallace, S. A. Conscious mechanisms in movement. In Stelmach, G. E. (Ed.) Information processing in motor control and learning. Academic Press, 1978.

Kroemer, E. Human engineering: The keyboard. Human Factors, 1972, 14, 51-63.

Logan, G. D. On the ability to inhibit complex movement: A stop signal study of typewriting. Journal of Experimental Psychology: Human Perception and Performance, 1982, 8, 778-792.

Norman, D. A., & Rumelhart, D. E. Studies of typing from the LNR Research Group. CHIP Report III, University of California, La Jolla, California, 920 93, September 1982.

Rabbitt, P. M. A., Vyas, S. M., & Fearnley, S. Programming sequences of complex responses. In: Rabbitt, P. M. A. and Dornik, S. (Eds.) Attention and Performance V. Academic Press, 1975.

Shaffer, L. H. Timing in the motor programming of typing. Quarterly Journal of Experimental Psychology, 1978, 30, 333-345.

Table 1

Average response time (msec) for entering conflicting and nonconflicting pairs of the same letter (meetings 5-7, dual letter blocks).

		Hand Symmetry		Spatial Congruence		Vertical Combined	
		First	Second	First	Second	First	Second
Letters that do not conflict	$\bar{x}$	618.0	644.0	559.0	591.0	585.4	591.6
(E, H, N, S)	SD	77.2	62.3	24.8	35.3	30.6	37.9
Letters that create a conflict	$\bar{x}$	626.0	652.0	646.6	669.7	554.3	551.8
(A, U, I, T)	SD	59.9	38.3	79.2	74.8	40.8	35.85
Difference conflict-no conflict		8.00	8.00	87.60	78.62	-31.12	-39.81

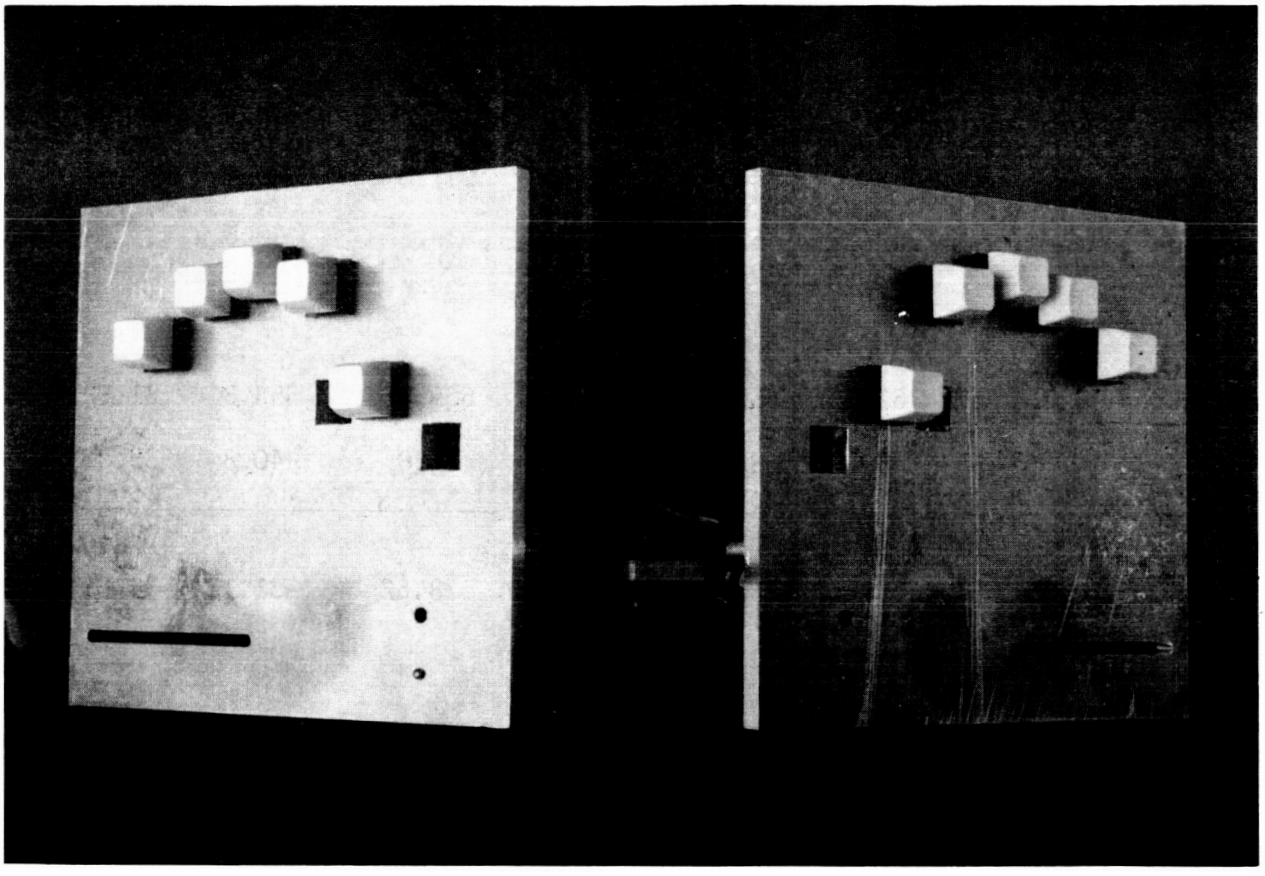


Figure 1 - Top down view of the two-hand chord keyboard. Note that each of the plates can be rotated to any angular position from horizontal to vertical posture.

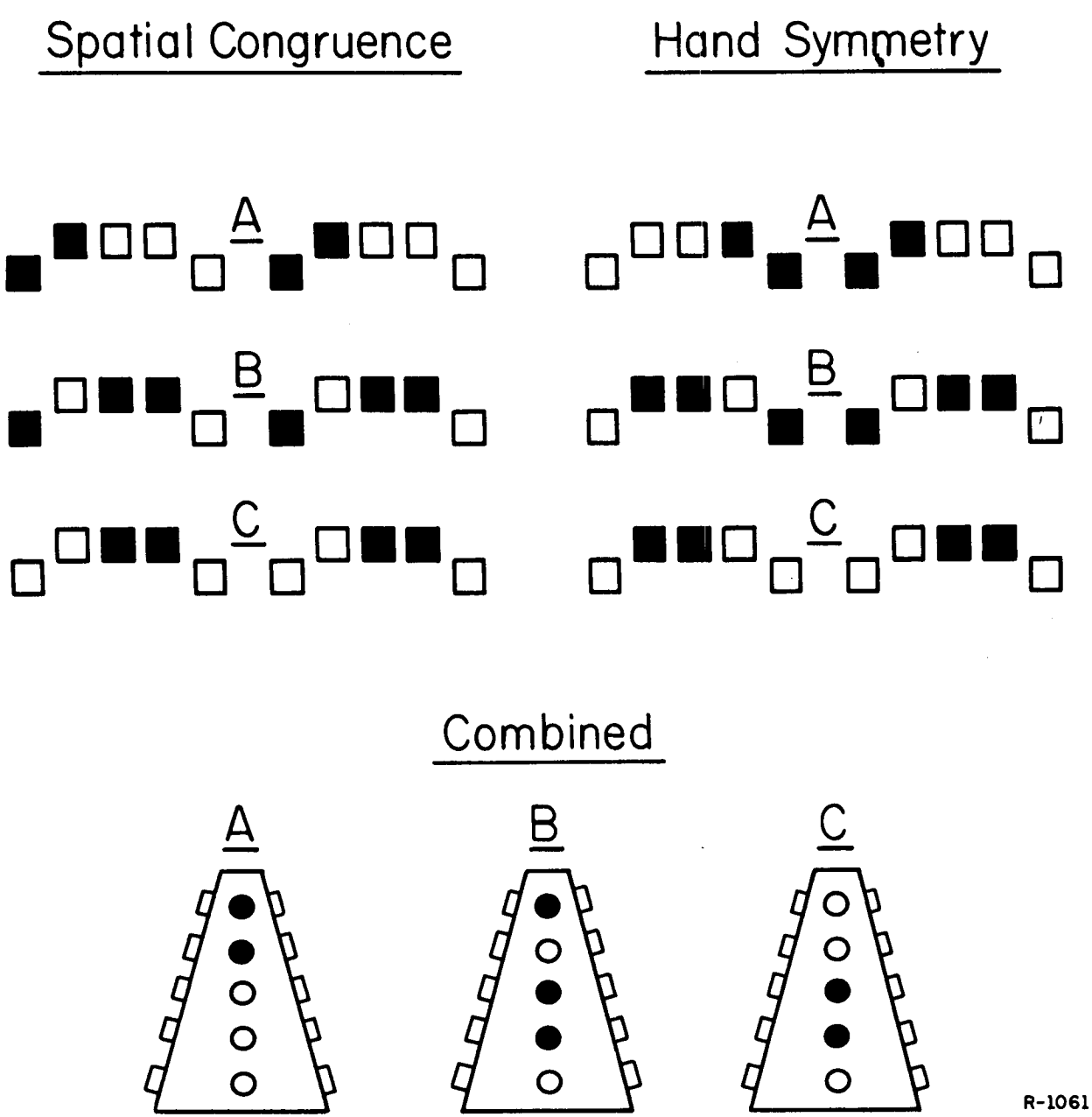


Figure 2 - Coding principles for associating letters with chord entries

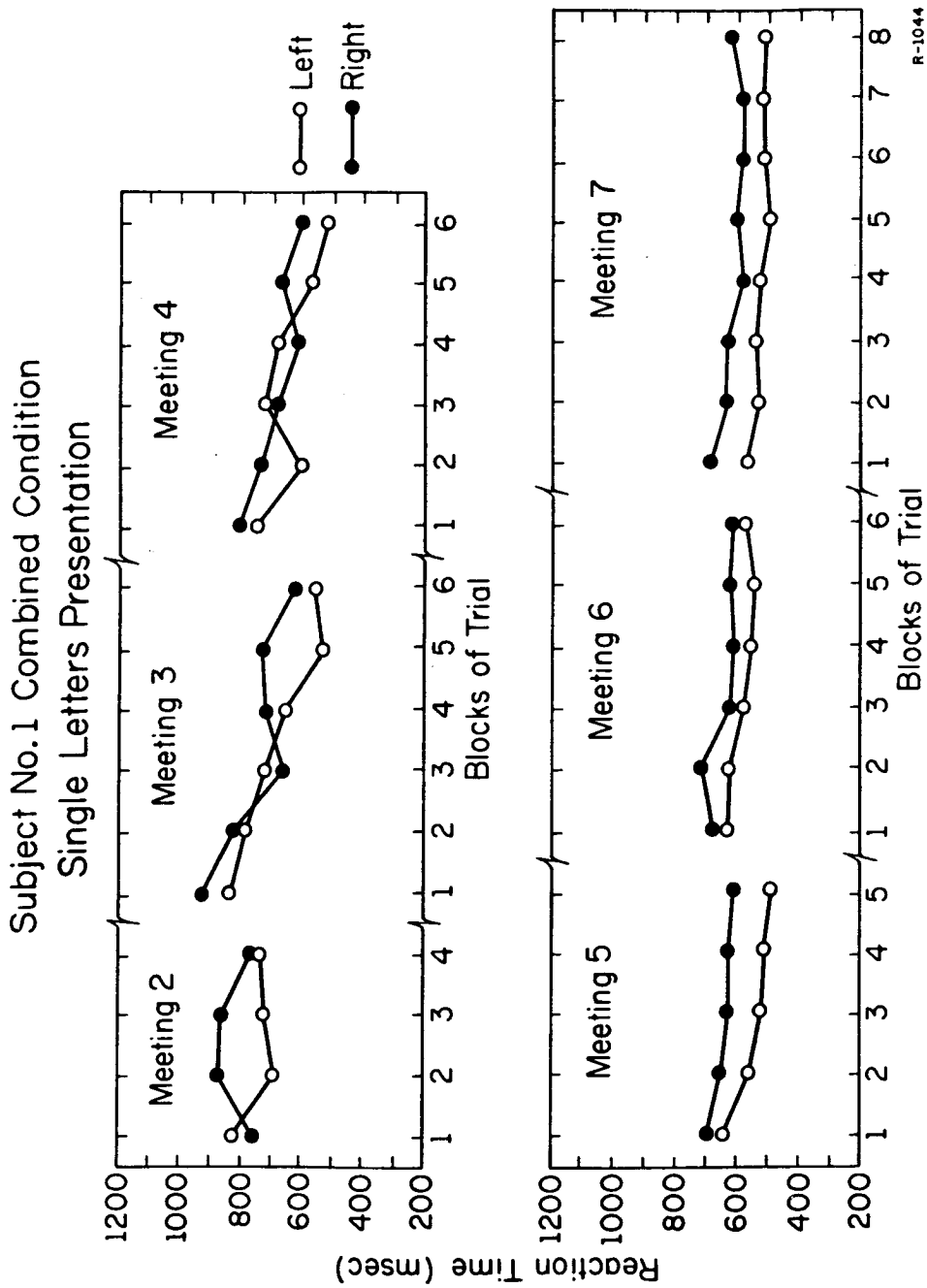


Figure 3 - The progress of training of one subject in single letter entries (averages for blocks of 52 trials)

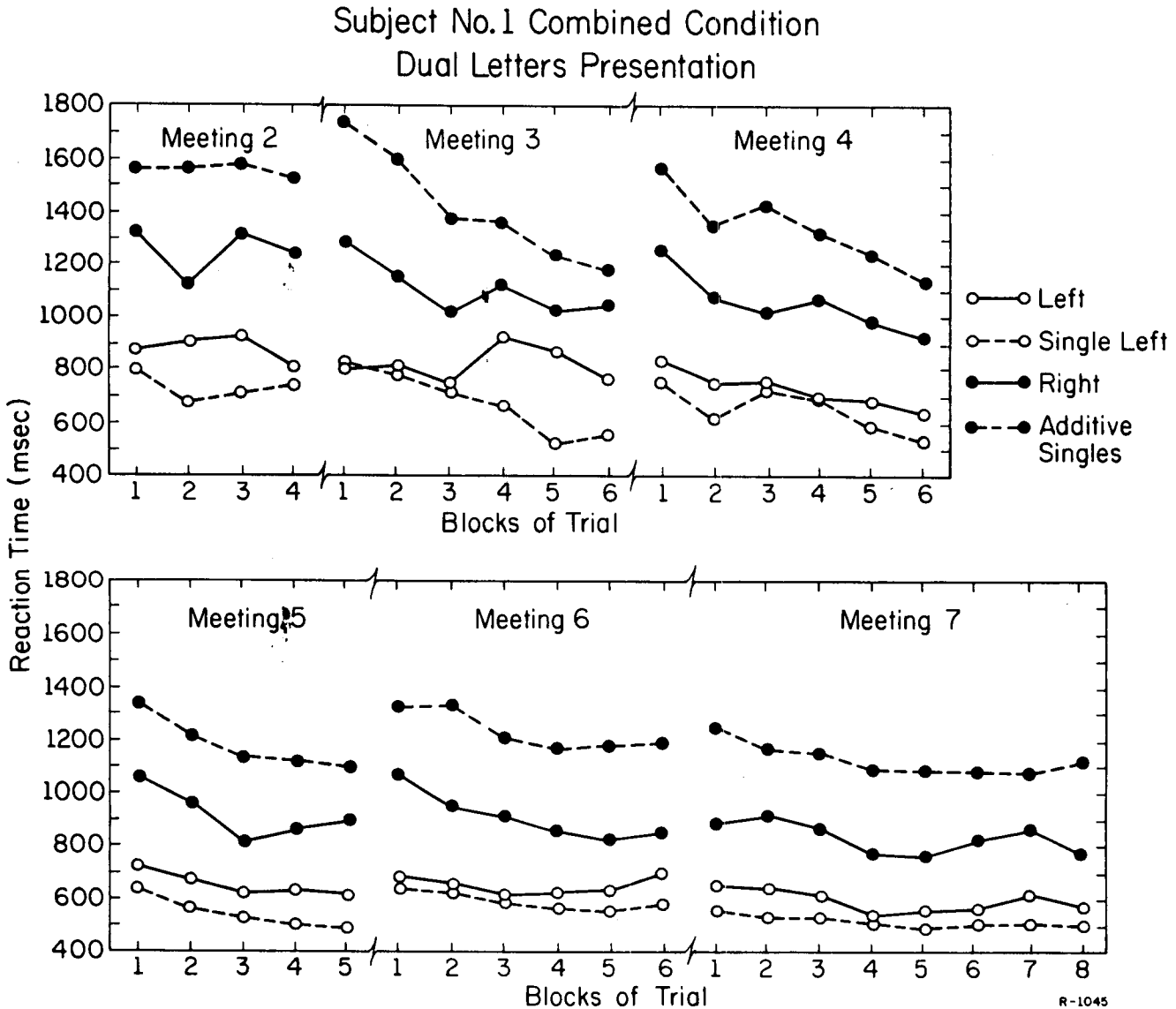


Figure 4 - The Progress of training for one subject in dual letter entries (mixed blocks, 52 trials in a block). The two inner curves represent the left and the right hand entries of a dual pair respectively. The lower curve is the time to type a single letter in single letter conditions. The upper curve is the time to type two letters in a sequence.

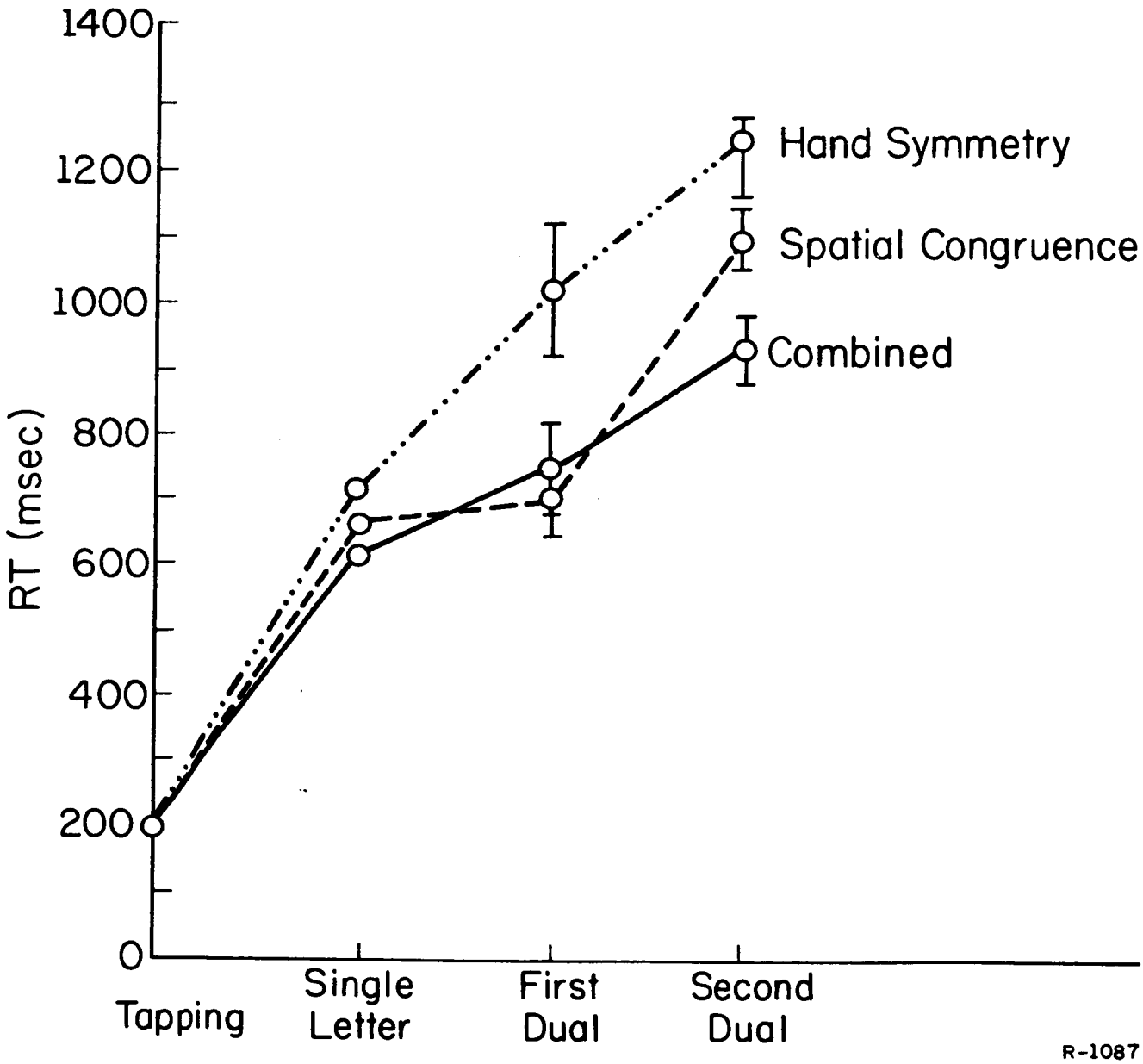


Figure 5 - Average response time to single and pairs of letters for the three experimental groups at the 7th training session

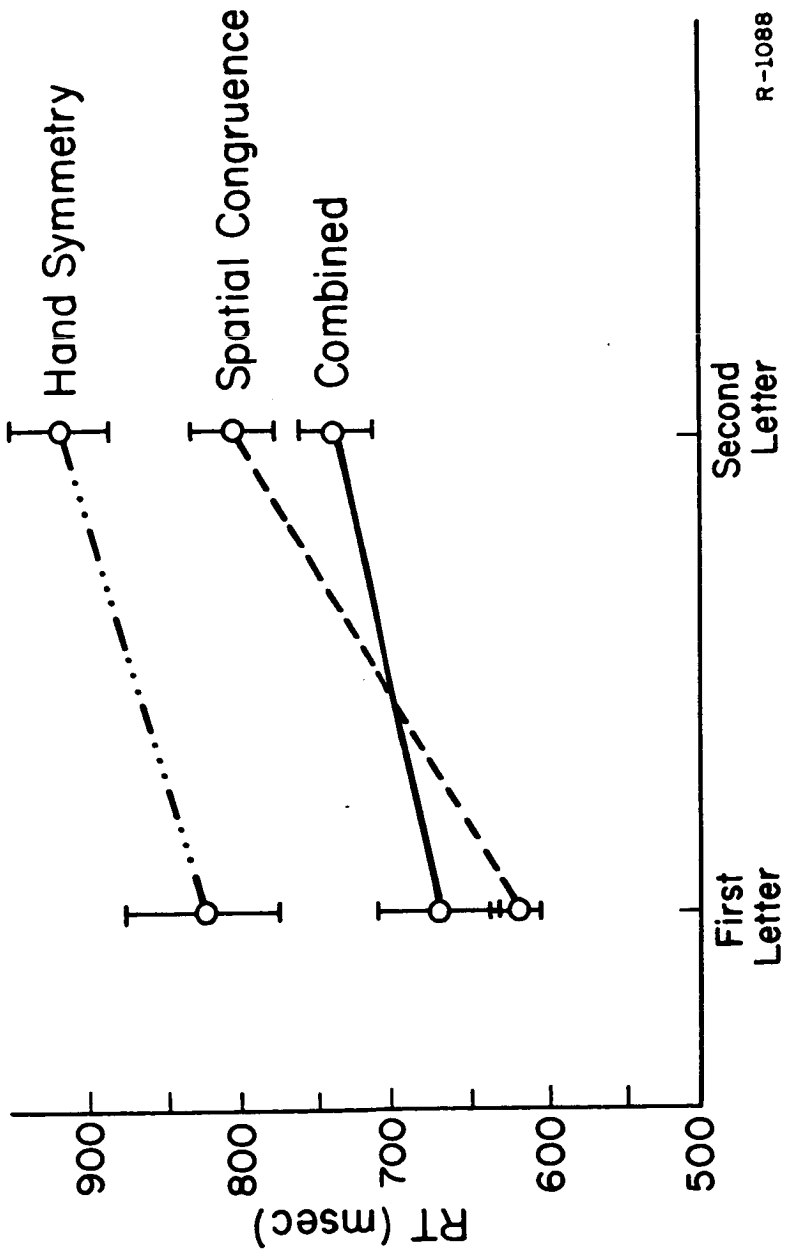
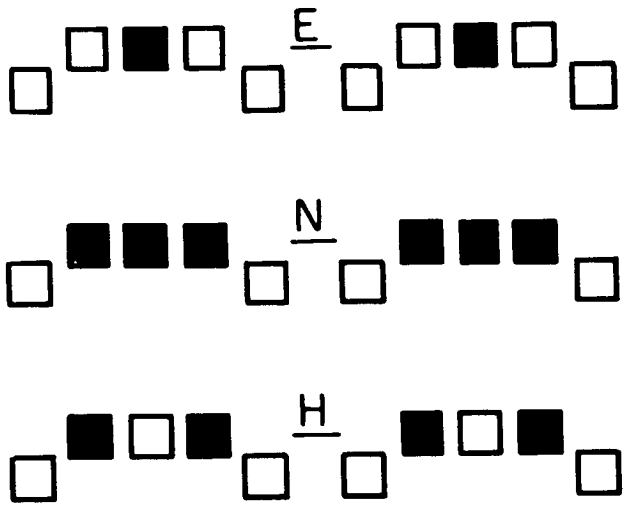
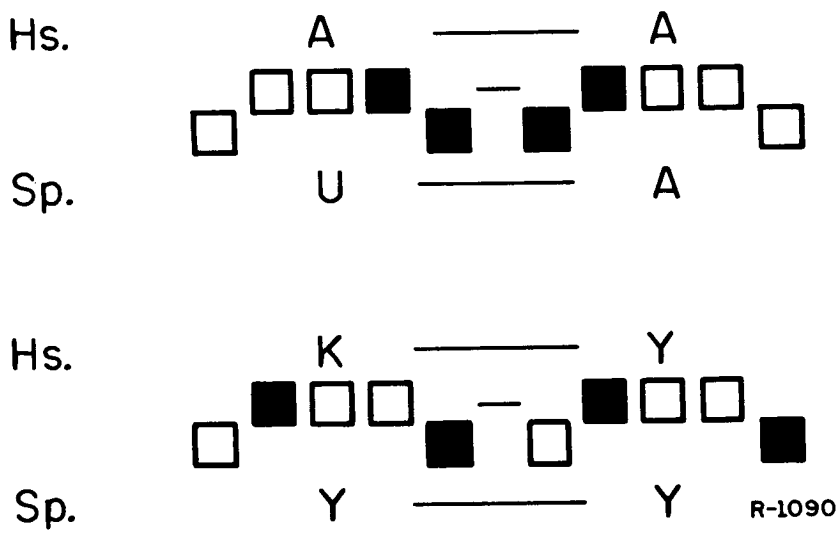


Figure 6 - Average response times in dual letter entries for the three experimental groups (7th meeting, dual-limited set blocks)

Letters with Equal Representation Under Hand Symmetry or Spatial Congruence



Letters that Create Conflicts



Hs. = Hand Symmetry

Sp. = Spatial Congruence

Figure 7 - Examples of letter codes that do not create conflict under any representation principle, and codes that create either spatial or hand symmetry conflict

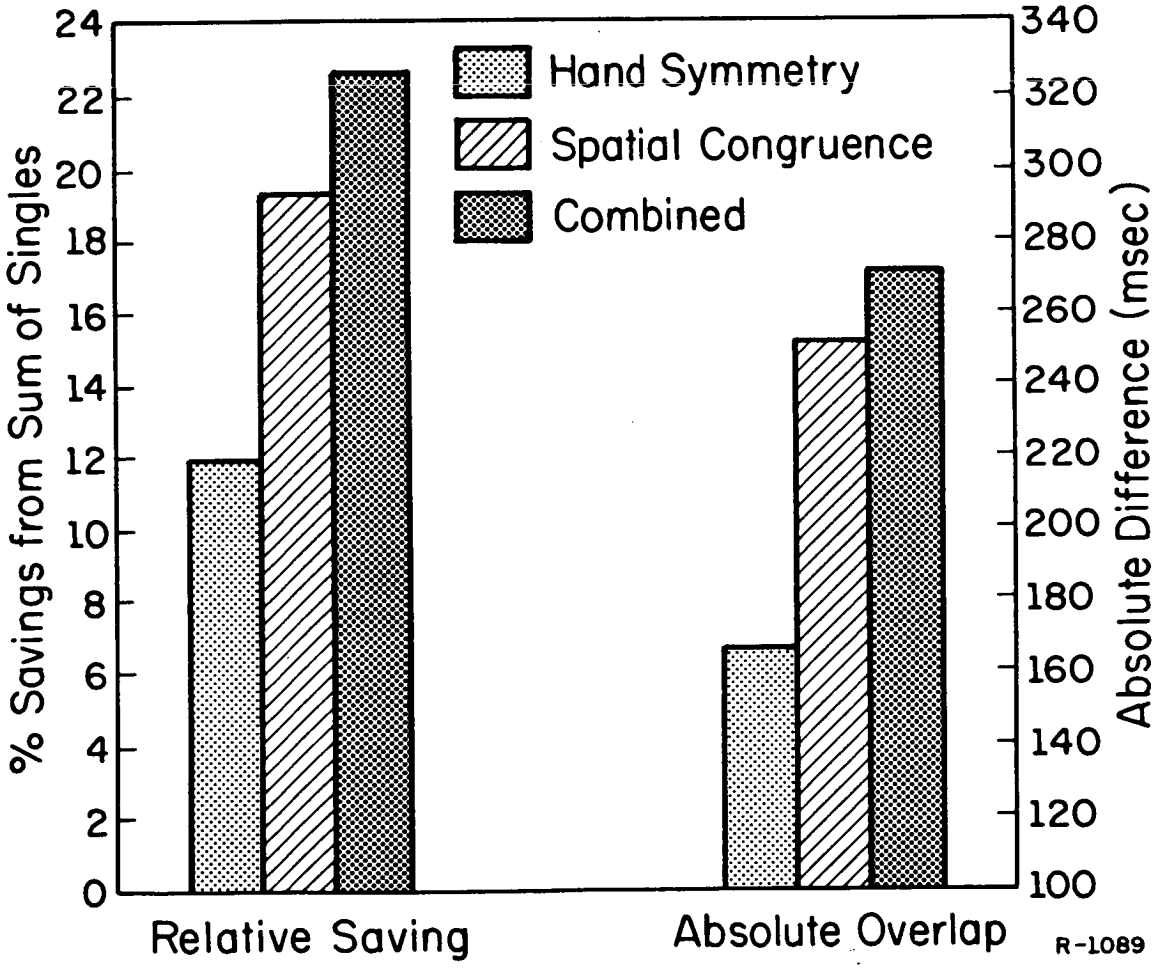


Figure 8 - Relative saving and absolute difference in typing a simultaneous pair of letters as compared with a sequential entry of two letters
(7th meeting)

MEMORY AND ACTION GRAMMAR FACTORS THAT INFLUENCE THE HUMAN-COMPUTER INTERACTION *

Richard A. Chechile
Rebecca N. Fleischman
Darleen M. Sadoski

Psychology Department
Tufts University
Medford, Mass. 02155

Abstract

In designing the software interface between the pilot and the avionic computer, it is important to understand the impact of design characteristics on the ability of the pilot to execute all of the time-shared tasks. Towards that goal, an experimental paradigm has been developed to study the human-software interface in a shared-attention task. The primary task consisted of editing the route-way point information and the concurrent secondary task involved the detection of digits greater than six on a random digit generator. The influence of two factors was studied on the shared-attention task. The first factor is the number of operations which needed to be recalled by the pilot from long-term memory. The second factor was the syntactic complexity of the human-software interface as measured by the length of an action grammar string representation. The number of memory retrievals did predict very well the accuracy and time to complete the time-shared tasks. While the string length did influence the time to learn the edit task, the results of three experiments showed that once the edit task was learned then the string length factor could not predict the shared-attention performance. These findings are discussed in terms of design guidelines for the human-software interaction.

* Funds for the support of this research have been provided in the form of a grant to the first author from the NASA-AMES Research Center, Moffett Field, California, grant no. NAG 3-51.

INTRODUCTION

There are many workstations, such as the airplane cockpit, where the operator must perform time-shared duties. In designing equipment for such environments it is essential to understand the relationship between the design features and the ability of the operator to maintain good performance on all of the time-shared duties. This research is directed at understanding the reasons that underlie the length effect. The length effect is the phenomenon of degraded performance in time-sharing when more operations are required of the operator, viz. Chechile, Butler, Gutowski, and Palmer, 1979. This prior work clearly showed that simple motor activity or eye movements were not the cause of the attentional difficulties since in their study the subjects could function with virtually no errors on either the primary or secondary tasks as long as the subjects were not required to recall subsequent primary task steps from long-term memory. These authors have argued that the act of retrieving information from memory requires cognitive effort such that a single task can be done without error. However, when a side task is introduced, the act of retrieving information from memory may require more cognitive effort than is possible in the available time, so both the primary and secondary tasks become prone to error. Thus, Chechile et. al. (1979) explain the length effect in terms of the number of memory retrievals required on the primary task.

In recent years a different approach to explaining the length effect has been proposed. Moran (1981) and Reisner (1981) have argued within the framework of structural or computational linguistics that the length effect is due to the length of the deep structure linguistic representation of the operator's task. This alternative approach will be illustrated later in the paper. Generally, the syntactic string length of the computational linguistic expression increases as a function of the number of operation required on the primary task. However, it is important to understand the basic mechanism of the length effect because there is not necessarily a perfect correlation between the number of memory retrievals and the syntactic string metric of structural linguistics. If, for example, syntactic string length is not a valid predictor of divided attention performance, then designing the human-software interface so as to minimize the syntactic string length metric would be counterproductive.

A series of three experiments have been conducted to contrast the memory hypothesis with the syntactic hypothesis. In each experiment the primary task consisted of editing route-way-point information. See Figure 1 for a typical route way display. In the upper right hand side of the CRT screen there were a series of random digits which appearred at the rate of one digit per .85 seconds. The secondary task of the experiment consisted of detecting a digit greater than six and denoting the event by pressing a specified key. Each subject was trained on all aspects of the experiment. This was immediately followed by

7

ID	ALT	SPD	PTA
CLARM	03400	380	23:01:20
FRANK	03700	400	23:04:40
ROCHE	03900	400	23:08:00
SANFO	04000	420	23:11:20

Figure 1. Sample route-way point display.

an experimental session which required the subject to make 25 separate edit changes of the route-way point information. The new information to replace current route-way data was presented for 5 sec. prior to the trial. The subjects were instructed to memorize the new data, but the subjects were not permitted to write the information down. The edit trials were self-paced, and the subjects could correct errors in the entered data in two ways. The first way was by means of a deletion key which functioned in the usual way during the data entry stage. The second way was by means of cancelling the entire results of the all the prior steps in the edit procedure and hence returning to the first step in the edit process.

EXPERIMENT 1

In this study, four different route-way point editing procedures were examined in order to understand the reasons for the length effect. A brief description of the four procedures is contained in Table 1. Each procedure began with calling up the route-way display. Also each procedure had a data entry step followed by an end of data (EOD) step. In terms of number of memory retrievals, the data entry step is considered as being 5.33 steps, on average, since 1/3 of the trials consisted of editing speed information (3 characters), 1/3 of the trials consisted of editing altitude information (5 characters), and 1/3 of the trials consisted of editing projected-time-of-arrival information (8 characters). Hence, in terms of the average number of memory retrievals, Procedure 1 required 9.33 retrievals, Procedures 2 and 3 required 11.33 retrievals, and Procedure 4 required 13.33 retrievals. The primary difference between Procedures 2 and 3 corresponded to when the data entry stage occurred in the command sequence.

These same four editing procedures can be described in terms of syntactic expressions within a formal grammar. The first step in the development of a grammar for route-way-point editing is to define a vocabulary. A vocabulary that is adequate for all the editing procedures used in all three experiments is shown in Table 2. Key symbols in the grammar are S, the beginning of a syntactic expression, and s, the end of the expression. The other symbols refer to either specific steps in the editing sequence, e.g. I for initiating the route-way-point display, or logical operations, e.g. ==> for the concatenation relationship. In terms of this vocabulary, a unique syntactic expression can be written for four edit procedures described in Table 1. These syntactic expressions are shown in Table 3. Within the formal grammar, the measure of the complexity of an expression, e(i), which is recommended by the advocates of syntactic analytic approach, is the string length, $|e(i)|$. The string length measures are also shown in Table 3.

The design of the four edit procedures leads to differential predictions regarding human performance. According to the memory retrieval hypothesis, Procedure 1 should be superior

Table 1. Description of the steps required for each of four Route-Way-Point (RWP) editing procedures.

Editing Procedures			
1	2	3	4
Call up RWPs	Call up RWPs	Call up Rwps	Call up Rwps
Enter Data	Enter Data	Identify I.D.	Denote Edit
EOD	EOD	Identify Value	Enter Data
I.D./Value	Identify I.D.	Enter Data	EOD
Exec. or Rej.	Identify Value	EOD	Call back RWPs
----------	Call RWPs back	Call RWPs back	Identify I.D.
----------	Exec. or Rej.	Exec. or Rej.	Identify Value
----------	----------	----------	Call RWPs back
----------	----------	----------	Exec. or Rej.

Note: Data Entry Averaged About 5.33 steps for each.

Table 2. Symbol definition for an action grammar for the edit task.

SYMBOLS	DEFINITION
S	Start of syntactic expression
s	End of syntactic expression
I	Initiate Route-Way-Point display
C	Change Route-Way-Point option
N	No Change in Route-Way option
D	Data before Parameter option
P	Parameter before data option
Bd	Begin data entry
#	Data entry
EOD	End of data entry
A(1)...A(12)	Specific location/parameters in display
O(4)	Four step option
O(5)	Five step option
R	Reject edit change
E	Execute edit change
/	Exit from Route-Way-Point procedure
=	Consists of relation
v	Logical OR operation
x ==> y	y follows x operation, i.e. concatenation
( )	Associative closure symbols

Table 3. Syntactic description of the four editing procedures.

e(1) = S ==> I ==> # ==> EOD ==> J ==> (E ==> s) v (R ==> S)

where J = (A(1) v A(2) v ... A(12)) . Note, |e(1)| = 45 .

e(2) = S ==> I ==> # ==> EOD ==> K ==> K ==> I ==> (E ==> s)

v (R ==> S) , where K = (A(1) ... A(4)).

Note, |e(2)| = 41 .

e(3) = S ==> I ==> K ==> K ==> # ==> EOD ==> I ==> (E ==> s)

v (R ==> S) , where |e(3)| = 41.

e(4) = S ==> I ==> C ==> # ==> EOD ==> I ==> K ==> K ==> I ==>

(E ==> s) v (R ==> S) , where |e(4)| = 45.

to Procedures 2 and 3 which in turn should be superior to Procedure 4. However, according to the syntactic string length hypothesis, Procedures 2 and 3 should be equivalent and superior to Procedures 1 and 4 which also should be equivalent.

A total of 100 undergraduates of Tufts University served as subjects in the experiment and received course participation credit. For each editing procedure, 25 subjects were randomly assigned, but five of these subjects in each edit procedure were tested only with the primary task as a control.

RESULTS AND DISCUSSION

The group means for the edit time and edit accuracy are shown in Table 4. Notice that the means for the control subjects who only performed the primary edit task are in each case shorter in edit time and greater in edit accuracy. Because of the unequal sample size between the experimental and control conditions, a series of t tests were performed. For edit time the difference between the experimental and control is statistically significant for Procedure 1, $t(23) = 2.39$, $p < .05$, and Procedure 3, $t(23)=2.91$, $p <.01$. For edit accuracy the difference between the experimental and control is statistically significant for Procedure 3, $t(23) = 2.33$, $p < .05$, and Procedure 4, $t(23) = 2.44$, $p < .05$. Consequently, there is a significant effect of the secondary task in reducing edit accuracy and increasing the time to edit.

Several dependent measures showed significant variation as a function of type of editing procedure. There is a significant difference across the edit designs for edit time, $F(3,76)=3.69$, $MSE=391.02$, $p<.02$, for edit accuracy, $F(3,76)=4.48$, $MSE=.09$, $p<.006$, and for the number of rejections of the prior steps, $F(3,76)=2.84$, $MSE=.12$, $p<.04$. The rejection means for Procedures 1 through 4 are respectively, .24, .17, .21, and .11. The hit and false alarm rates on the secondary digit detection task did not differ significantly across the various editing procedures with the respective averages being .415 and .0625.

Of central interest is the relative explanatory power of the predictors of divided attention performance. To address this question, regression analyses were done using edit time and edit accuracy as criterion variables and using syntactic string and the average number of memory retrievals as predictors variables. The predictor of syntactic string length was not significantly correlated either to edit time, $r =.00314$, or to edit accuracy, $r =.01351$. However, the average number of memory retrievals predictor was highly correlated with the edit time, $r =.78315$, as well as edit accuracy, $r =.81757$. Consequently, the results of Experiment 1 provide strong support for the Chechile et. al. (1979) explanation in terms of the number of memory retrievals as the mechanism underlying of the length effect in divided attention. However, it may also be possible that there simply was not enough manipulation of the syntactic string length across the four editing procedure. Using edit procedures with greater

Table 4. The mean edit time and edit accuracy for all groups.

Mean Edit Time		Mean Edit Accuracy		Predictors	
Control	Exper.	Control	Exper.	|e(i)|	n
18.7	27.7	.95	.88	45	9.33
25.0	32.0	.96	.84	41	11.33
18.0	28.5	.96	.79	41	11.33
32.3	33.9	.94	.77	45	13.33

differences in syntactic string length might result in some confirmation of the prediction that originates from the formal language approach. The possibility that syntactic length was not varied in an extreme enough fashion in Experiment 1 will be explored in the next experiment. Note, however, that Experiment 1 did show large and important differences in divided-attention performance that are accurately explained in terms of the memory retrieval hypothesis.

EXPERIMENT 2

The purpose of Experiment 2 is to vary more widely the syntactic string length between editing procedures without increasing the average number of memory retrievals. Towards that end, three new editing procedures were developed with the first two procedures requiring an average of 11.33 memory retrievals, and the third procedure requiring either 9.33 or 10.33 retrievals depending on the subject's choice. A syntactic representation of the three procedures is shown in Table 5. Procedure 1 consisted of first calling up the route way display, followed by the data entry and end of data steps. The operator then pressed one of four keys which corresponded to the next four route-way locations. Once a location key was depressed, then a separate display was shown for the specified location. The operator then depressed one of three keys which specified the parameter. Next the operator called back the full route-way display, and lastly the operator decided between executing the change or rejecting the change. In terms of the syntactic string length, Procedure 1 has a length of 39. The syntactic string length for Procedures 2 and 3 are 105 and 142, respectively. Basically the reason for the marked increase in string length is associated with the fact that the last two procedures provide a number of options to the operator, and when these options are expressed in terms of the action grammar, the string length increases. In Procedure 2, the operators had the option (after calling up the route-way display) of either editing the display or not editing the display. Furthermore, if the operator decided to edit the display, then there was an option of either entering the data first or parameter/location first. In Procedure 3, the operator had the immediate option of whether or not to have the option of data or parameter first. If the subject chose not to exercise the choice of the parameter and data order, then the default sequence was a data first entry system.

If syntactic string length influences performance on a shared-attention task, then Procedures 1, 2, and 3 should be increasingly more difficult conditions. In order to examine this question, 81 subjects (27 per editing procedure) were tested in a divided-attention study which followed the same method as that used in the prior experiment.

Table 5. Syntactic description of the editing procedures for Experiment 2.

Procedure 1: 11.33 steps with a string length of 39.

e(1) = S ==> I ==> # ==> EOD ==> K ==> L ==> I ==> (E ==>s)
v (R ==> S) , where L = (A(2) v A(3) v A(4))

Procedure 2: 11.33 steps with a string length of 105.

e(2) = S ==> I ==> (C ==> (D ==> # ==> EOD ==> J ==>
(E ==> s) v (R ==> S)) v (P ==> J ==> # ==> EOD ==>
(E ==> s) v (R ==> S))) v (N ==> s) .

Procedure 3: 9.33 or 10.33 steps with a string length of 142.

e(3) = S ==> (O(4) ==> # ==> EOD ==> J ==> (E ==> s) v (R ==>S))
v (O(5) ==> (D ==> # ==> EOD ==> J ==> (E ==> s) v
(R ==> S)) v (P ==> J ==> # ==> EOD ==> (E ==> s) v
(R ==> S)) .

Table 6. Results for Experiment 2.

MEASURE	11.33 steps length 39	11.33 steps length 105	9.33/10.33 steps length 142
Edit time	29.09	26.05	21.76
Accuracy	0.94	0.85	0.87
Pacing	1.89	1.72	1.74
Predicted accuracy	0.90	0.88	0.88

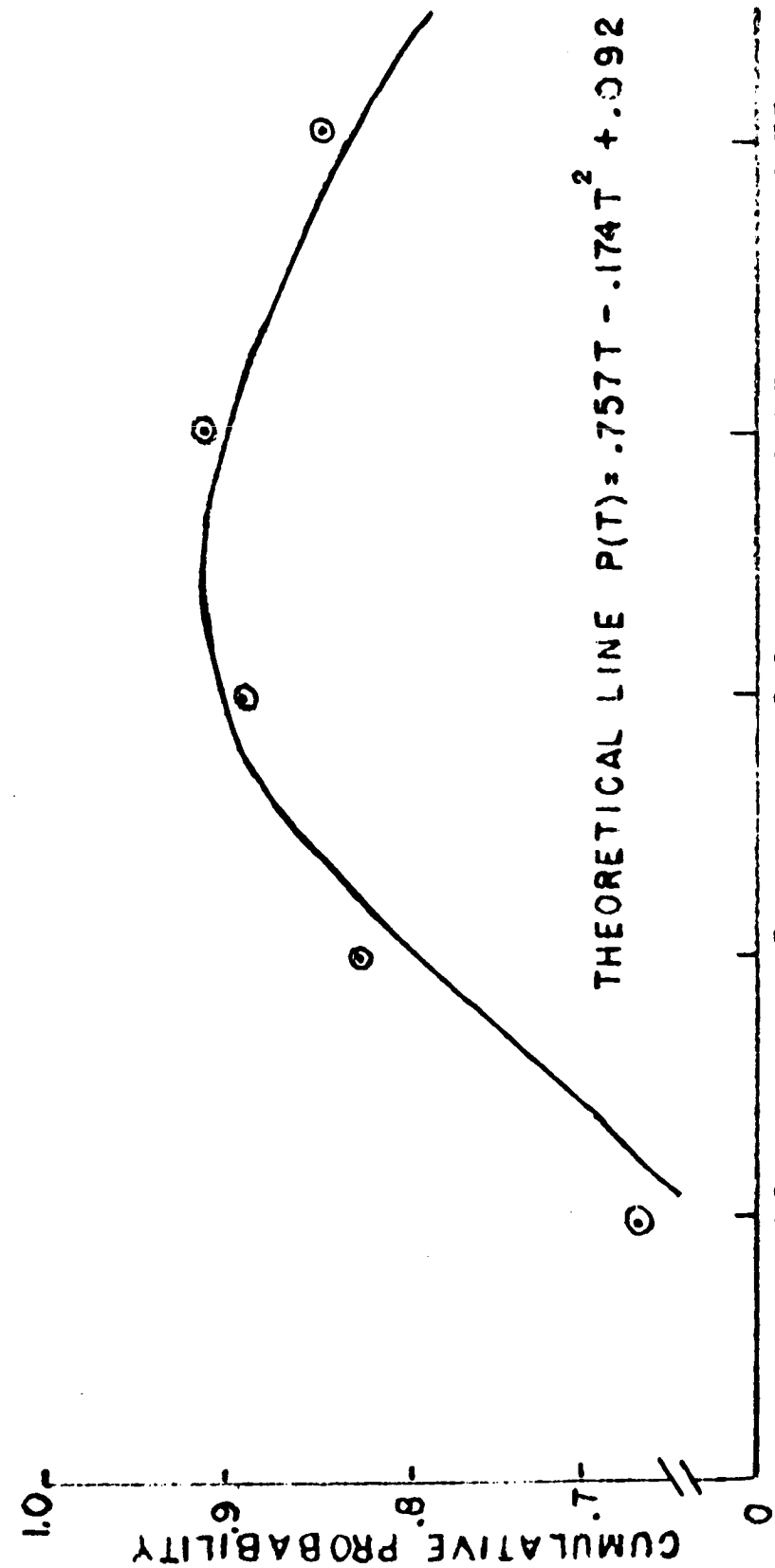


Figure 2. Mean cumulative probability correct on the primary task (averaging across the number of steps and feedback conditions) as a function of temporal pacing.

RESULTS AND DISCUSSION

The major results from Experiment 2 are shown in Table 6. Contrary to the predictions from the syntactic string length hypothesis, the edit times significantly differed between three edit procedures, $F(2,78)=7.05$, $MSE=51.91$, $p<.01$, but in exactly the opposite direction. The fastest edit time is for Procedure 3 which has longest string length, and the slowest edit time is for Procedure 1 which has the shortest string length.

In terms of edit accuracy, there is also a significant difference across the three procedures, $F(2,78)=8.34$, $MSE=.0072$, $p<.01$. While the condition with highest accuracy is the procedure with the shortest string length, there is, nevertheless, no lower accuracy for Procedure 3, with the longest syntactic string length, than there is for Procedure 2 which had intermediate string length. Initially the high accuracy for Procedure 1 seemed most puzzling. However, a more fine-grained analysis showed that the average time per step on the primary task was longer for Procedure 1 than Procedures 2 and 3. Perkins (1981) showed on a similar divided-attention task that there was an inverted-U shaped function of accuracy as the temporal pacing of the primary task varied. Figure 2 is a graph of the Perkins (1981) results. Notice that the optimal pacing on the primary task is about 2 sec. per keystroke which is a value close to the calculated time per keystroke on the primary task with Procedure 1. As shown in Table 6, the time per keystroke on the primary task is faster with Procedures 2 and 3. Perhaps the high edit accuracy for Procedure 1 is a result of slower pacing by the subjects. To explore that possibility further, a Chi-Squared statistical test was performed between the predicted number of correct edits based on the Perkins (1981) inverted-U shaped function. The test failed to show a departure from predictions based on simply temporal pacing, i.e. $\chi^2=1.97$, $p>.5$. Consequently, the high accuracy in Procedure 1 is most likely an artifact of slower pacing on the primary task. We believe the reasons that underlie the slower pacing are due to the three changes in the information on the CRT screen in Procedure 1. The three changes in the displays occurred by means of a relatively slow scroll which could have functioned to slow the subjects down, and thereby increase their edit accuracy.

The other performance measures (i.e. the rejection rate, the digit detection hit rate, and the detection false alarm rate) did not significantly vary across the three designs. The averages for these three measures are .097, .42, and .07, respectively, for the rejection, hit, and false alarm rates.

EXPERIMENT 3

The research design of the prior two experiments did not totally separate the number of memory retrievals factor from the

syntactic string length factor. A factorial analysis of variance design, however, does have the feature of providing an orthogonal separation of these factors. For this reason, Experiment 3 consists of a 2 x 2 x 2 completely randomized factorial analysis of variance design with the three factors being the average number of memory retrievals, the syntactic string length, and amount of syntactic interference. The first factor of average number of memory retrievals is either 11.33 or 13.33. The second factor of syntactic string length was either high or low where the high condition consisted of adding 46 to the low string length value. The third factor of syntactic interference is introduced here to assess a different notion discussed by advocates of the syntactic analysis for human factors designing, namely the consistency of syntax. In the low interference condition each keystroke is unique in its meaning, whereas in the high interference condition the same keystroke had two different meanings. For example, Reisner (1981) has argued that the multiple use of the same step for different operations would increase the complexity of syntax and would lead to an increased tendency for errors in human performance. Including the third factor into the research plan, allows for a direct test of Reisner's hypothesis.

A total of 104 undergraduates (13 per condition) served as subjects in a similar divided attention task as the one described previously. In the low syntactic string length condition, the subjects could change the route-way information by means of a data first entry system. In the high string length condition, the subjects could use either a data-first or a parameter-first system with the option adding 46 symbols to the action grammar syntax. In the high interference condition, the function key that ended data entry was the same as the key used to exit from the route-way procedure itself. In the low interference condition, two separate keys were used to perform the separate functions.

Also, in Experiment 3 a record was kept of the learning rates of the subjects. For example, syntactic string length may not influence divided attention performance once the operator learns the task. However, syntactic string length might be very important predictor of the learning of the task. Consequently, the learning rate information is required in order to evaluate this hypothesis.

Finally, a subjective workload rating was added as a new feature in Experiment 3. After 20 edit trials, the subjects were asked to give a workload rating on a 1 to 20 scale. The subjects had been instructed on the use of the workload rating scale prior to the divided-attention task. The reason for including the workload measure was simply to relate the same theoretical questions of the memory retrieval hypothesis and the syntactic hypothesis to the perceived workload of the operator.

RESULTS

Table 7. The summary results of three separate ANOVA's for Experment 3 on the edit accuracy, edit time, and workload measures.

		F VALUES		
SV	DF	ACCUR.	TIME	WORKLOAD
MEMORY STEPS	1	3.77 *	11.94 **	5.08 *
STRING LENGTH	1	0.02	1.76	0.04
MEM. X LENG.	1	0.02	1.34	0.00
INTERFERENCE	1	0.14	6.39 *	1.43
MEM. X INTER.	1	0.62	0.11	0.32
LENG. X INTER.	1	0.04	0.01	0.04
MEM.XLEN.XINT.	1	0.00	1.06	1.61

	MEANS	
	11.33	13.33
ACCURACY	0.87	0.83
TIME	26.65	31.16
WORKLOAD	15.20	16.40

The results of the analysis of variance on the edit accuracy, edit time, and workload ratings are shown in Table 7. The findings are quite consistent, and clearly show that only the number of memory retrievals is a statistically significant factor for all three measures. The means for the 11.33 and 13.33 memory retrieval conditions are also shown in Table 7. The means indicate that more memory retrievals increases edit time, reduces edit accuracy, and increases workload. Moreover, none of the other factors are statistically significant except for the interference factor which is significant on the edit time measure. The means for the edit times of the low and high interference conditions are 27.26 sec. and 30.56 sec., respectively.

The number of trials and the training time was used to evaluate the learning phase of the experiment. For these measures, the syntactic string length factor was the only source term in the analysis of variance that was statistically significant, $F(1,96)=29.8$, $MSE=13.54$, $p<.0001$ for the number of trials measure, and $F(1,96)=12.97$, $MSE=65.4$, $p<.0005$ for the learning time measure. The means indicate that the high syntactic string length condition requires about 6 additional minutes of training and about 4 additional training trials. Apart from the speed of learning the task, the syntactic string length metric did not predict performance on any divided attention measure.

DISCUSSION AND CONCLUSIONS

The results from the three experiments strongly indicate that the syntactic string length hypothesis is incorrect. The effect of syntactic string length is to influence the learning of the task. However, once the computer task is learned then the execution of the task in a time-sharing fashion can be achieved equally well for long and short syntactic strings. The three experiment plus the earlier work by Chechile et. al. (1979) and by Perkins(1981) show that the number of memory retrievals is a critical factor in an attention-sharing task. Consequently, in the design of the human interface with the computer system it is extremely important to minimize the number of memory retrievals required of the operator. If computer interaction is the only responsibility of the operator, then the memory retrieval process is not subject to very many errors. However, when the operator must also share attention with other instruments and other tasks, then the process of retrieving information from memory can require more cognitive effort than can be provided easily under the time pressures of the shared-attention task. Hence, both the time to complete the primary task as well as the accuracy on the primary task is influenced negatively.

One benefit of the null effect of syntactic string length on divided-attention performance is to free the software engineer to use longer string lengths. For example, the provision of

options to the operator greatly lengthens the syntactic string length. These options might make the software more general or more powerful. The results of the three experiments presented in this paper indicates that there is no apparent cost in terms of divided-attention performance of increasing the syntactic string length as long the number of memory retrievals do not increase. The only cost of lengthening the syntactic string would be in the time it takes to train the operator on the task.

REFERENCES

1. Chechile, R. A., Butler, K., Gutowski, and Palmer, E. A. Division of attention as a function of the number of steps, visual shifts, and memory load. Proceedings of the Fifteenth Annual Conference on Manual Control, 1979, 71-81.

2. Moran, T. P. The command language grammar: A representation for the user interface of interactive computer systems. International Journal of Man-Machine Studies, 1981, 15, 3-50.

3. Perkins, R. The effects of temporal pacing, interruption, and feedback in a divided attention task., Master Thesis, Tufts University, 1981.

4. Reisner, P. Using a formal grammar in human factors design of an interactive graphics system. IEEE Transactions on Software Engineering, 1981, SE-7, 229-240.

Manual Control Interfaces for the Tremor-Disabled

Michael J. Rosen, Principal Research Scientist*
Patrick O. Riley, Doctoral Candidate**
Bernard D. Adelstein, Doctoral Candidate*

*Mechanical Engineering Department
Massachusetts Institute of Technology

**Harvard-M.I.T. Division of Health Sciences and Technology

Introduction

From a clinical standpoint, the management of abnormal involuntary movement is a worthwhile target for reseach because of the substantial numbers of people involved, because of the potential for return of function and because of the intractability of many cases to conventional modes of treatment. Abnormal involuntary movement is seen in cerebral palsy, multiple sclerosis, Parkinson's disease, Friedreich's Ataxia, chronic alcoholism, cerebellar injury, monosymptomatic essential tremor and other congenital, acquired and degenerative neurological conditions. It is difficult to establish from available statistics the number of people in the population of the United States, for example, who have lost or impaired motor function due to involuntary movement. Incidence statistics for major disease states which may produce involuntary movement along with estimates of percentages of those populations in which it does, in fact, appear lead to an estimated total of 800,000.

The work described in this report is aimed specifically at investigating the effects of control interface characteristics which are hypothesized to provide improved accuracy in the presence of pathological "intention tremor". This abnormality, commonly seen in people with multiple

sclerosis and lesions of the cerebellum and basal ganglia, may be observed as a rhythmic oscillation of a limb induced or aggravated by attempted voluntary movement. Its amplitude may be comparable to purposeful acts in severe cases, so that loss of independent function may be complete. The frequency of oscillation is dependent on the limb segment studied but frequencies in the 2 to 6Hz range are typical. In earlier work of the authors and their colleagues (1), it was observed that this tremor contaminates voluntary activity in a simple additive way (so that, for example, voluntary tracking of a periodic target movement may be extracted from movement records by ensemble averaging triggered in synchrony with the target). Intention tremor may be disabling in spite of functionally adequate strength.

The functional penalty entailed by the presence of tremor is obvious. A long term goal of the authors and their colleagues in the Biomechanics and Human Rehabilitation Group is to develop compliant orthoses which apply loads across joints, loads chosen to selectively attenuate tremor with respect to general voluntary limb activity. While that aim is actively being pursued, this project is concerned with the technologically more limited goal of determining optimal chartacteristics for interfaces to environmental control systems, non-vocal communication devices, and electric wheelchairs. The design of proportional controls which allow more accurate manipulation

of such commercially available assistive technology can lead to substantial gain in independent function for tremor patients.

As detailed below, this research is investigating both mechanical properties and signal processing aspects of control interfaces. The expectation that tremor may be modified by these means arises from several sources. The literature of both psychology of movement and physiology includes demonstration that parts of the spectral content of normal control of limb position may be adjusted by manipulation of the visual display through which the task is presented (2). Normal physiological tremor has also been shown by numerous authors (3,4) to be adjustable in amplitude and frequency by application of inertial and elastic loads. The experiments of Stiles and colleagues are notable examples (5,6). While far less frequently reported, modulation of pathological tremor has also been observed both clinically (7) and experimentally (8,9). The present authors and colleagues have reported experiments in which viscous damping was successful in producing selective reduction of intention tremor during one degree-of-freedom tracking tasks (1).

The feasibility of mechanical modification of tremor is also suggested by hypothetical tremorogenic mechanisms. Systems theory suggests three models or contributing phenomena. The biomechanical resonance hypothesis (10) credits the underdamped muscle-limb plant with tuned

oscillations driven by broad-band muscle force oscillations. Closed neuromuscular loops are also capable of limited unstable oscillations (11,12), especially considering the presence of delay elements and pathologically increased gain. Simplest of all, there is the possibility of a Central Nervous System oscillator which drives the neuromuscular loops at tremor frequencies (13,14). [At this point, the work of the authors and colleagues suggests a dominant role for the third mechanism (15)].

Under each of these hypotheses, mechanical loading is a rational approach. Force sensing under isometric restraint eliminates the limb mass term from the biomechanical plant, reducing the system order and hypothetically eliminating the possibility of resonance. Applied loads may also be viewed either as a potential source of series compensation for loop oscillators or mechanical filtering given a central drive.

The objective of this project has been to test the effectiveness of particular interface characteristics for improving control accuracy of tremor-disabled individuals. The goal in the first year of study was to observe the range of effects which could be obtained rather than to establish statistically significant generalizations for each clinically distinct movement disorder group. This work represented an extension of earlier work by the investigators (16) on tremor in three respects. Whereas past studies required that subjects generate one degree-of-freedom movements, this project required control

in two degrees-of-freedom and involved several muscle systems. A second change was the focus of this work on development of interface hardware and processing for control of electronic assistive devices, rather than optimal mechanical characteristics for wearable compliant orthoses. In addition, the control tasks subjects were asked to perform were less abstract, including discrete target acquisition to simulate menu selection, as well as dynamic tracking.

Hypotheses and Methodology

Test subjects were required to perform a series of visual target tracking task in two dimensions. The tests were performed using the experimental setup shown in Figure 1. The subject sat either in a chair provided or in his/her own wheelchair facing the oscilloscope on which the target and cursor were displayed. The joystick was rigidly mounted either to the wheelchair or to a lab table with a fixture which permitted height and attitude adjustment. The weight of the subject's forearm was supported by a commercial feeding assist orthosis which permitted movement in the horizontal plane and rotation in the vertical plane about the forearm support point. This arrangement minimized subject fatigue, provided a degree of trial-to-trial uniformity, and allowed the subject to choose the way the joystick was grasped. At the same time the subject's arm was free to undergo tremor in most degrees-of-freedom.

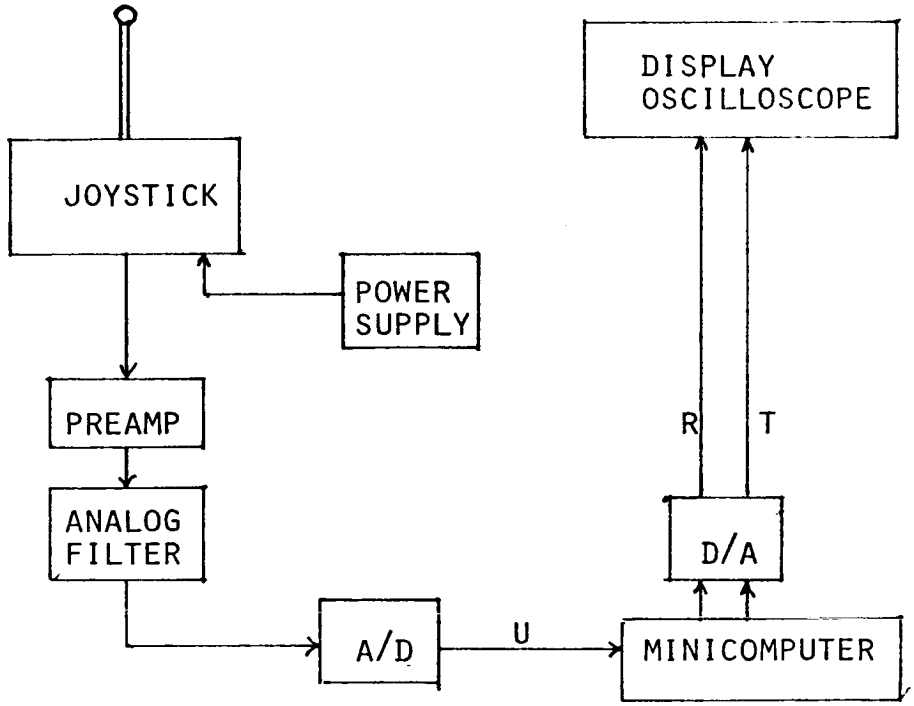


Figure 1. Test Setup

The minicomputer used was an ADAC System 2000 which incorporates A/D and D/A capability. Two joysticks were used; a commercial two-axis force sensing joystick with a stiffness permitting only 5 thousandths inch deflection for a 20 pound applied force, and a commercial displacement joystick modified to provide maximum compliance. In the experiments to date, the analog filter was used to filter the joystick output to below half the sampling frequency. Filtering of the digitized joystick output for tremor attenuation was done with a digital filter implemented in software. The target and response were sampled at 40Hz and all data recording was digital.

For each type of tracking task a number of parameters were used to evaluate subject performance. The continuous target tracking error was measured as ΔH_{TR} defined as :

$$\Delta H_{tr} = |1.0 - H_{tr}|$$

$$\text{where } H_{tr} = \int G_{tr}(\omega)d\omega / \int G_{tt}(\omega)d\omega$$

with the integrals evaluated over the frequency band of target power (less than approximately 2Hz).

The tremor power was defined as:

$$\text{TREMOR POWER} = G_{rr}(\omega)\{1-\gamma^2_{tr}(\omega)\}d\omega$$

where $\gamma^2_{tr} \equiv$ coherence of the target and response.

Here the integral was evaluated from 2Hz to the anti-aliasing filter frequency ($\sim$14Hz). The signal to noise ratio was defined as:

$$\text{SNR} = \int G_{tr}(\omega)d\omega / \text{TREMOR POWER}$$

where the frequency integral was evaluated over the frequency band of interest. The control effort factor was

$$U_{fac} = \int U^2 dt / \int T^2 dt$$

where T is the target signal and U is the digitized joystick output.

In the case of discrete target tracking the parameters measured were Settling Rate, Settling Time, Acquisition Rate, Acquisition Time, and Error Power. Settling Rate was defined as the fraction of target transitions for which the response was within ± 20% of the transition magnitude of the

target for at least the last 0.5 seconds of the allotted period (approx. 6.5 seconds). Settling Time is the time from target transition to arrival of the response within the ± 20% band for those cases where "settling" occurred. Acquisition Rate is the fraction of target transitions in which the response came within and remained within ± 20% band for at least 0.5 seconds during the period. Acquisition Time is the time from target transition to the beginning of acquisition. Error Power was defined as the average of the square of the target to response distance from the beginning of the first acquisition until the end of the period (time of the subsequent target transition).

The experimental design is represented in Figure 2. This testing matrix permitted the head to head comparison

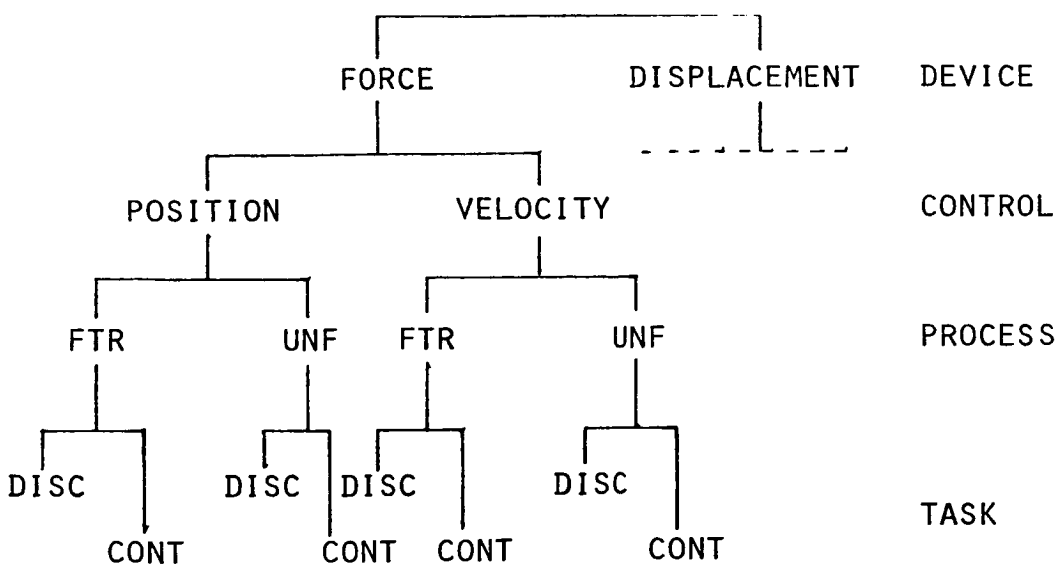


Figure 2. Experimental Design

of: 1) performance with filtering to performance without filtering, 2) performance using velocity control to performance using position control, and 3) performance using a force sensing device to performance using a displacement device. Since velocity control is a form of filtering, the combination was considered redundant and was deleted from later trials.

The experimental protocol was the same for each device and essentially the same with slight variation with respect to total number of trials for all subjects. For each sensing device and control mode the subject was allowed to familiarize himself with the setup by driving the cursor around the full extent of the screen, then to various points designated by the tester. The subject then practiced target tracking for two to three runs with either type of target. Testing with normals had indicated that this was sufficient practice to reach a plateau of nearly constant performance.

The subject then performed four to six discrete target tracking tasks, alternately without and with filtering. The subject next performed three to four continuous target tracking tasks at various gain settings, followed by four to six continuous target tracking tasks alternately with and without filtering. Each task lasted approximately forty-five seconds, with twenty-five seconds of data taken for processing from the middle. During the actual testing, trials were run at about two minute intervals.

This sequence of trials was run again for the alternate control mode, then the whole sequence would be repeated for he alternate device. Testing with a given device would last approximately one hour. Because of subject fatigue or possible loss of interest after extended testing, it was sometimes necessary to conduct the tests of each device on separate days. The order of device testing was random and at least one of the split-test subjects was back tested with the original device to assure constant performance. Repeated trials of each type would permit checking for learning; detailed statistical analysis has not been performed to date; but casual review of the data has shown no obvious trends.

Results

Four subjects participated in the initial series of tests, each with an action tremor of different etiology. Subject JBG is disabled by familial cerebellar ataxia while the tremor of subject CRS results from midbrain stroke. Subject EPS is cerebral palsied and subject RVM's movement disorder is secondary to multiple sclerosis. Due to the severity of the movement disorder, subject CRS participated in only discrete target tracking tests. Subject JBG was tested twice; the first time using his right hand, the second time using his left.

The following six tables represent a summary of the test results. Three comparisons were made: 1) performance with and without 4 Hz low-pass filtering of the joystick output signal; 2) performance using velocity control against performance using position control; and 3) performance using a force sensing (isometric) joystick compared to performance using a displacement joystick. In each case a plus (+) indicates that the parameter has changed in the direction of improved performance, i.e. higher SNR, lower tremor power, shorter settling time, etc. A minus (-) indicates degraded performance. An equal sign (=) indicates that the compared cases are equivalent at the 90% confidence level, and a zero (0) indicates that there was not adequate data for a comparison. Subjects' data are represented in the order JBG(R), JBG(L), EPS, RVM for the continuous target tracking task and JBG(R), JBG(L), CRS, EPS, RVM for the discrete target tracking task.

When comparing filtered and unfiltered tracking performance, only position control trials were considered since velocity control is itself a form of filtering. The continuous target tracking results summarized in Table 1 are consistent with the hypothesis that filtering improves SNR by reducing the tremor power component of the control signal.

Table 1. Filtered, with respect to Unfiltered Continuous Target Tracking

	SNR	Tremor Power	ΔH_{TR}	U_{FAC}
Force Joystick	++==	++==	====	0===
Displacement Joystick	==++	===+	====	0==+

Table 2. Filtered with respect to Unfiltered Discrete Target Tracking

Settling Rate	Settling Time	Acquisition Rate	Acquisition Time	Error Power
-==+=	===+=	===+=	=====	====-

An example of filtered and unfiltered tracking is presented in Figure 3. Neither the underlying tracking fidelity nor control effort is effected significantly by filtering. In one case SNR improved without any other parameters changing, probably reflecting a favorable, but less than statistically significant change in both tremor power and tracking error (ΔH_{TR}).

The results of the discrete target tracking test shown in Table 2 are more difficult to interpret. Fewer trials of this type were performed, hence the need to pool isometric and displacement joystick trials. In at least one case, subject EPS improved discrete performance with filtering was demonstrated.

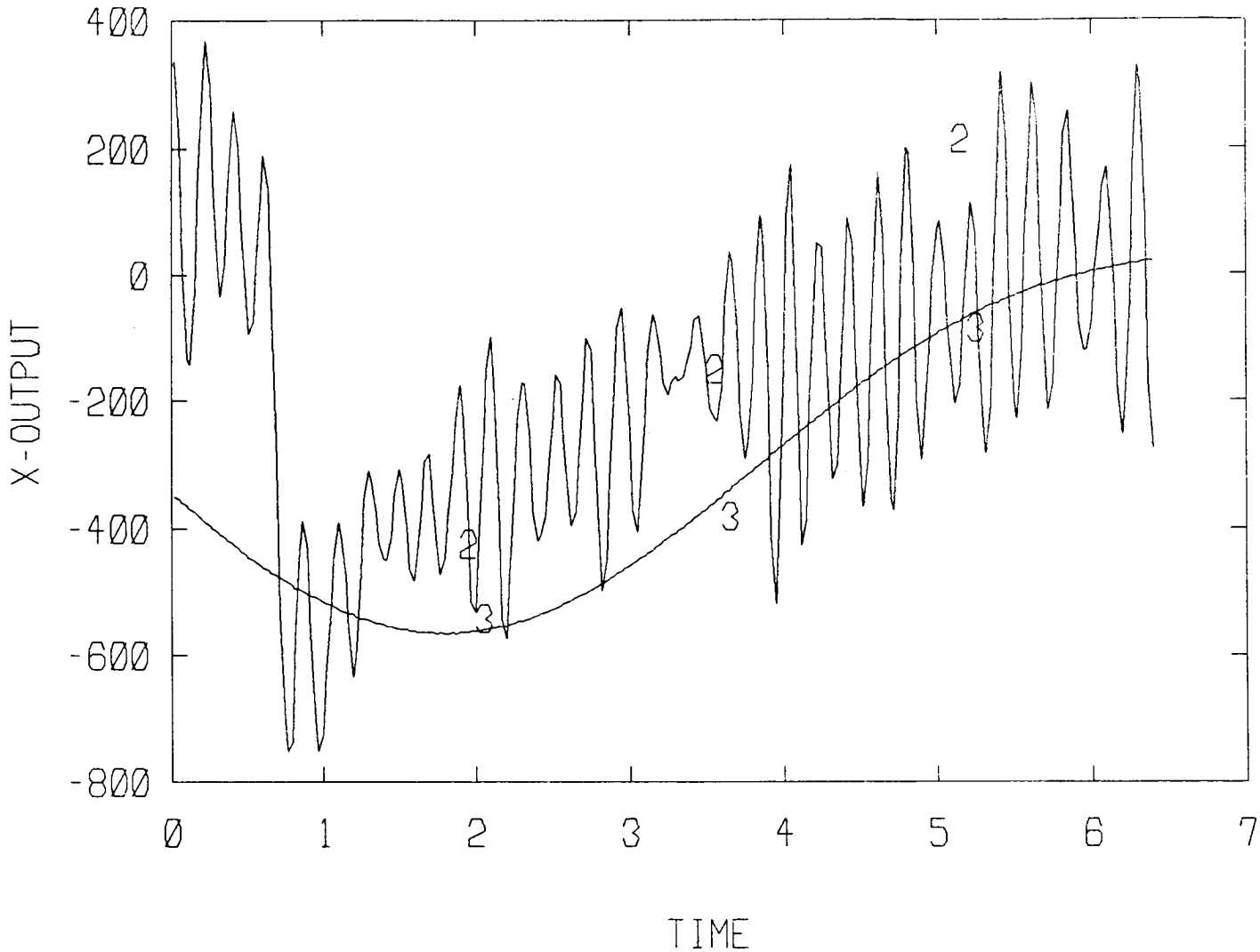


Figure 3a. Tracking Without Filtering

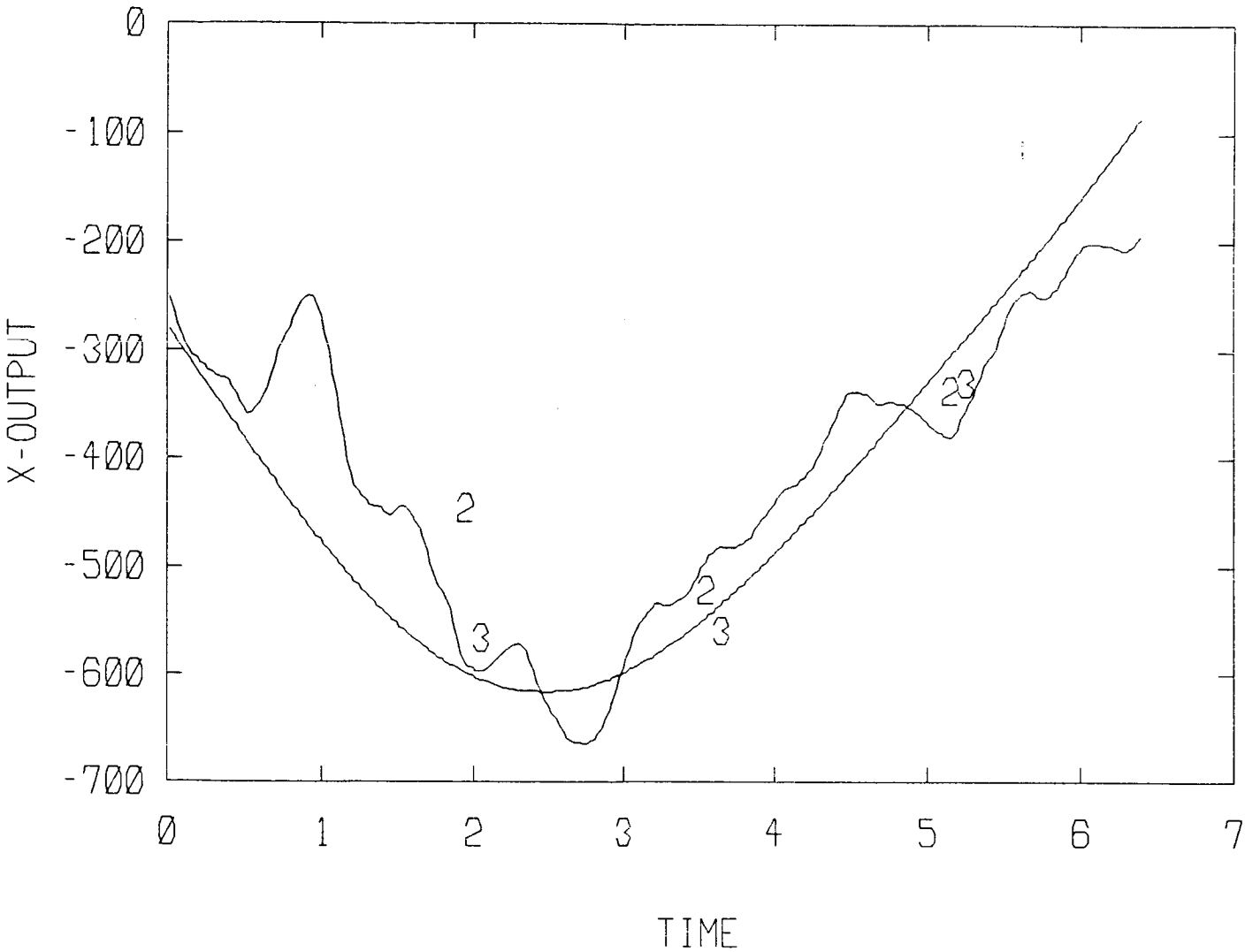


Figure 3b. Tracking With Filtering

Table 3. Velocity vs. Position Control - Continuous Target Tracking, Unfiltered

	SNR	Tremor Power	ΔH_{TR}	U_{FAC}
Force Joystick	++=+	++=+	===-	0=--
Displacement Joystick	-==+	-==+	=-=-	0---

Table 4. Velocity vs. Position Control - Discrete Target Tracking, Unfiltered

Settling Rate	Settling Time	Acquisition Rate	Acquisition Time	Error Power
-==+=	===+-	-=-++	==-=+	-+++=

Only unfiltered trials were considered for the velocity/position control comparisons. Table 3 indicates that velocity control also improves SNR by attenuating the tremor component of the control signal. The increase in tremor power (minus sign) was seen with subject JBG (R) using the displacement joystick. This result was not repeated during the left handed trials. There is a possible penalty of increased tracking error and a pronounced tendency to require additional control effort in the case of velocity control. The latter occurs despite the very high gain afforded by velocity control at very low frequency.

Again the results for discrete target tracking shown in Table 4 show a greater degree of variability. Velocity control does not appear to adversely effect acquisition time or settling time. In general, error when "on target", i.e. error power, seems to be lower when velocity control is employed.

Table 5. Force vs. Displacement Joystick
Continuous Target Tracking

	SNR	Tremor Power	ΔH_{TR}
Position Control	--=+	--++	---+
Velocity Control	===+	==++	===+

Table 6. Force vs. Displacment Joystick - Discrete Target Tracking, Unfiltered

	Settling Rate	Settling Time	Acquisition Rate	Acquisition Time	Error Power
Position Control	====-=	-=0-0	===-=	-===+	=-=-=
Velocity Control	=-=+=	+=+=+	=====	=====	=-++=

When comparing the force sensing joystick to the displacement joystick, neither type of test yielded a simple result pattern. In the case of position control one subject, JBG, showed degraded performance on both left and right handed trials while subject RVM showed improved performance in all areas. For subject EPS, decreased tremor power was offset by increased tracking error to yield no improvement in SNR. For velocity controlled continuous target tracking, the performance with the isometric joystick was equivalent or superior. The force sensing joystick showed better results with velocity control than with position control. These results, however, are not uniform for all subjects. Control effort was not compared due to lack of an adequate means of equating the effort score for isometric force generation to that for unopposed movement.

Discussion

The use of a low-pass filter with a 4Hz cutoff frequency appeared to improve SNR in continuous target tracking, principally by attenuating the tremor component of the control signal. Subsequent analysis of the tremor power frequency spectra of the subjects tested indicates that significant tremor power occurred in the frequency range from 2.5 Hz to 4Hz. For filtering to be maximally effective, therefore, use of a lower cutoff frequency may be necessary. A compromise may be needed to minimize attenuation in the frequency band needed for adequate control.

Velocity control also appears to improve SNR chiefly by attenuating tremor. The effect on tracking fidelity is not necessarily adverse, but may occur. When tracking fidelity is adversely effected, it may be due either to the subjects' lack of familiarity with this control mode, or to the effect of attenuating the control signal in the frequency range of the target thereby reducing the control responsiveness. The force-sensing joystick is not uniformly superior to the displacement joystick and, in fact, may sometimes be disadvantageous.

The usefulness of each interface characteristic appears from our data to be a function of the individual disorder and the other system features. Case by case assessment will be necessary until more generalized data for various etiologies has been

gathered. Each of our subjects in this study presented an action tremor due to a different disorder so that differences in response due to individual variation could not be distinguished from possible etiology-related effects. A much larger and broader sample would be required for the question to be evaluated statistically. The protocol would also have to include clinical evaluation and classification of the subjects.

There are at least three possible explanations for the degree of variability, subject-to-subject, in the results of the discrete target tracking task. First, fewer data points were obtained for each subject in each control made for discrete target tracking than for continuous tracking. It was desirable to test the subjects in a small number of sessions occurring within a short time of each other to minimize the effect of day-to-day and long term variability in their condition. The length of the individual sessions was also limited by subject physical and mental fatigue. These combined constraints limited the amount of data.

Secondly, the protocol scheduled discrete target tracking tests after familiarization and practice but before the continuous target tracking tasks. As a result, the continuous target tracking task may have been less contaminated by the effects of learning. The two to three practice runs which were sufficient to stabilize the performance of normal subjects may

not have been wholly adequate for pathological subjects. Here again, subject fatigue limited the amount of practice feasible.

The third possible source of variability may have been that, while a number of indices were tried in order to evaluate the discrete performance, all parameters which were highly sensitive to subject attention and alertness. In the case of the continuous tracking, tremor power is a relatively unconfounded measure of pathological movement, independent of how well the subject is tracking voluntarily. SNR is effected by tracking effort but not dominated by it. In contrast, the measures of tremor during discrete target tracking all depend on the assumption that the subject is putting equal effort into each response and is employing an unchanging strategy. Neither assumption is particularly robust.

Conclusions and Recommendations

The testing to date demonstrates that 4Hz low-pass on-line filtering can improve SNR with little or no increase in control effort or loss of tracking fidelity. Velocity control can also result in higher SNR, usually with some increase in control effort. Isometric control devices seem to offer an advantage over displacement devices only under specific circumstances, not well defined by the scope of the testing conducted so far.

Since the pathological action tremor of the upper extremity has significant power in the 2.5 to 4.0 Hz range, it appears advisable to evaluate the use of filtering with a cutoff frequency lower than 4.0 Hz. The effect of low frequency attenuation on control responsiveness and tracking fidelity will have to be evaluated.

To date, comparison of the results of on-line filtering to results obtained by post-filtering the input control signal of unfiltered tests has not shown any significant differences. This would seem to indicate that there is no interaction via the visual feedback path between the subject's motor output and the effect of filtering on the visual display process. This needs to be verified with additional subjects, hopefully using an on-line filter which produces a more statistically determinable effect.

References

1. Rosen, M.J., Dunfee, D.E. and Adelstein, B.D. Suppression of Abnormal intention tremor by application of viscous damping. Proceedings, 4th Congress Internat. Soc. Electrophysiol. Kinesiol., Boston, August, 1979.

2. Sutton, G.G. and Sykes, K. The effect of withdrawal of visual presentation of errors upon the frequency spectrum of tremor in a manual task. J. Physiology, vol. 190, pp. 281-283, 1967.

3. Joyce, G.C. and Rack, P.M.H. The effects of load and force on tremor at the normal human elbow joint. J. Physiol., London, 240:375-396, 1974.

4. Robson, J.G. The effect of loading upon the frequency of muscle tremor. J. Physiol., London, 149:29P-30P, 1959.

5. Stiles, R.N. and Randall, J.E. Mechanical factors in human tremor frequency. J. Appl. Physiol., 23:324-330, 1967.

6. Stiles, R.N. Mechanical and neural factors in postural hand tremor of normal subjects. J. Neurophysiol., 44:40-59, 1980.

7. Hewer, R.L., Cooper, R. and Morgan M.H. An investigation into the value of treating intention tremor by weighting the affected limb. Brain, 95:579-590, 1972.

8. Chase, R.A., Cullen, J.D., Sullivan, S.A. and Ommaya, A.K. Modification of intention tremor in man. Nature, 4983:485-487, 1965.

9. Ring, N.D., Brown, A.W.S. and Lord, M. An investigation into the control of involuntary movement. Future Priorities in Orthotics and Prosthetics, Biological Engineering Society, London, March, 1980.

10. Rietz, R.R. and Stiles, R.N. A viscoelastic-mass mechanism as a basis for normal postural tremor. J. Appl. Physiol., 37:852-860, 1974.

11. Stein, R.B. and Oguztoreli, M.N. Reflex involvement in the generation and control of tremor and clonus. Prog. Clin. Neurophysiol., vol. 5, ed. J.E. Desmedt, pp. 28-50 (Karger, Basel, 1978).

12. Watanabe, A., Kikuchi, H., Fukumoto, I. and Saito, M. Analysis of micro-vibration frequency of human limbs during stroke movement. Digest, 11th Int. Conf. on Med. Biol. Eng., Ottawa, 1976.

13. Joffroy, A.K. and Lammare, Y. Rhythmic unit firing in the precentral cortex in relation with postural tremor in a deafferented limb. Brain Res., 27:386-389, 1971.

14. Elble, R.J. and Randall, J.E. Motor unit activity responsible for 8- to 12-Hz component of human physiological finger tremor. J. Neurophysiol., 39-370-383, 1976.

15. Adelstein, B.D. Peripheral mechanical loading and the mechanism of abnormal intention tremor. Master of Science Thesis, Dept. of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, July, 1981.

16. Adelstein, B.D. and Rosen, M.J. The effect of mechanical impedance on abnormal intention tremor. Proceedings, 9th Ann. Northeast Bioengin. Conf., March, 1981.

Acknowledgement

This work has been supported in part by a grant from the National Institute of Handicapped Research, Department of Education, Grant # G008300074; by the Eric P. and Evelyn E. Newman Laboratory for Biomechanics and Human Rehabilitation Fund; and by a Medical Engineering and Medical Physics Fellowship from the Harvard-M.I.T. Division of Health Sciences and Technology.

The visual representation of the efforts in telemanipulation

J.P. GAILLARD, S. GALERNE, E. COLLE
(Université de Paris Val de Marne)

ABSTRACT

Nuclear industry requires activities in radioactive cells where humans cannot enter. Elements and objects are manipulated through a telemanipulation system which has been widely reported (e.g. Vertut, 1976). Recently new activities such as the dismandling of old nuclear reactors are being added. A rapid analysis of these new activities indicates that the actual conception cannot satisfy the level of activity required. In particular, operators would be required to supervise their action on a video screen. This mode of control incompatible with the actual master-slave manipulator which imposes the operators to stand up, to displace themselves and to produce large movments.

A new concept of a syntaxer without force feedback in the handle arises. A general study has been performed concerning the way to feed back to the operator forces exerted by the manipulator on the environment (Gaillard, 1981). Its results showed a visual dominance effect for the operator treating this kind of information.

The aim of the present study is to define the visual presentation of gripper forces on a video screen. A systematic comparison has been made between a classical vectorial presentation of their amplitude and orientation and a presentation involving an operator's representation of the forces exerted on a symbolic object.

I.- INTRODUCTION

In the present experiment we compare the effect on performance and training of different visual displays. Our aim was to facilitate the operator's treatment with a representation of the forces in order to reduce the cognitive cost to treat the information without affecting the performance.

Two approaches were studied :

1. Presenting the efforts as forces in a conventional way.

2. Presenting the efforts as non visible forces exerted on an object.

In the first manner vectors could represent direction and amplitude of the forces exerted by the manipulator efforts but not the effect on the environment.

In the second one forces exerted in telemanipulation remained "non visible" in their properties.

Cognitive treatment of such important information when manipulating objects could be expected to be more or less effective depending of the compatibility of the visual presentation of the efforts with their image in memory. We hypothetised that an adequation between the memorized image and the presentation of the forces on a screen should facilitate the operator's task. However, if we would present properties of the forces which are not part of the memory image we should render the operator's task more difficult as could be the case of the former approach.

In earlier experiments we studied the possibility to feedback information to an operator in a telemanipulation control task through another sensory channel than the visual one (Gaillard, 1981a, 1981b). Indeed Bertsche, Logan, Pesch and Winget (1977) noted that "utilization of a visual force feedback display required the operators to time share his visual patterns between the displays and the work surface". Then we tried to avoid this difficulty in testing an electrotactile display. Results clearly indicated that this way to feedback forces exerted by a telemanipulor was not satisfactory we found that the time completion task and the number of corrections were significantly higher when compared with a simple visual force feedback althoug the number of errors remained the same.

The interpretation we made about the increase of completion time was not contradictory to Winget's observations but more elaborate. We hypothetised that task completion time is varying in function of a number of factors which are contradictory in their effects.

1. A cognitive cost to treat informations displayed. Informations which are displayed and treated are time consuming.

2. An improvement in decision making. Informations which are treated allowed operators to make more efficient decisions ; reduce uncertainty and save time in the completion of a task.

However, observations made on the time to complete a task was not sufficient to draw conclusions. A shorter time could be the result of an improvement in decision making and not a lack of information feedback.

In the present experiment we used the number of failures as an additional dependent variable to evaluate the weight of the two main factors on the operators' performance.

An improvement of the decision making process could be obtained by an adequate presentation of the forces exerted by the telemanipulator on the environment as we suggested in representing these efforts as "non visible" forces exerted on objects.

II.- EXPERIMENTATION

1. Methods

Material : the experimental system was composed of a CEA-La Calhène MA 23 type manipulatof with six degrees of freedom plus Gripper. The manipulator was controlled by a syntaxer with six degrees of freedom. Three to control the translations in cartesian coordinates and three to control the orientation of the gripper. The velocities of the translations were controlled by micro-displacements of the syntaxer, whereas a position control of the orientations was used (Dupourqué, Gaillard, Quétin, 1981).

The forces on the three axes reported at the end of the gripper were presented on a video screen. Two types of presentations were used :

1. A conventional presentation of the forces exerted by the manipulator on the environment composed by four vectors : three vectors for each of the cartesian components and the resultant of these vectors (photo 1).

2. The presentation of the forces effectuated on a symbolic object, represented by a square on the screen, by the movments of both the

object and the shadows, projected both on a vertical plane and a horizontal plane simulated by a perspective drawing (photo 2).

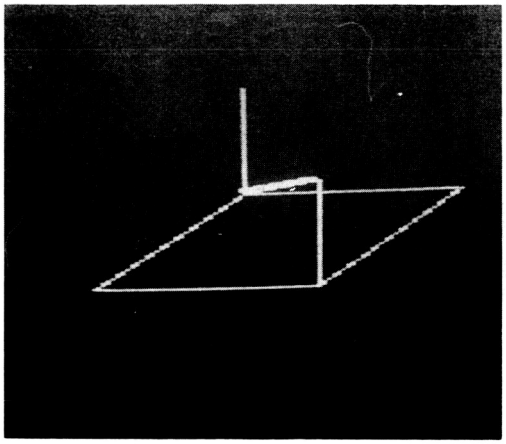


Photo 1

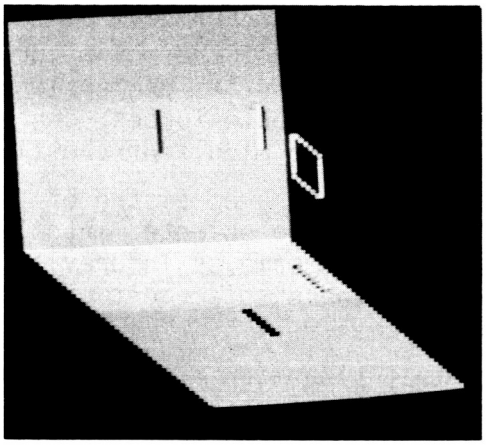


Photo 2

The presentations were implemented on an Apple II computer with as its input the torques of the three motors controlling the displacements of the manipulator corresponding approximately with cartesian coordinates.

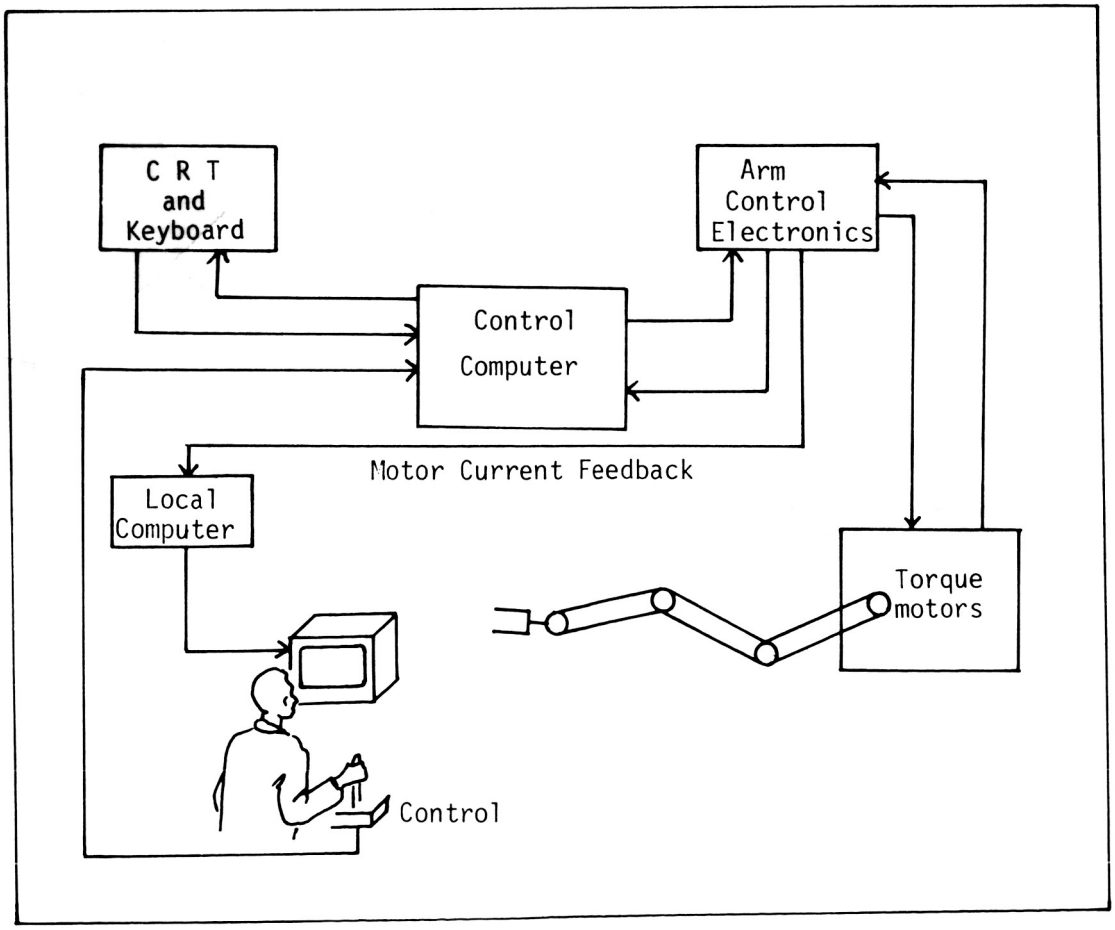


Figure 1 : The experimental system.

For the first presentation type :

The modules of the three cartesian components was proportional to the amplitude of the motor currents in the corresponding directions.

For the second presentation type :

The amplitude of the motor currents in the corresponding directions moved the square and its projections over the screen. The zero corresponding to negligeable forces, contened the square and its shadows on fixed marks in the middle of the figure.

Task : The task consisted of the displacement of a slide along a rod that could be oriented into 36 different possible directions. The task was the same as in previous experiments (Gaillard, 1981b) and consisted of the displacement of the slide along the rod with de manipulator.

In that situation a control command of the gripper not in parallel with the rod's orientation increased the forces exerted. To accomplish the task the operator had to displace the slide by controling the forces.

The execution time was measured automatically at each trial by a counter controlled by two microswitches placed at each end of the rod. In addition the number of failures was counted. A failure consisted of the triggering of an automatic safety stop when the forces reached a certain level, there by automatically uncoupling the manipulator.

Operators : Three operators were volunteers and had a training session before the experiment. The training consisted of an initiation to the control language for two of them and the execution two sets of trials, one for each type of presentation of the forces. The third operators was experienced in the use of the manipulator and had a short training before each session.

Design and procedure : three conditions accross 30 trials were defined as follows :

- 10 trials without synthetic feedback on the video screen. The operator had to work only by watching the manipulator and the rod. In that condition force information was deduced by rotation of the gripper constrained by the rod.

- 10 trials with a conventional presentation type of forces by vectors on the video screen, placed in front of operator.

- 10 trials with the second presentation type of forces exerted on the symbolic object, the square on the video screen.

The first condition was considered as a control condition and the two others as experimental ones. The orientation of the rod was modified at each trial in order to avoid that the operator may work in open loop condition. At each trial the experimenter fixed the new orientation of the rod, then the operator grasped the cursor with the gripper and started the trial for the control condition. Before the experimental conditions the operator looked at the new orientation, then the experimenter placed a plywood panel between the operator and the scene of manipulation to hide it to him. In that conditions the operator had to execute the trial, remembering the orientation of the stem and controling with the visual feedback of the forces on the screen.

At each trial the operator was informed of the execution time, with the chronoscope placed next to the video screen. When the chronoscope stoped it means that the cursor was at the end of the rod and the trial was finished successfully. Under neither one of the two experimental conditions the operators were not informed whether cursor was sliding on the rod or not.

III.- RESULTS

The task completion time for the different conditions was not significantly different (table 1).

Control condition	Visual "vectors"	Visual "square"
23,04	18,87	25,19

Table 1 : Mean task completion time in seconds

The failures rate calculated by the number of failures divided by the number of trials for each condition was as follows (figure 2).

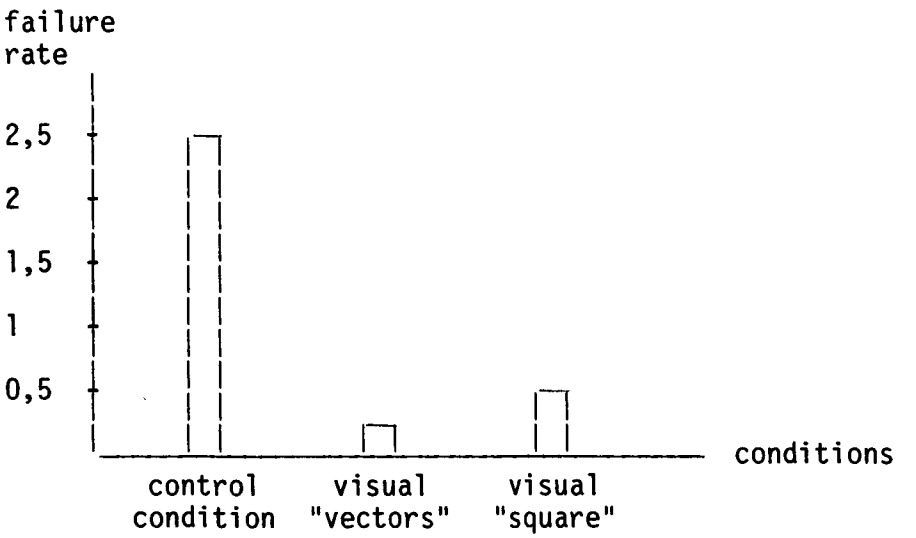


Figure 2 : Failure rate as a function of conditions

The learning effect was calculated by compairing the first five trials to the last five trials for each condition (figure 3).

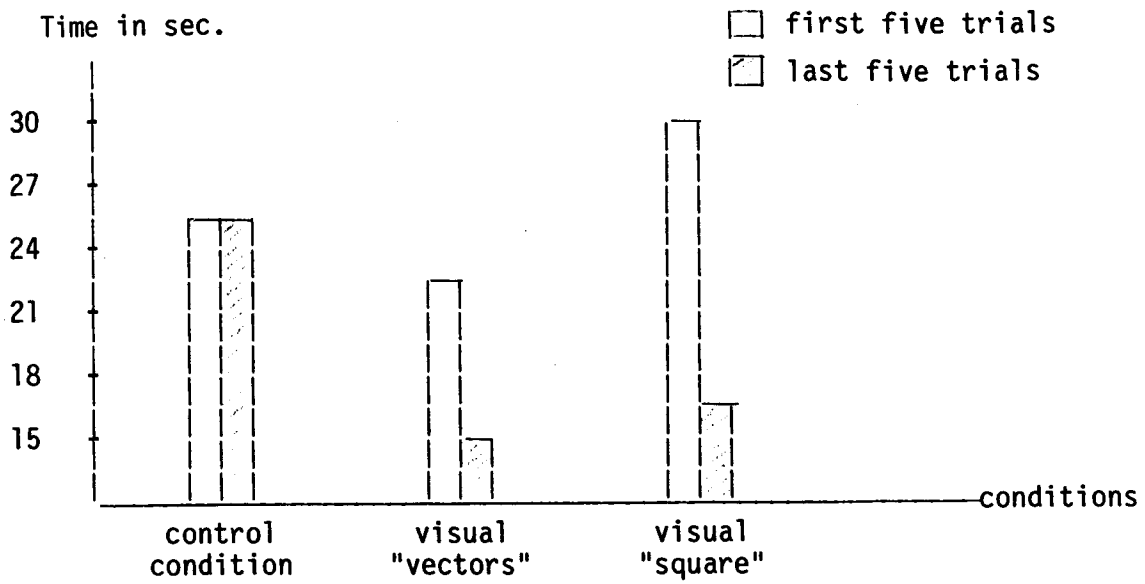


Figure 3 : Learning effect as a function of conditions

Comparisons with the t test were, respectively for the three conditions : $t = 0,03$ (28) ; $p > .10$; $t = 1,27$ (28) ; $p > .10$ and $t = 2,78$ (28) ; $p < .01$.

The mean times between the first part of the "vectors" condition were not significantly different to the first part of the "square" condition. $t = 1,24$ (28) ; $p > .10$.

IV.- DISCUSSION

The results on task completion time clearly indicate we are not allowed to speculate on the amount of feedback information to the operator under the different conditions. Two methodological interpretations are possible :

1. Due to the small number of subjects and the large variance obtained no difference in the amount of information processed could be established.

2. Under the experimental conditions the scene was not visible to operators, whereas under the control condition it was but forces were not fed back to them. One could argue that the mean quantity of information processed under the first condition was about the same as under the others.

However, the number of failures is more informative and distinguishes between different experimental conditions the task was facilitated by the forces feedback on the screen when the rod was out of view. Our opinion is based on the following observations. Indeed in the experimental conditions we feedback informations which was roughly well adapted to the type of task the operators had to realize. An adapted feedback information is the class of information which has values which may continuously vary during the task. This variation could be dependent both on operator actions and external factors.

We suggest that in remote control operation, informations which are not continuously varying should not be permanently presented. Ergonomic studies in aeronautic (Stanton and Cook, 1981 ; Falzon, 1982) lead to an analogous conclusion for two complementary reasons.

First, the increasing complexity of an aircraft produces an increase in the number of displays, and therefore amount of information presented. Computers should be used to share the treatment of the information with the pilot and present him the right information at the right moment. Secondly, the information presented should depend of the operator's cognitive representation of the task. These results reinforced an assumption that quantitative aspects which could be analysed with the task completion time have to be completed with qualitative aspects concerning the operators ability to produce adequate motor decision in telemanipulation.

The learning effect we obtained confirms that information about forces was adequate for this task. Interpretations have to be based both on : differences in time between the first and second part and on the absolute value of execution time at the biginning and at the end of the session. Under the control condition no learning effect occured. As suggested by our assumptions, observation of the rod does not give the operator a knowledge of the result of his motor action on the telemanipulator where as visual force feedback does. Under these two experimental conditions only the "visual square" presentation leads to a significant learning effect but operator performance was not better at the end of the session with this presentation as compared to the "vectorial" presentation. These results suggest that the "vectorial" presentation could be more appropriate to the cognitive representation of the forces exerted by the telemanipulator than the "visual square" type presentation. However, the lack of significant differences between vector and visual square presentations does not allowed us to conclude in favour of one of then.

The learning effect we obtained under only one experimental condition clearly indicates task completion time is varying with the type of the visual feedback. The concept of operative image (Ochanine, 1971) could give us a general frame for more research to conceive visual presentations of feedback information adapted to its human cognitive representation. The experiment reported was a first attempt to include cognitive studies to design the teleoperator's working station.

REFERENCES

1. BERTSCHE, W.R., COGAN, K.P., PESCH, A.J. and WINGET, C.L. Operator performance in undersea manipulator systems : studies of control performance with visual force feedback. Final report, unpublished manuscript, 1977.

2. DUPOURQUE, V., GAILLARD, J.P. et QUETIN, N. Organes de commande manuelle : Etude bibliographique. "Préévaluation d'une poignée à six degrés de liberté". Rapport final non publié. Convention n° 80/CNES/816, 1981.

3. FALZON, P. Structure des dispositifs de présentation d'information et compatibilité avec les représentations mentales. Communication au 18e Congrès de la Société d'Ergonomie de Langue Française, Paris, 1982.

4. GAILLARD, J.P. The effects of electrotactile and visual force feedback on learning and performance in a telemanipulation task. In proceedings of the first european annual conference on human decision making and manual control. Delft University of Technology, 1981a.

5. GAILLARD, J.P. The persistance of a visual dominance effect in a telemanipulation task : a comparison between visual and electrotactile feedback. In proceeding of the 17th annual conference on manual control. University of California, Los Angeles, California, 1981b.

6. OCHANINE, D.A. L'homme dans les systèmes automatisés, Paris, Dunod, 1971.

7. STANTON, L. and COOK, C.W. Multiple man-machine interfaces. In proceeding of the 17th annual conference on manual control. University of California, Los Angeles, California, 1981.

8. VERTUT, J., CHARLES, J., COIFFET, P., and PETIT, M. Advances of the new MA-23 force reflecting manipulator system. In proceeding of the Ro Man Sy - 76 Symposium, Warsaw, Poland, 1976.

COMPENSATION FOR OBJECT MOVEMENT IN REMOTE MANIPULATION

Kazuo Tani
Mechanical Engineering Laboratory
Namiki, Sakura-mura, Ibaraki 305, Japan

Thomas B. Sheridan
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, U.S.A.

ABSTRACT

This paper proposes a technique for master/slave control with compensation for object movement to facilitate the human operator's manipulation of a moving object. The technique was implemented by means of digital computer control, and evaluated by experimentally comparing the performance of the operator with and without such compensation.

An experimental system was built which consists of a master/slave manipulator, a moving table for the moving object, and a computer which controls both the manipulator and the table. A computer control scheme for master/slave and compensation for object movement was developed in consideration of the kinematics and dynamics of the manipulator. The computation time in this scheme proved practical and permitted a system sampling frequency of 20 Hz.

Experiments were carried out with human operators performing manipulation tasks using computer aided master/slave control. Their performance was compared in three situations: no object movement, compensation for object movement, and no compensation. The comparison of the compensation and no compensation situations showed that the compensation reduced the operation time by 26-41 % in the peg moving task and increaced the accuracy by two and a half times in the rectangle tracing task. In the valve turning task, however, no significant improvement was observed.

Thus, it was concluded that the compensation for object movement can improve the performance of the human operator significantly in certain kinds of tasks.

The research described in this paper was carried out at the Man-Machine Systems Laboratory, Department of Mechanical Engineering, Massachusetts Institute of Technology under Grant 04-7-158-44079, MIT Sea Grant Program, Office of Sea Grant, National Oceanographic and Atmospheric Administration, Department of Commerce and Contract N00014-77-C-0256, Work Unit number NR196-152, Office of Naval Research, Engineering Psychology Programs.

1. INTRODUCTION

With the advancement of undersea development, the necessity to perform various manipulation tasks undersea is increasing and such work has to be done in deeper places. Work by human divers is suffering from higher costs and greater danger, while work by manipulators mounted on submersibles is gaining economy and safety advantages[1]. Undersea work includes construction tasks and maintenance and repair tasks, and requires many unrepeated actions. Therefore, play-back robots very often used in manufacturing factories cannot be applied, and manipulators operated by humans are suitable instead. Operation of a manipulator, however, is a tough task for a human operator. Usually he operates the manipulator with switches and joysticks in rate control or with a master in master/slave control. In these methods, the load on the human operator is so great that the operator gets tired in a short time, reducing his performance.

A special problem in the undersea application of a manipulator is the relative movement between the manipulator base and the object to be handled by the manipulator. The submersible equipped with a manipulator may fix itself on the seabed, but the object may drift because of the water flow; or, inversely, the object may be fixed and the submersible with a manipulator may move about; or both may move. In any case there is a relative movement between the manipulator base and the object. Consider the case where the operator uses a master/slave manipulator. He has to make a motion which is the superposition of the object pursuit motion and the motion of the proper manipulation done on the object. It is easily imagined that the operation is much harder than if there were no such relative movement.

If the movement of the object relative to the manipulator base can be measured, compensation for object movement is possible which eliminates the object pursuit component of the motion and allows the operator to perform the same motion as if the object were stationary.

In this study, we proposed a concept of master/slave control with compensation for object movement, constructed an experimental system according to the concept, and evaluated the compensation for object movement by carrying out experiments with human operators and comparing their performance in manipulation tasks.

Development of the technique of master/slave control with compensation for object movement, the experimental setup, and the control software will be described below, and the method and the results of the performance comparison experiments by subjects will be discussed.

2. MASTER/SLAVE CONTROL WITH COMPENSATION FOR OBJECT MOVEMENT

In the usual force-reflecting master/slave control, a servo loop is constructed where the master position is the command to the slave position control and the slave position is the command to the master position control. In master/slave control with compensation for object movement, on the other hand, both the master position and the slave position relative to the moving object are controlled

in a master/slave way. That is, the object position is added to the master position to give the command to the slave position control, and is subtracted from the slave position to give the command to the master position control. Here, the addition and the subtraction are done on the Cartesian coordinates (x), while the drive of the manipulator is done on the joint angles coordinates (θ). Therefore, transformations between coordinates: $\theta \rightarrow x$ and $x \rightarrow \theta$ are needed, which would be unnecessary in the usual simple master/slave control where only θ coordinates were used. Fig. 1 shows the conceptual block diagram. x_m and θ_m in this figure represent the present master position in the respective coordinates, x_{mr} and θ_{mr} present the command to the master position control, x_s, θ_s, x_{sr}, and θ_{sr} represent corresponding values for the slave position. x_t represents the present position of the object.

In the realization of such a system, computer control is mandatory, for the computational processing such as coordinate transformations can be performed easily. Stability, controllability, and object pursuit should be implemented carefully in consideration of the manipulator dynamics to make a good system. Details of these items will be discussed in Chapter 4.

3. EXPERIMENTAL SETUP

The main devices used in the experiment were a master/slave manipulator and a moving table which are shown in Fig. 2 and a computer controller.

The manipulator was Argonne National Laboratory E2, articulated, had six degrees of freedom besides grip, and is modelled as shown in Fig. 3 [2]. The master and the slave had the same construction, electrical and mechanical, except the shape of the gripping part or the hand. Each joint angle was driven by a DC servo motor, 10 W, 115 V, 60 Hz. A servo circuit was provided to measure each joint angle and drive it to follow the respective given command. The signal of the angle also was sent to the computer. The servos of the each pair of correponding master and slave angles could be switched into two modes: master/slave mode and computer control mode. In the master/slave mode, the computer was irrelevant and a master/slave control loop was constructed by connecting the corresponding master servo and slave servo. In the computer control mode, on the other hand, the master servo and the slave servo operated independently in response to the respective commands given by the computer.

The moving table for the object movement had a carriage which moved along three axes (x, y, z) driven by stepping motors. The speed and the direction along each axis were given by the computer. The maximum speed was 107 mm/s along x and y and 90 mm/s along z, and the strokes were 350, 400, and 50 mm along x, y, and z, respectively.

The computer PDP11/34 was used for the control. A data aquisition and distribution system ANALOGIC AN5400 was used for the computer to communicate with the experimental devices. The AN5400 had 32 channels of analog input, 16 channels of analog output, 32 bits of digital input, and 32 bits of digital output, and was connected to PDP11/34 through a universal interface module DR11-C.

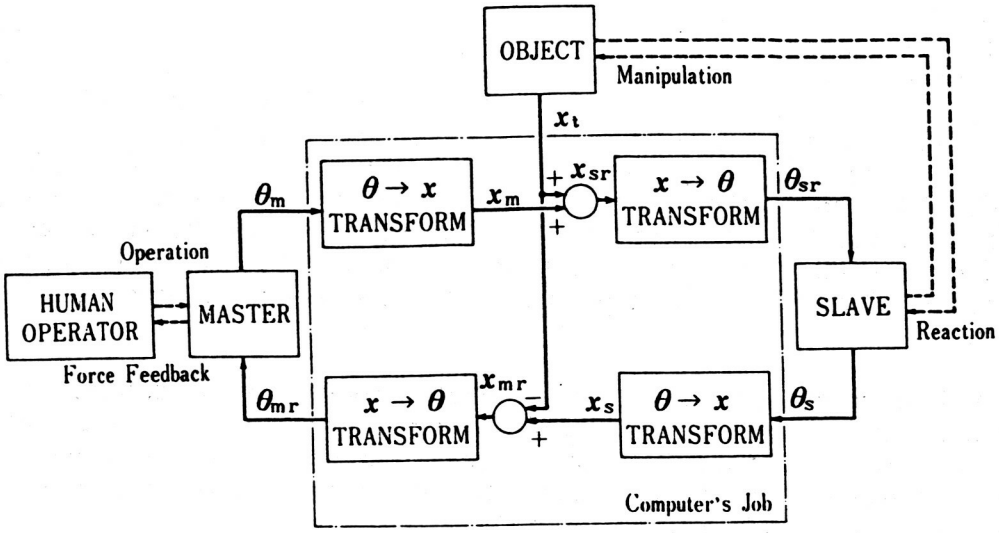


Fig. 1 Conceptual scheme of master/slave control with compensation for object movement.

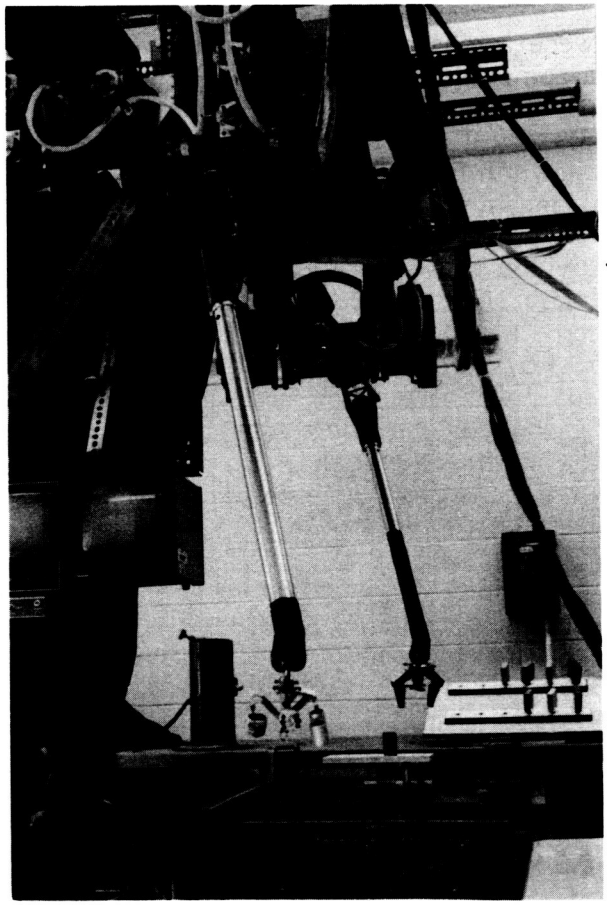


Fig. 2 Master/slave manipulator and moving table.

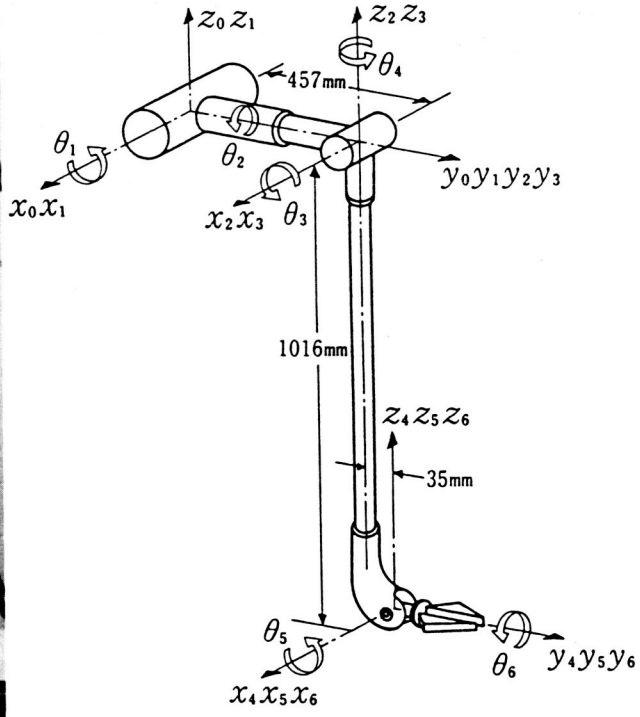


Fig. 3 Definitions of coordinate systems and rotation angles of the manipulator. Angles θ_k are the rotations of coordinate frame k. Angles are assumed zero as shown.

4. SYSTEM SOFTWARE

In implementing master/slave control with compensation for object movement done by computer control, the computer software is very dependent on the hardware conditions such as the structure and the dynamics of the manipulator. The consideration of the hardware used in the experiment, made in the course of the implementation, will now be discussed.

4.1. Kinematics of the manipulator

The system requires the coordinates transformations $\theta \rightarrow x$ and $x \rightarrow \theta$. Among them, the transformation $\theta \rightarrow x$ can be expressed as an explicit function:

$$x = f(\theta) \tag{1}$$

whose details will be discussed later. The transformation $x \rightarrow \theta$, on the other hand, generally cannot be expressed in an explicit function. For this reason we adopted an indirect method of the Resolved Motion Rate Control (RMRC) which was proposed by D. E. Whitney[3].

Denote x the position vector at the time t

$$x = (x_1, x_2, \cdots, x_n) \tag{2}$$

and θ the joint angles vector

$$\theta = (\theta_1, \theta_2, \cdots, \theta_m) \tag{3}$$

Using a Jacobian matrix

$$J(\theta) = \begin{bmatrix} \frac{\partial x_1}{\partial \theta_1} & \cdots\cdots & \frac{\partial x_1}{\partial \theta_m} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \frac{\partial x_n}{\partial \theta_1} & \cdots\cdots & \frac{\partial x_n}{\partial \theta_m} \end{bmatrix} \tag{4}$$

the following relation is obtained:

$$\dot{x} = J(\theta)\dot{\theta} \tag{5}$$

Eqn. (5) shows a linear relation between $\dot{x}$ and $\dot{\theta}$, and $\dot{\theta}$ can be solved for a given $\dot{x}$ if $m = n$ and $J(\theta)^{-1}$ exists:

$$\dot{\theta} = J(\theta)^{-1}\dot{x} \tag{6}$$

If it is assumed that for a small time interval Δt

$$\Delta\theta = \dot{\theta}\Delta t, \quad \Delta x = \dot{x}\Delta t \tag{7, 8}$$

$\Delta\theta$ is obtained by the following expression:

$$\Delta\theta = J(\theta)^{-1}\Delta x \tag{9}$$

The application of the RMRC method requires formulation of $f(\theta)$ in Eqn. (1), introduction of a new parameter to eliminate the system redundancy, formulation of the solution $J(\theta)^{-1}$ and $\Delta\theta$, and addition of a degree of freedom for the hand orientation control. Details of these will be discussed below.

As shown in Fig. 3, we call the coodinate system fixed to the manipulator base Frame 0, and the coordinates in it x_0. Likewise, we call the coordinate system fixed to the k-th joint angle Frame k, and the coordinates in it x_k. In particular we call Frame 4 and Frame 6 the wrist frame and the hand frame, respectively.

Table 1 Transformation matrices A_0^4 and A_0^6

$$A_0^4=\begin{bmatrix} a_{11}^4 & a_{12}^4 & a_{13}^4 & x \\ a_{21}^4 & a_{22}^4 & a_{23}^4 & y \\ a_{31}^4 & a_{32}^4 & a_{33}^4 & z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

a_{11}^4=S3S2S4+C2C4
a_{12}^4=S3S2C4−C2S4
a_{13}^4=C3S2
a_{21}^4=C1C3S4+S1S2C4−S1S3C2S4
a_{22}^4=C1C3C4S−1S2S4−S1S3C2C4
a_{23}^4=−(S1C3C2+C1S3)
a_{31}^4=S1C3S4−C1S2C4+C1S3C2S4
a_{32}^4=S1C3C4+C1S2S4+C1S3C2C4
a_{33}^4=C1C3C2−S1S3
x=35(S3S2C4−C2S4)−1016C3S2
y=35(C1C3C4−S1S2S4−S1S3C2C4)
+1016(S1C3C2+C1S3)+457C1
z=35(S1C3C4+C1S2S4+C1S3C2C4)
−1016(C1C3C2−S1S3)+457S1
$x^4=(a_{11}^4, a_{21}^4, a_{31}^4)$
$y^4=(a_{12}^4, a_{22}^4, a_{32}^4)$
$z^4=(a_{13}^4, a_{23}^4, a_{33}^4)$

$$A_0^6=\begin{bmatrix} a_{11}^6 & a_{12}^6 & a_{13}^6 & x \\ a_{21}^6 & a_{22}^6 & a_{23}^6 & y \\ a_{31}^6 & a_{32}^6 & a_{33}^6 & z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

a_{11}^6=(S3S2S4+C2C4)C6−[C3S2C5−(S3S2C4−C2S4)S5]S6
a_{12}^6=(S3S2C4−C2S4)C5+C3S2S5
a_{13}^6=[C3S2C5−(S3S2C4−C2S4)S5]C6+(S3S2S4+C2C4)S6
a_{21}^6=(C1C3S4+S1S2C4−S1S3C2S4)C6
+[(S1C3C2+C1S3)C5+(C1C3C4−S1S2S4−S1S3C2C4)S5]S6
a_{22}^6=(C1C3C4−S1S2S4−S1S3C2C4)C5−(S1C3C2+C1S3)S5
a_{23}^6=−[(S1C3C2+C1S3)C5+(C1C3C4−S1S2S4−S1S3C2C4)S5]C6
+(C1C3S4+S1S2C4−S1S3C2C4)S6
a_{31}^6=(S1C3S4−C1S2C4+C1S3C2S4)C6
−[(C1C3C2−S1S3)C5−(S1C3C4+C1S2S4+C1S3C2C4)S5]S6
a_{32}^6=(S1C3C4+C1S2S4+C1S3C2C4)C5+(C1C3C2−S1S3)S5
a_{33}^6=[(C1C3C2−S1S3)C5−(S1C3C4+C1S2S4+C1S3C2C4)S5]C6
+(S1C3S4−C1S2C4+C1S3C2S4)S6
$x^6=(a_{11}^6, a_{21}^6, a_{31}^6)$
$y^6=(a_{12}^6, a_{22}^6, a_{32}^6)$
$z^6=(a_{13}^6, a_{23}^6, a_{33}^6)$

The transformation from x_4 and x_6 to x_0 is expressed using transformation matrices $A_0^4(\theta)$ and $A_0^6(\theta)$ as follows:

$$x_0=A_0^4 x_4, \quad x_0=A_0^6 x_6 \tag{10, 11}$$

respectively. Table 6 shows the matrices A_0^4 and A_0^6. Here $\sin\theta_1$, $\cos\theta_1$, $\sin\theta_2$,... are abbreviated as S1, C1, S2,..., respectively. Vectors x^4, y^4, z^4, x^6, y^6, and z^6 are defined in this table. x^4, for example, is the unit vector along the x axis on Frame 4 as seen from Frame 0.

Now consider obtaining joint angles $\theta_1 \sim \theta_6$ which place the hand, i.e. the origin of Frame 6, in a certain position of Frame 0 (x_d, y_d, z_d). Transformation matrices A_0^4 and A_0^6 show that θ_5 and θ_6 are irrelevant. Since the system, to solve $\theta_1 \sim \theta_4$ for a given (x_d, y_d, z_d), is a redundant one, we consider introducing a new parameter. Suppose, keeping the y axis vector of the hand frame y^6 horizontal, we express y^6 as follows:

$$y^6=(c, 1, 0)(1+c^2)^{-1/2} \tag{12}$$

where c is an arbitrary value. Since the vector x^4 is perpendicular to y^6 regardless of θ_5 and θ_6, the inner product of x^4 and y^6 is zero:

$$(a_{11}^4, a_{21}^4, a_{31}^4) \cdot (c, 1, 0)=0 \tag{13}$$

which is rewritten

$$ca_{11}^4+a_{21}^4=0 \tag{14}$$

Thus we introduce a new parameter $p(\theta)$

$$p=ca_{11}^4+a_{21}^4 \tag{15}$$

This parameter should be controlled to be zero to meet the supposition. The RMRC method can be applied now if we rewrite

$$x=(x, y, z, p) \tag{16}$$
$$\theta=(\theta_1, \theta_2, \theta_3, \theta_4) \tag{17}$$

In this method, given the destination position (x_d, y_d, z_d, c), first evaluate $\boldsymbol{x}$ of Eqn. (16) using the present value of the joint angles $\boldsymbol{\theta}$, and then evaluate $J(\boldsymbol{\theta})^{-1}$. In this case we adopted a method to express $J(\boldsymbol{\theta})^{-1}$ in advance as an explicit function of $\boldsymbol{\theta}$ using the method of Banachiewicz[4], and then to substitute $\boldsymbol{\theta}$ with the present value. Next evaluate the value

$$\Delta \boldsymbol{x}=(\Delta x, \Delta y, \Delta z, \Delta p)=(x_d-x, y_d-y, z_d-z, -p) \tag{18}$$

and $\Delta \boldsymbol{\theta}$ from Eqn. (9). In this case Δt is the sampling period.

In the construction of the program to calculate the values of $\boldsymbol{x}$ and $J(\boldsymbol{\theta})^{-1}$, economy of the calculation is possible because there are common terms among expressions. Thus we reduced the calculation time of the RMRC program to 8.8 ms. This program should run twice during one sampling period, once for the master and once for the slave. The calculation time was short enough for the Δt = 50 ms.

The requirement to keep the vector $\boldsymbol{y}^6$ horizontal:

$$a_{32}{}^6=0 \tag{19}$$

gives θ_5. Define the θ_6 value as θ_{6h}, given by the temporary condition to keep the hand vector $\boldsymbol{x}^6$ horizontal:

$$a_{31}{}^6=0 \tag{20}$$

Thus we can obtain $\theta_1 \sim \theta_5$, θ_{6h} with the condition to keep the hand vectors $\boldsymbol{x}^6$ and $\boldsymbol{y}^6$ horizontal, which is a system with four degrees of freedom. If we add the freedom of θ_6, that is, add an arbitrary angle θ_{6d} to the base angle θ_{6h}, then the system has five degrees of freedom. Thus the master/slave control system shown in Fig. 1, by the application of the above results, now has five degrees of freedom, that is, x_d, y_d, z_d, c, and θ_{6d}. c and θ_{6d} are irrelevant from the compensation. The grip was not controlled in the computer control mode but was left in the master/slave mode. Thus the system has one degree of freedom less than the manipulator hardware.

4.2. Stability and controllability of the manipulator

In the realization of the system shown in Fig. 1, problems arising from the manipulator dynamics need to be solved. Here the measures taken for the stability and controllability of the system will be discussed.

The use of a digital computer required discrete time sampling for program control. The sampling period was set at 50 ms, for a period of 100 ms did not produce a sufficiently smooth movement of the manipulator.

In the computer controlled master/slave, simply using the master position as the command to the slave position control and the slave position for the command to the master position control caused an undamped oscillation of the manipulator. The following control scheme was used to prevent this vibration:

$$\boldsymbol{x}_{sr}=a\boldsymbol{x}_m+(1-a)\boldsymbol{x}_s \tag{21}$$

$$\boldsymbol{x}_{mr}=a\boldsymbol{x}_s+(1-a)\boldsymbol{x}_m \tag{22}$$

where a $(0 \leqq a \leqq 1)$ is called the master/slave coefficient, $\boldsymbol{x}_m$ and $\boldsymbol{x}_s$ are positions at the sampling time t, and $\boldsymbol{x}_{mr}$ and $\boldsymbol{x}_{sr}$ are commands given by the computer. The experiment showed that a value over 0.7 for a caused vibration of the manipulator.

The above control scheme still had a problem in that the master was so stiff that the operator found difficulty moving it. Use of the positive feedback of the velocity terms suggested the following control scheme:

$$x_{sr}=ax_m+(1-a)x_s+b(x_s-x_{s2}) \quad (23)$$

$$x_{mr}=ax_s+(1-a)x_m+b(x_m-x_{m2}) \quad (24)$$

where b is called the positive feedback coefficient, x_{m2} and x_{s2} are the positions of the master and the slave two periods in the past. These values were adopted, for the use of the values one period in the past x_{m1} and x_{s1} caused vibration. A value over 0.5 for b caused a self-initiated movement of the manipulator.

The above discussions are represented schematically by the block diagram of Fig. 4 with the part enclosed by the dashed line excluded, where G_m and G_s are the transfer functions of the master and slave servos, D_1 is the sampling, and D_2 the delay of two periods. In the experiments to be described below, values a = 0.7, b = 0.5 were used. These coefficients, a and b, are meant to compensate for the decline of the stabilibity and controllability caused by the delay from the sampling and computation inevitable for a computer control system. Details of the coefficients, such as their ranges for the system to be stable, will not be discussed in this paper.

4.3. Control of compensation for object movement

Suppose that, in master/slave control with compensation for object movement, the master stays still. Then the slave should make the same movement as the object. Fig. 5 is the block diagram of the computer control that makes the slave do the same movement as the object. Here the computer determines $G_s^{-1}\boldsymbol{X}_t$ from the sampled data $D_1\boldsymbol{X}_t$ and outputs it to the slave. G_sD_1 can be regarded as a delay if the dominant frequencies of x_t are small enough as compared with the fundamental frequencies of G_s and D_1. Then $G_s^{-1}D_1^{-1}$ is regarded as a lead. In other words, the computer assumes the object position a certain time in future x_{te} from the sampled data x_t and outputs it to the slave. The lead time was measured from the experiment. It varied as a function of the axes and was 100 ms along x and y axes and 150 ms along z, if expressed as multiples of the sampling period for simplicity.

Embedded with the object movement pursuit control shown in Fig. 5 in addition to the master/slave control, the overall system is schematized by Fig. 4. Note that x_t theoretically does not have any influence on the master.

In the experiment the object was moved in a predetermined manner so that x_{te} was easily determined. In practical applications, however, preview control will be needed, which predicts the future object position from the measured object position [5].

The control scheme of the computer aided master/slave control with compensation for object movement was finally as follows:

$$x_{sr}=ax_m+(1-a)(x_s-x_t)+b[(x_s-x_t)-(x_{s2}-x_{t2})]+x_{te} \quad (25)$$

$$x_{mr}=a(x_s-x_t)+(1-a)x_m+b(x_m-x_{m2}) \quad (26)$$

where x_{t2} is the object position two periods in the past.

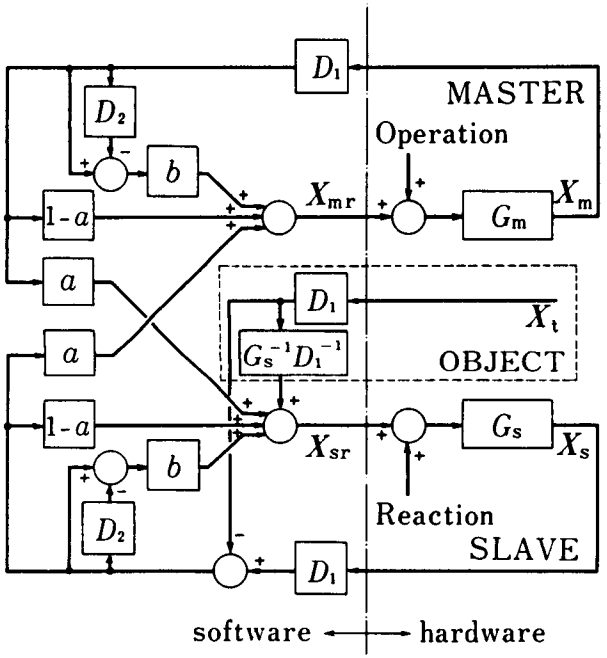


Fig. 4 Block diagram of computer aided master/slave control with compensation for object movement.

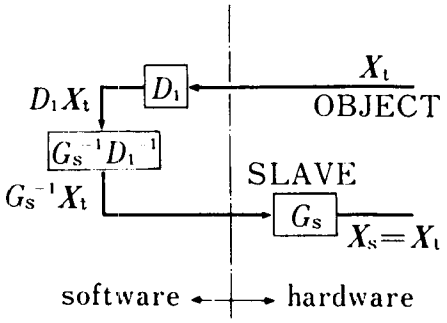


Fig. 5 Block diagram of computer controlled object pursuit motion by the slave.

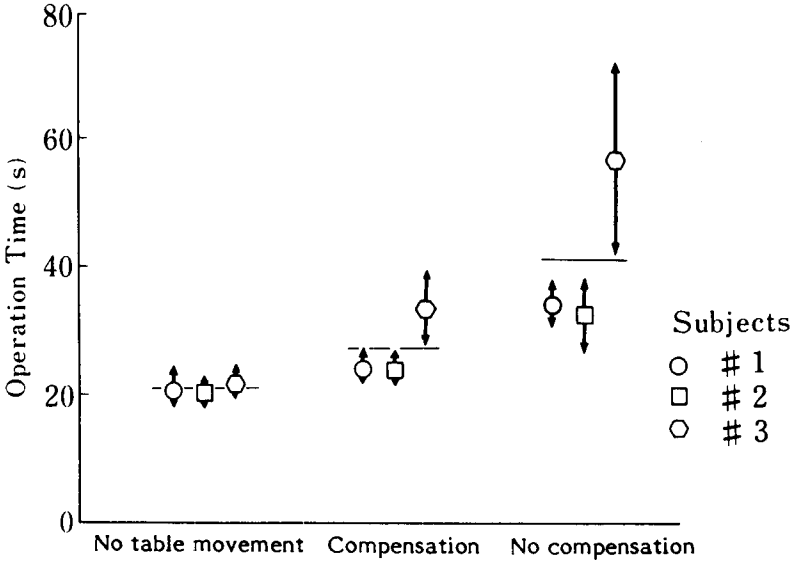


Fig. 6 Operation time for the peg moving task.

5. EXPERIMENTS ON THE PERFORMANCE OF OPERATORS

5.1. Experiments

Experiments were carried out with the experimental system for computer aided master/slave control with compensation for object movement to evaluate how much the compensation for the object movement improves the manipulation performance of the operator. This performance was compared with three situations: (1) no relative movement between the manipulator base and the object, (2) relative movement but no compensation, and (3) relative movement and compensation. The movement of the table, which is the relative movement between the object and the manipulator base was given by the following equations:

$$x = 36 \sin(2\pi t/5.0) - 142 \sin(2\pi t/14.5) \quad (27)$$

$$y = 36 \sin(2\pi t/6.5) - 142 \sin(2\pi t/18.5) \quad (28)$$

$$z = 10 \cos(2\pi t/5.0) - 10 \cos(2\pi t/6.5) \quad (29)$$

where position x, y, or z is expressed in millimeters and time t in seconds.

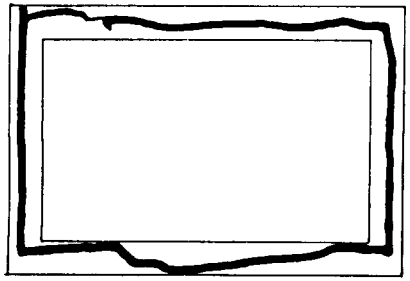
Three kinds of tasks were chosen for the operator to do: peg moving, rectangle tracing, and valve turning. The peg moving task was chosen as an example of a point-to-point task, the rectangle tracing task an example of a continuous path task, and the valve turning an example of the partial constraint to the task motion from the mechanism of the moving object. The operator did the tasks as he directly watched the slave and the object. This is equivalent to no compensation of the field of view for the object movement. Each subject operator performed five runs for each task and each situation.

In the peg moving task, a peg of 10 mm in diameter and 32 mm in length with a plate of 66 x 61 x 20 mm to be gripped by the slave hand was used. A board with eight holes of 11 mm in diameter arranged in a 2 x 4 array with 76 mm intervals was mounted on the moving table. Seven of the holes were champhered by 2 mm and used for the experiment. The subject was told to insert the peg into the first hole and move it to the next hole and so on around the board and back to the first hole. The time for the complete cycle was measured between the first insertion and the second insertion into the same first hole.

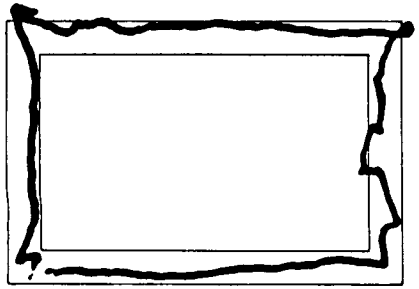
In the rectangle tracing task, a piece of paper with two parallel and concentric rectangles of 102 x 153 mm and 76 x 127 mm drawn on it was placed on the moving table and the middle of the space between the two was regarded as the reference path. The subject was told to trace the reference path with a marking pen gripped by the slave. Eight largest deviation values were chosen from the error, which is the distance between the trace and the reference path, and their average value was calculated. The following method was used to select the eight values.

Take the maximum value of the error. Discard all the error within 25 mm from the point of the value taken along the reference path. Take the maximum value from the remaining error. Repeat this operation until eight values are taken (see Fig. 7 (c)).

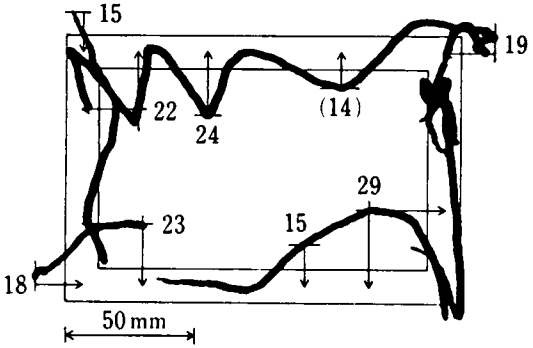
In the valve turning task, a valve with a handle of about 70 mm in diameter was placed on the table. The time during which the subject turned the valve 720 degrees was measured. The subject gripped the valve with the slave hand and turned the valve about 180 degrees, then released it and turned back the hand to grip the valve again.



(a) No table movement (e = 4.0)



(b) Compensation (e = 7.0)



(c) No compensation (e = 20.6)

Fig. 7 Examples of rectangle tracing.

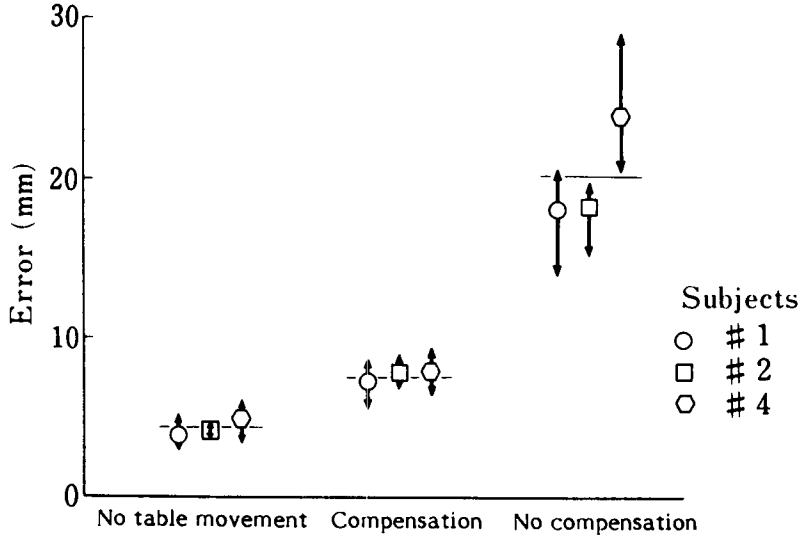


Fig. 8 Comparative errors in rectangle tracing.

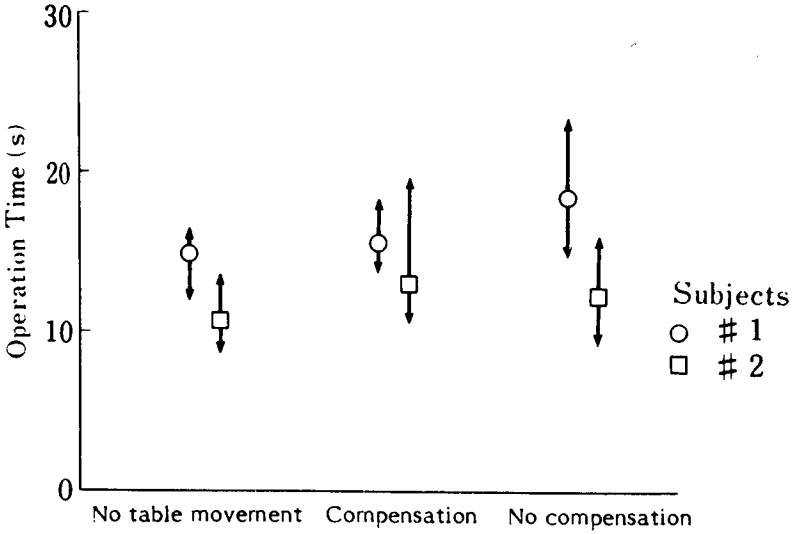


Fig. 9 Operation time for the valve turning task.

5.2. Results of the experiments

Fig. 6 shows the results of the peg moving task. The marks show the average task time and the arrows show its variation. The task time increased by 18-51 % (29 % average) with compensation as compared with no relative movement, but decreased by 26-41 % (32 % average) as compared with no compensation.

Fig. 7 shows examples of the rectangle tracing. Fig. 8 shows the comparison of the errors. The average error in the compensation situation was 162-189 % (178 % average) of the error in the no movement situation, but 33-42 % (38 % average) of the error in the no compensation situation.

Fig. 9 shows the results of the valve turning task. No significant performance difference was observed between different situations in this experiment.

5.3. Discussion

The effect of compensation for object movement on operator performance is most apparent in the rectangle tracing task among the three kinds of tasks. It is also apparent in the peg moving task. We think that these two kinds of tasks are mainly positioning tasks, one continuous path and the other point-to-point, and therefore both benefited from the compensation.

On the other hand, the improvement by the compensation in the valve turning task is hardly significant. We suppose the reason as follows: The difficulty in the valve turning task lies not in the positioning but rather in turning the master hand about the $\boldsymbol{y}^6$ axis while maintaining the hand vector $\boldsymbol{y}^6$ to coincide the valve turning axis. As the turning motion has nothing to do with the compensation, the benefit from the compensation on the performance is small. Sharing the task, in which the operator does the positioning of the manipulator and the computer does the turning automatically, as is shown in the literature [6], should lead to a different result.

The range these experiments have covered is by no means wide, but they have shown that operator manipulation performance can be improved by automatic compensation for object movement for certain kinds of tasks.

6. CONCLUSION

The experiments we conducted showed that compensation for object movement improves the performance of the human operator when it is incorporated in master/slave control of a manipulator. Toward practical application of master/slave control with compensation for object movement, the following topics should be investigated.

(a) Measurement of object movement.
(b) Display of the working environment to the operator. For example, application of compensation for object movement to the operator's field of view should be considered.
(c) Application to supervisory control systems. The concept of supervisory control was proposed [7], where the operator occasionally gives orders in a high level language to a semi-autonomous computer, which, with its capability of sensing the environment, directly and continuously controls the manipulator in an autonomous way for a relatively short period of time. Compensation for object movement has potential to contribute to the improvement of performance as a subsystem in a supervisory control system.

ACKNOWLEDGEMENT

We thank Dr. Dana Yoerger and other members of the Man-Machine Systems Laboratory for their helpful advice and suggestions as well as for their cooperation in preparing and carrying out the experiments.

REFERENCES

[1] T. B. Sheridan and W. L. Verplank: Human and Computer Control of Undersea Teleoperators, Man-Machine Systems Laboratory Report, Massachusetts Institute of Technology (1978).

[2] T. L. Brooks and T. B. Sheridan: SUPERMAN: A System for Supervisory Manipulation and the Study of Human/Computer Interactions, MIT Sea Grant Program, MITSG-79-20, Massachusetts Institite of Technology (1979).

[3] D. E. Whitney: Resolved Motion Rate Control of Manipulators and Human Prostheses, IEEE Trans. on Man-Machine Systems, MMS-10-2, 47/53 (1969).

[4] S. H. Crandall: Engineering Analysis, 26/31, McGraw-Hill (1956).

[5] T. B. Sheridan and W. R. Ferrell: Man-Machine Systems, M.I.T. Press (1974).

[6] T. L. Brooks and T. B. Sheridan: Experimental Evaluation of the Concept of Supervisory Manipulation, 16th Annual Conference on Manual Control, M.I.T., Cambridge, MA (1980).

[7] W. R. Ferrell and T. B. Sheridan: Supervisory Control of Remote Manipulation, IEEE Spectrum, 4-105 81/88 (1969).

A SENSORY SAMPLING MODEL OF THE HUMAN OPERATOR OF TELEOPERATED DEVICES

Kevin Corker
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California 91109

ABSTRACT

The human operator (HO) as a control element of teleoperated devices, specifically teleoperated manipulators, has a multiplicity of roles in such systems. The range of operator action is from that of a direct manual controller of the remoted manipulator to that of a supervisor and monitor of complex automated or computer augmented task performance. Models of that operator must take into account the diversity of control tasks ascribed the human as well as the dynamic value of the information fedback to the operator to aid control. Experiments were conducted to characterise operator information seeking and control in bilateral teleoperation involving manual control and computer augmented information display. Preliminary results indicate:
1) that current models of neuromuscular control techniques describing the kinematics and dynamics of limb control are applicable to manual control of remote teleoperators, and
2) sensory sampling in the information seeking aspect of HO control follows task dependent utilities through which robust strategies are developed.

INTRODUCTION

The multiple roles of the HO in control of teleoperated manipulators impose constraints on attempts to model that operator's behavior in information seeking and control. On the one hand, the operator serves as a supervisory monitor of information and high level planner in control. On the other hand, because of limitations in algorithmic control, environmental and task uncertainties, and/or restrictions on sensor data, the operator provides direct manual control of the teleoperator.

Ferrell and Sheridan (1967) proposed a paradigm for consideration of this operational diversity for which they coined the term Supervisory Control. This conceptually complete structure assigns to the operator both manual and executive control functions. The resolution of the role of the HO into distinct and hierarchically arranged task components continues to provide a useful paradigmatic framework for the descriptive models of the human controller.

Following from this division of tasks, many researchers have sought to characterise the operators function in either the supervisory or manual control portions of the model. In a wide

ranging study Brooks (1979) used the supervisory framework to structure research on the interaction of a semiautonomous manipulator and the human controller of that device. Tulga and Sheridan (1980) investigate and model operator behavior in the face of a muti-task, multi-decision level control. White (1981) uses the paradigm to model large scale process control incorporating a control theoretic description of the operator. Other control theoretic approaches have been extended to include some decision theoretic models of information processing by the operator in the supervisory mode (Pattipatti et al., 1980). At the level of manual control in the supervisory paradigm, suggestions have been made for the explicit inclusion of optimal control formulations for modelling of neuromuscular control and skill development (Pew and Baron, 1978).

The work cited above serves as an example of the ample framework for modelling the process of operator control of a teleoperated system (TOS). However, the approaches above do not directly address several critical concerns for the behavior of that operator in the actual context of manipulation. First, the human operator serves interactively at the two extremes of control (supervisory and manual) and at the full range of intermediate levels of information processing and task assignnment. Descriptive unification of the extremes of this functional range can be achieved by consideration of information seeking in control. The informational components for each level of the control task are analyzed to abstract information about their interrelations and then synthesised into a hypotheses for appropriate control. These hypotheses about meaning of fedback information are used to direct the search for new information. In this sense, the control is both information driving and information driven at all levels of the task. This is particularly true in the case of TOS control because the teleoperator interacts with the environment; therefore, manipulation is inherently affective and effective. Second, the informational constraints on control processes based upon the value of information in control planning, or in specification of the controlled parameters for action, are not systemaically evaluated (cf. Howard, 1967, Sheridan, 1970).

It is the purpose of the research presented here to begin a preliminary investigation of the content of the information feedback and control interaction of the HO as a control element in a bilateral teleoperator system. The experiments reported assume a supervisory paradigm and concentrate on the actvity associated with direct manual control of the teleoperated manipulator. Future work will include refinement of the requirements for direct control, and be expanded to include the human in the supervisory control mode.

EQUIPMENT

The research was conducted at the Advanced Teleoperation Laboratory of the Jet Propulsion Laboratory, California Institute of Technology. The Laboratory has, over the past several years, developed an integrated control and display test bed for the investigation of teleoperator control techniques.

The primary equipment used in these experiments was as follows: 1) a universal force reflecting hand controller (FRHC), 2) a control computer to impliment the required kinematic and dynamic transforms, 3) an integrated sensor and display system to provide the operator with the feedback necessary for sensor based control, and 4) a task board for timing and accuracy measurement in generic manipulation tasks. A brief description of the equipment is provided as it impacts the experimental methodology. For more detailed description the reader is referred to Bejczy and Salisbury, 1980, Handlykken and Turner, 1980, or Bejczy 1980.

Controller

The universal FRHC provides a generalized bilateral force-reflecting control of teleoperated manipulators. In a departure from the standard practice of master/slave control systems requiring kinematic and dynamic replication between the master and the slave, the FRHC control function is implemented through a hand controller which is dissimilar to the particular slave arm both dynamically and kinematically (see Figures 1 and 2). The

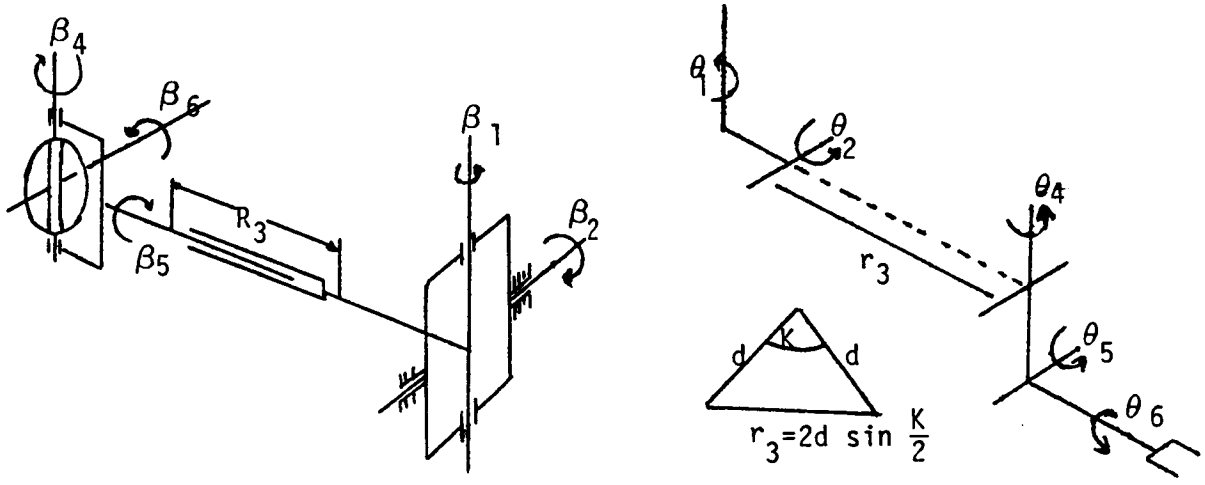


Figure 1. Hand Controller and Slave Arm Kinematics.

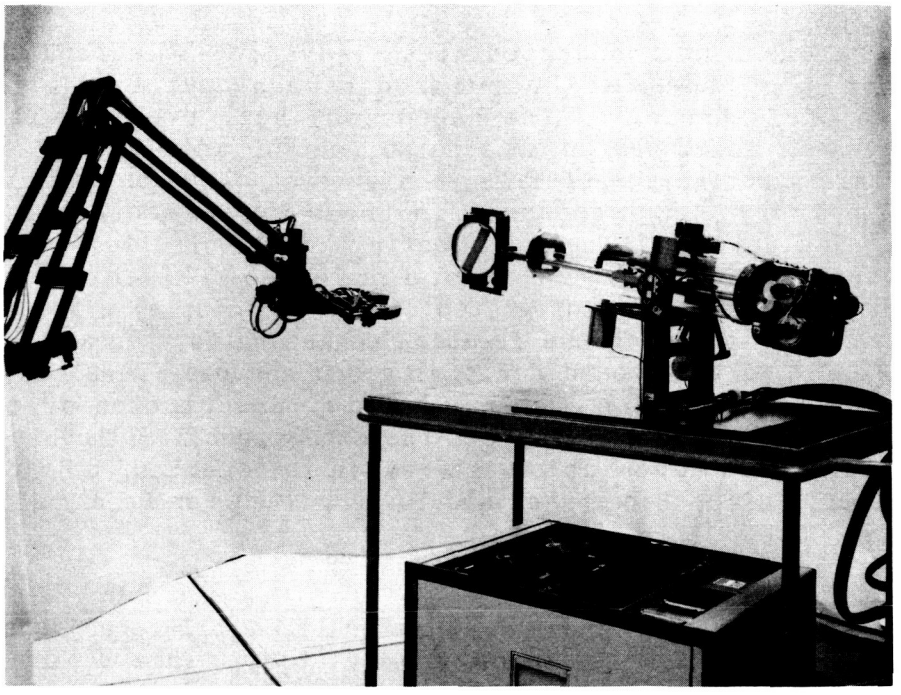


Figure 2. Universal Force Reflecting Hand Controller and NASA/CURV Teleoperator.

hand controller is a six degree of freedom (DOF) isotonic joystick which can be backdriven by commands from the control computer. The NASA/CURV teleoperator is a six DOF manipulator equiped with a force-torque sensor at the wrist link of the end effector, as well as proximity and touch sensors mounted on the gripper portion of the end effector (Bejczy, 1980).

Real time control algorithms resolve hand controller position commands (in the 30 cm. cubic range of motion of the FRHC) into appropriate manipulator joint motions to position the manipulator end point in its work space. Control algorithms also provide postion and velocity error backdrive commands to the hand controller linking the operator's hand position to the the end effector position of the manipulator in its workspace. Additionally, the control algorithms resolve forces and torques sensed through the force-torque sensor at the base of the hand and backdrive the FRHC in proportion to the sensed force. This provides the operator with a sense of feel in relation to the forces and torques impinging on the end effector.

The mechanical design of the hand controller is directed toward providing a transparent control device to the human operator. The system operates through a series of low friction, low backlash pulleys and gimbals. It is counter-balanced through a weighted idler mechanism which serves to make the hand grip neutral with respect to gravity, and maintains tension on the hand gimbal drive cables. The back drive motors and potentiometers, located at the opposite end of the translation axis from the hand gimbal serve to balance the system and provide low effective inertia at the hand grip (Bejczy and Salisbury, 1980).

The manipulator can also be decoupled from the FRHC controller and positioned using potentiometer joint control to position the end effector independently of the hand controller position.

Feedback

The kinesthetic feedback, providing the following: 1) proportional position error feedback, 2) proportional velocity error feedback, and 3)force-torque backdrive from the hand controller, can be varied in its effect through software manipulation of the gains associated with those feedback signals. This manipulation of feedback gain can be made for each axis of motion, independently.

The JPL Advanced Teleoperator Lab has integrated several other sensor based feedback mechanisms for operator control into a universal control station.

Visual feedback is provided by the following: 1) direct visual access from the control station to the work site, 2)through COHU television cameras mounted to provide mono TV, and 3) stereo TV provided through interlacing the input video from two cameras, mounted at approximately interpupillary distace, to a single CRT. The user then wears piezoelectric ceramic lead lantham zirconate titinate (PZLT) glasses, which shutter in snych with the TV interlace to provide the output of the two cameras alternatively to the right and left eye of the observer. The binocularly disparate image is subjectively fused to provide stable stereo picture.

Symbolic display is provided for proximity and touch information in the form of a CRT displayed schematic representation of the end effector and gripper surface, which is responsive to changing sensor conditions. Finally, the force-torque information from the force-torque sensor is displayed in a bar graph and numerical output format. The three orthogonal forces and three orthogonal torques, sensed at the wrist, are displayed as a histogram emerging to the right or left of a central neutral position shown on a CRT (see Figure 3).

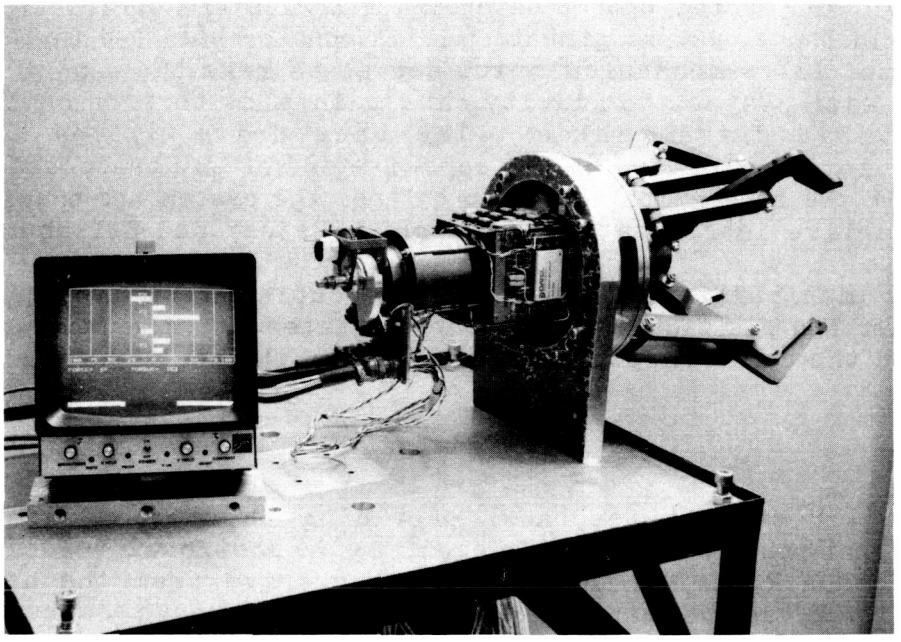


Figure 3. Force-Torque Display and Shuttle RMS Configured Gripper

EXPERIMENTS

Two experimental procedures have provided preliminary data charaterising the method of control and information seeking, through sensory sampling, of the HO in control of a bilateral manipulator system.

Experiment One

The purpose of the first experiment was to characterise the method of operator control in a bilateral manipulator with regards to the neuromuscular components of control. Recent research in the neuromotor control of position and trajectory formation have emphazised the importance of the dynamic characterisitics of the limb in the production of planar motion trajectories and achievement of final position (cf. Bizzi et al., 1982, Bizzi et al., 1978, Cooke, 1979, Hollerbach, 1980). The results of these investigations indicate that the achievement of a final position for the end point of the upper limb can be modelled as a balance of tensions between oppposing pairs of antagonist muscle groups. This balance can be considered to be defined by the equilibrium point of the intersecting length/tension curves of the appropriate muscle groups. Further, this equilibrial point is considerd to be a unitary control

variable in the specification of final position. Trajectory formation can be considerd to be accomplished by a sliding equilibrium point between inital and final position (Bizzi et al., 1982). An experiment was performed, using the equipment desibed above, to consider the impact of systematically varying the hand controller feedback dynamics on controller positioning in cartesian space.

Procedure

The manipulator was positioned in a central point in its workspace using the potentiometer control. The FRHC control was decoupled from the manipulator. The subjects (Ss) were blindfolded and asked to choose a target position to which to move the FRHC in its 30 cm cubic workspace. Despite the decoupling of control , kinesthetic feedback (resultant from position and velocity mismatch between the controller and the manipulator) was still available to the Ss. Therefore, the target position chosen by the Ss had associated with it a characteristic force at the hand controller. The accurate attainment of the target position was trained to an asymptotic criteria for each subject. When the Ss could move to the target position from a standard starting position to criteria, two manipulations of the kinesthetic feedback were introduced according to a balanced rendomized block design.

The feedback gains to the hand controller were varied along the three major orthogonal axis of controller elevation, extension, and azimuth. The gains were varied from full to one half to zero feedback. Each axis was varied independently. (The exact value of the feedback at the hand controller is obviously a function of the controller kinematics during the motion; however, the force measured at the hand reference frame did not exceed two pounds.) Additionally, the direction in which the feedback was provided was varied between assistive and resisitive. This was accomplished by changing the position of the manipulator through reflection of its end point about the Ss' target position. This had the effect of potentiating motion from the start position and attempting to drive the hand beyond the target point into concurrence with the manipulator end point. (See Figure 4.)

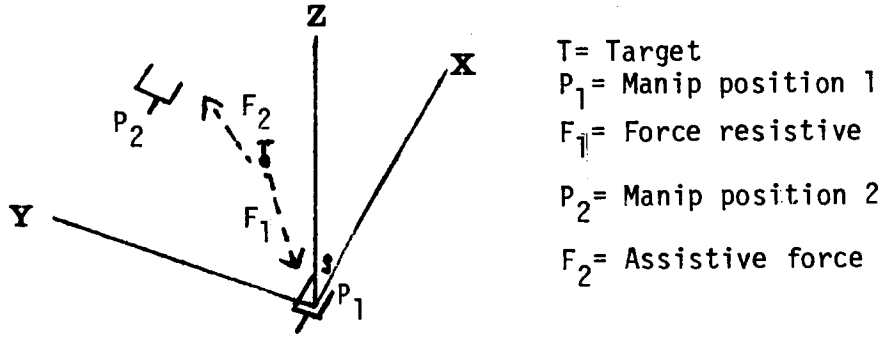


Figure 4. Example: Target Position, Reference Frame, and Feedback

Results

The results of these experiments are preliminary. However, it has been establised that the HO can attain a stable target point in cartesian space based on kinesthetic feedback. Systematic variation of forces associated with the kinesthetic feedback characterisitic of a final position effects attainmant of that position in a way that implies that the target position can be considered to be defined as an equilibrial point among forces of contributing muscle groups, i.e., position is muscle force encoded.

These preliminary results support previous research data with the general finding that the control of manipulation proceeds most effectively if it is performed with force and dynamic feedback (cf. Herzog, 1969, Merhav, 1976). The results also have implications for control of TOSs in space. The microgravity of the orbital environment introduce changes in the dynamic properties of the controlling limbs as well as changes in the dynamics of trajectory formation. Further investigation of the controlling variables in human limb movement, and the sensitivity of those control variables to environmental perturbation are ongoing at the Advanced Teleoperator Laboratory, JPL.

Experiment Two

The purpose of the second experiment was to determine, within the constraints of a prescribed manipulation task, the information seeking strategy of the operator of a teleoperator device. A theoretical formulation of this problem was presented by Sheridan (1970) under the rubric of Supervisory Control. The approach maintained that the task of a supervisor could be considered to optimally sample the controlled process assuming a cost for information and a normally distributed time varying cost for uncertainty. The optimization procedure, dependent upon cost per sample and information value, defined sampling strategies between distributional knowledge of the probability of the state of the process and continuous sampling, providing instantaneous certainty about the state of the process.

In the control of teleoperator devices, especially space teleoperators, there is a significant cost in providing feedback information to the operator. The costs occur because of high informational and attentional demands in space operation, computational and equipment volume constraints, and bandwidth and power constraints. The present experiments were undertaken to assess the sensory sampling strategy of the operator of a teleoperator with computer augmented displays. Description of the process of information valuation in teleoperator control is a step in the development of efficient and computer aided sensory data selection in Supervisory Systems.

Procedure

Subjects were trained to asymptote on a peg placement and peg insertion task using the manipulator, FRHC, and feedback modalities described above. During training, each subject was instructed as to the use of each feedback method (direct view, mono TV, stereo TV, kinesthetic feedback, and schematic displays) and was given a demonstration of each modality's information content during representative portions of the placement task. During this training, individual strategies for the use of feedback during portions of the task emerged. These were tabulated by recording frequency and dwell time for the modality sampled as a function of the ongoing task.

At asymptotic performance for each subject, the following constraint was placed on the subject's use of feedback: in selecting one or more of the feedback modalities available (excluding direct vision) the subject would incur a cost for that information through a reduction of available task completion time. (Available task time was one standard deviation of completion times during training added to the asymptotic completion time.) That reduction was a linearly additive function dependent on the number and type of feedback modlities chosen. The relative contribution of the structure of this cost function to feedback selection strategy is currently being investigated.

The performance measures were running task time (i.e., task performance time with feedback selection and initialization time subtracted), and frequency and dwell time for the sensory samples.

Results

Again, the results reported are a function of preliminary analysis. The results indicate that individual establish sensory sampling strategies which are robust in the face of cost for use of the information considered important. That is to say, that subjects maintained their established strategies to the point of being unable to complete the task under those strategies. Also, some strategies yield significantly better performance than others. These strategies could be considered as approaching optimality in the sense that the sensory data provided appropriate information for sampled period as reflected by the time performance measure. The distribution of these sampling periods are a task dependent function. A systematic description of the form of this sampling distribution is currently being undertaken.

CONCLUSION

These experiments and preliminary data are directed toward characterising the behavior of the human operator in control of a

teleoperator device. That characterization should take into account the roles of the operator in teleoperated systems and the control parameters contingent upon those roles. The results of the first experiment with the operator as manual controller are being further analysed to model the interaction between operator limb dynamics as a control parameter and teleoperator dynamics as response and feedback parameters. The results of the second experiment are being examined to attempt to define a distribution of task optimal sampling information to reduce control burden and aid in design of intelligent control interface.

REFERENCES

Bejczy, A., Sensors, controls, and man-machine interface for advanced teleoperation. Science, 1980, 208 (4450), pp. 1327-1335.

Bejczy, A. K., and M. Handlykken. Experimental results with a six-dgree-of-freecom force-reflecting hand controller. Proceedings of the Seventeenth Annual Conference on Manual Control, Los Angeles, California, October 15, 1981, 465-477.

Bizzi, E., Accornero, N., Chapple, W., and Hogan, T. N., Arm trajectory formation in monkeys, Experimental Brain Research, 1982, 46 (1) 139-143.

Bizzi, E., Dev, P., Morasso, P. and Polit, A., Effect of load disturbances during centrally initiated movements, Journal of Neurophysiology, 1978, 41 (3), 542-556.

Brooks, T., Superman: A system for supervisory manipulation and the study of human computer interactions. Master's Thesis, Massachusetts Insitute of Technology, May, 1979.

Cooke, J. D., Dependence of human arm movements on limb mechanical properties. Brain Research, 1979, 165 (2), 366-369.

Ferrell, W. R. and T. B. Sheridan, Supervisory Control of Remote Manipulation, IEEE Spectrum, 1967, Vol. 4 (10), 81-88.

Handlykken, M., and Turner, T., Control systems analysis and synthesis for a six degree-of-Freedom universal force-reflecting hand controller. Proceedings of the Nineteenth IEEE Conference on Decision and Control, December 10-12, 1980, Albuquerque, New Mexico, 1197-1205.

Herzog, J. H. Proprioceptive cues and thier influence on operator performance in manual control. NASA CR 1248, January, 1969.

Howard, R. Information value theory, IEEE Transactions on Systems

Man and Cybernetics, 1966, SSC-2 (1), 22-26.

Merhav, S. J. and Ya'Acov, B. O., Control augmentation and workload reduction by kinesthetic information from the manipulator, IEEE Transactions on System Man and Cybernetics, 1976, SMC-6,(12), 825-835.

Pattipati, K. R. Dynamic decision making in multi-task environments: Theoretical and experimental results. Ph.D. Thesis, University of Conn., August, 1980.

Pew, R. W., and Baron, S., The components of an information processing theory of skilled performance based on an optimal control perspective. In G. E. Stelmach (Ed.) Information Processing in Motor Control and Learning, New York: Academic Press, 71-79.

Sheridan, T. B., On how often a supervisor should sample, IEEE Transactions on Systems Science, 1970, SSC-6 (2),140-145.

Tulga, M. K. and T. B. Sheridan, Dynamic decisions and workload in multi-task supervisory control, IEEE Trasactions on Systems Man and Cybernetics, 1980, SMC-10 (5),217-232.

White, T. N. Modelling the human operator's supervisory behavior, in Proceedings of the First European Conference on Human Decision Making and Manual Control, H. Stassen (ed.), May, 1981, 203-207.

MANUAL CONTROL INTERACTION WITH ROBOTIC SYSTEMS

John Garin

Westinghouse Industry Automation Division

Pittsburgh, Pennsylvania

ABSTRACT

John Garin, Industry Automation Division, Westinghouse Corporation, will discuss manual control in the development of large scale robotics for hostile industrial environments. Although such systems involve removal of direct human contact with the work place, human operators usually function in a more centralized capacity, perhaps in a central command - control work station. Human surveillance of the remote work place involves development of appropriate CCTV systems. Repair and maintenance of the robots may require the use of teleoperator systems with associated remote control display problems. Programming of robots for task performance involves development of integrated control systems. The human remains part of the system, and assumes functions of a higher order than these functions delegated to robots.

The human interaction with the robotic device in an industrial environment is somewhat different than with a teleoperated device, since the human doesn't have to be an integrated part of the control loop. Most robot devices today, have the operator interface, either in the form of a teach pendant, a keyboard and CRT, or both. The teach pendant is a device that is an umbilical to the robot. This device has an average of twelve (12) switches each used to give a variety of discrete commands such as:

- job function
- speed control/velocity
- direction of motion clockwise/counter-clockwise
- hand effector open/close
- emergency stop
- teach mode or jog mode etc.

The teach pendant requires minimal feedback information to the operator. The feedback is usually visual, the robot moves or doesn't move. Things like record functions are indicated by a light or an alpha numeric LCD display.

The other form of interaction is through a keyboard and CRT. This usually requires an individual with a programming background and a significant amount of training related to the robotic system.

Typically, most robots have a form of teach pendant as standard man-machine interface. The CRT and keyboard is optional. In today's industrial robot, not too much attention has been paid to manual controls or human factor as it relates to man-machine interface. The primary concern is the robotic device.

MANUAL CONTROL USING COORDINATES RELATIVE TO THE EFFECTOR IN TELEOPERATION

N. QUETIN
Laboratoire de Robotique d'Evry
Université Paris XII
Quartier des Passages
91000 EVRY (France)

The work presented here concerns the manual control of teleoperators. The main objective is to improve performance by the choice of an ergonomic control interface (control transducer + langage + feed-back of information). This applies to remote control operations in the industrial field (forging, welding) and to intervention by remote control in a hostile medium (spatial, undersea, nuclear).

The variables under remote control in an up to date system are essentially movements and operations carried out by the effector of the teleoperator ; being signals generated by either a transducer and/or by a computer system. The totality of the coordinates of translation and rotation from which the movements of the effector will be determined constitutes the "command reference".

The choice of a command reference adapted to the task and conditions of telemanipulation (manipulator of a fixed or carrier borne control point, direct oversight of the scene or TV monitoring) is decisive vis-à-vis the results obtained.

In a given task there are priviliged references "linked to the task" in which these gaps are easily defined.

Example : A reference linked to the task of inserting a cylinder in a hole situated on an inclined plan e is that defined by the normal $\vec{N}$ to this plane.

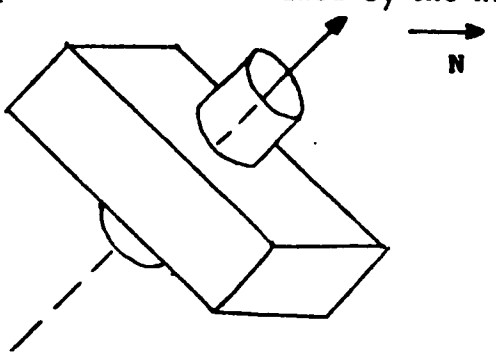


Fig. 1. Reference linked to the task of inserting a cylinder.

A command reference will be adapted if it is near to the task to be carried out (save for uncertainty relative to play in mechanical movements and to flexing of the system) and natural for the operator in relation to the conditions of the teleoperation.

1) THE REFERENCES LINKED TO THE EFFECTOR

Amongst all the references possible, the ones of particular interest are those linked to the effector, these relative references (mobile in the work area) present three main advantages:

- They are of considerable general use, as they are adapted to a wide range of tasks.

- They allow a reduction in the number of variables controlled by the operator, by restricting the command reference to the sole variables characterising the task under way.

- They allow suppression of the notion of absolute reference fixed in the area of work, a notion which loses all meaning for the operator in many cases of teleoperation:

- When the operator has to be mobile in relation to the effector in order to oversee the operation (forging, welding).

- When the operator is monitoring the scene on TV.

In these different situations, the effector (or its image) will constitute a permanent reference for the operator.

Moreover, where the scene is being monitored on TV from a camera carried on the set in proximity to the effector, a reference linked to the latter will coincide with the visual reference of the operator.

To allow us explain the two first advantages of such references, we must detail the use of a cartesian reference linked to the effector (XYZ) which determined the translations of a reference point and the axes of rotation of the effector around it.

I.I. USE OF A CARTESIAN REFERENCE LINKED TO THE EFFECTOR

This reference, very natural for the operator, is seen in Fig. 2 ; to simplify the schema of this paper we present a case where the effector is a grip with two fingers.

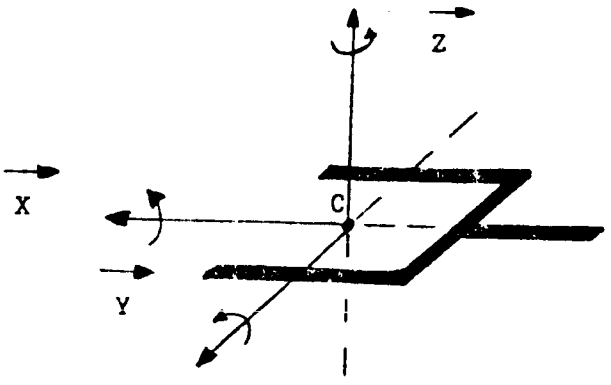


Fig. 2 CARTESIAN REFERENCE LINKED TO THE EFFECTOR

Remark 1 The use of such a reference can of course be generalised, whatever the tool held as soon as position of this tool in relation to the terminal organ is known.

The aim of the following examples is to show how such a reference linked to an effector enables the number of variables controlled by the operator to be restricted to just those defining the task under way.

Example : TRANSLATION OF THE EFFECTOR IN A GIVEN DIRECTION IN SPACE

After having correctly oriented the effector by means of the three rotations around the axes $\vec{X}\vec{Y}\vec{Z}$, the operator is able to control the task by using the translation variable alone X, Y, Z depending on the orientation wanted .

Example 2 : TRANSLATIONS OF THE EFFECTOR IN A GIVEN PLANE IN SPACE

After having correctly oriented the effector, the operator controls the task by using two variables of translation of cartesian reference X Y; Y Z, or X Z, depending on the orientation wanted for the last one.

We have shown in fig. 3 the case where the two variables controlled are translations of the effector according to X and Y.

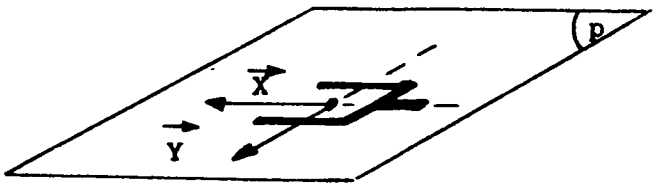


Fig. 3. Translation of the effector in a plane (according to X and Y.)

Example 2.I. If the effector has to be moved in a plane parallel to the material plane of the work area, and at a strictly constant distance from it (visual inspection by camera, various projections) the management of the distance between the effector and the plane can be undertaken automatically by the control system, using, for instance, proximity detectors.

Example 2.2. If the effector has to be moved on a material plane of the work area, applying a strictly constant pressure upon it (microwave, V.S. inspection) the management of the force exercised by the effector in the direction normal to the plane (which coincides with one of the axes of the cartesian reference linked to the effector) can be automatically, undertaken by the control set using, for instance force detectors (compliance programmable in the normal direction of the plane).

Remark I By using both types of control described in examples 2.1 et 2.2. the operator is able to control more precisely the translations of the effector parallel to or on the material field of work.

Remark 2 To carry out the operations described in examples 1 and 2 in a reference of absolute cartesian coordinates, simultaneous control of three independant variables (associated with the three translations) is required.

Example 3 MOVING THE EFFECTOR PARALLEL TO A SURFACE IN SPACE

After having correctly orientated the effector, the operator can use a mode of control restricted to the four cartesian variables linked to the effector: for example, translations along X, Y and the two rotations around them. This mode of control allows the operator to move along two curvilinear coordinates of the surface (Fig. 4).

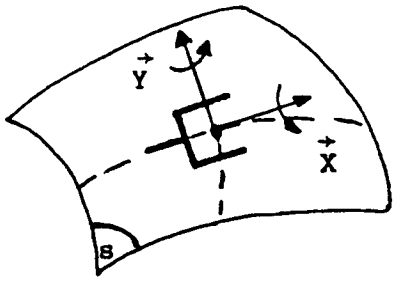


Fig. 4 - MOVING THE EFFECTOR PARALLEL TO A SURFACE IN SPACE

Examples 3.1. and 3.2.

They can be directly deduced from examples 2.1. and 2.2. by substituting "plane" for "surface".

Remark I To effect the task described in example 3 in an absolute cartesian reference the simultaneous control of six variables is required.

CONCLUSION

These examples show how the use of a relative reference allows the number of variables controlled by the operator to be restricted to just those defining the task.

In the second part we are going to describe the coupling of a manual syntaxor with six degrees of freedom to a manipulator MA 23 in such a reference.

II. Application for coupling a manual syntaxor to a manipulation MA 23

The syntaxor and the MA 23 are shown in fig.5 .

II.I. DESCRIPTION OF THE MANUAL SYNTAXOR

This syntaxor (Fig.6) has been realised in the laboratory with the aim of obtaining a control organ with 6 degrees of freedom which will allow the operator to work seated and to remain within the zone of manual comfort. This syntaxor would be portable and so of reduced dimensions.

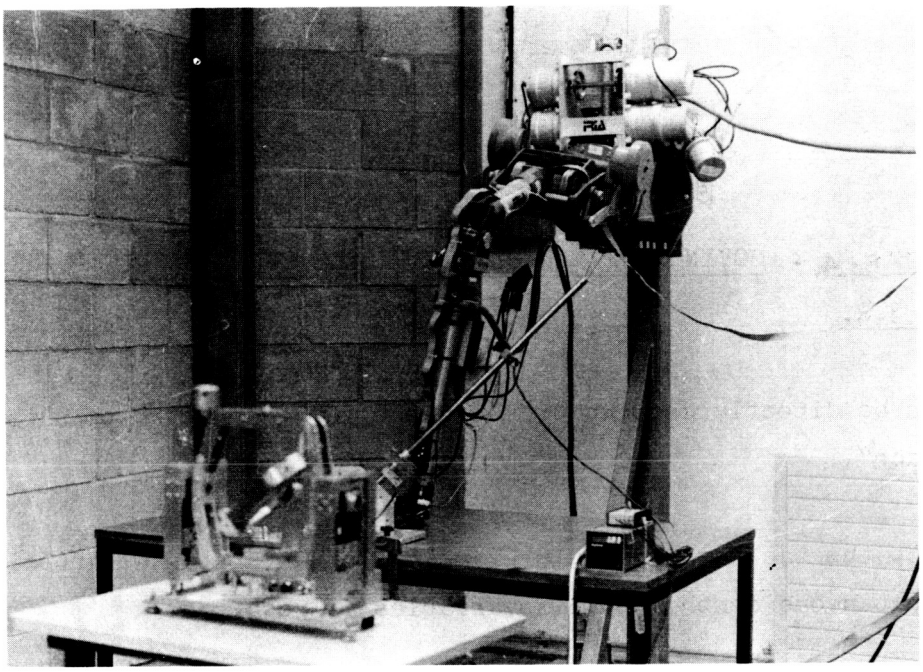


Fig. 5.

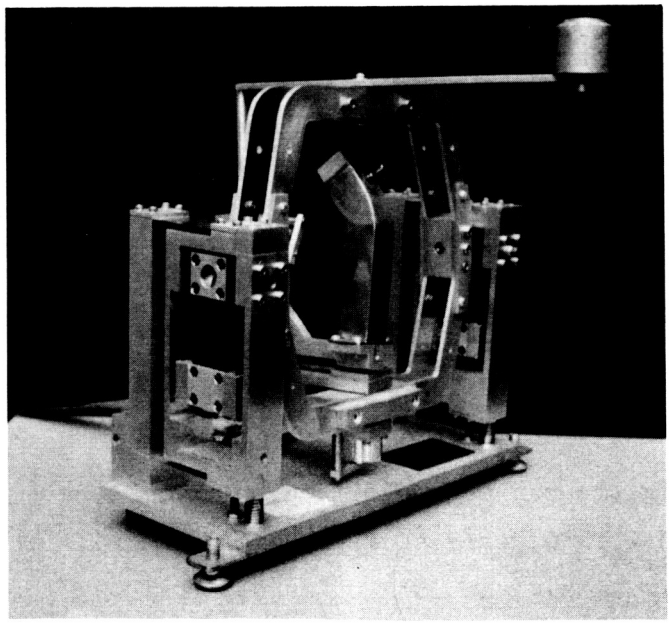


Fig. 6. THE SYNTAXOR WITH SIX DEGREES OF FREEDOM

II.2. DESCRIPTION OF THE SYNTAXOR/ MA 23 COUPLING

We have chosen to use three microdisplacements of the handle according to the axes $(\vec{a}, \vec{s}, \vec{n})$ to control the speed of point C of the handgrip in the cartesian reference $\vec{a}_p, \vec{s}_p, \vec{n}_p$ linked to the latter and the rotations of the handle around $(\vec{a}_s, \vec{s}_s, \vec{n}_s)$ to control the angles of rotation of the handgrip around $(\vec{a}_p, \vec{s}_p, \vec{n}_p)$. This coupling is shown in Fig. 7, the handle held by the operator is represented by its section (s).

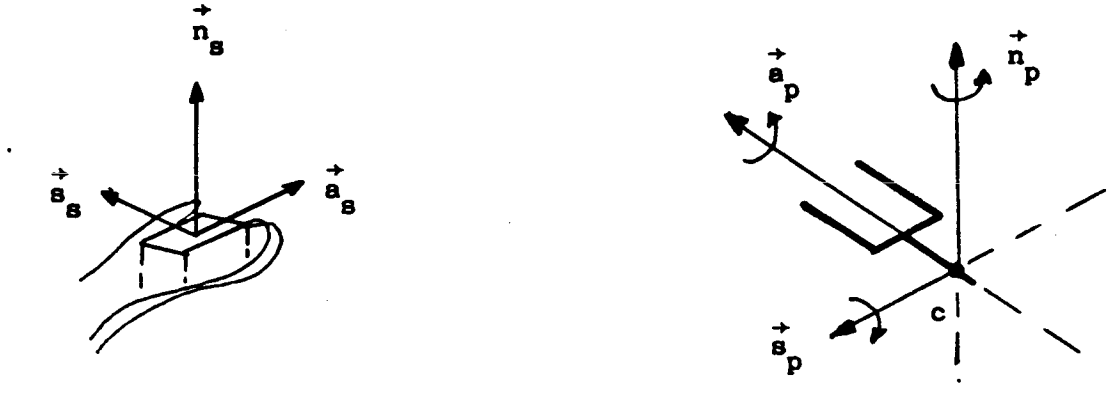


Fig. 7. COUPLING OF A SYNTAXOR (S) AND THE HANDGRIP MA23 (p)

The six degrees of freedom correspond to three rotations of the handle around a fixed point and three microdisplacements of itself around three axes of orthogonal reference $\vec{X}, \vec{Y}, \vec{Z}$. Moreover two contactors allow the generation of logical signals which can be used, for instance to open and close the terminal organ. The equilibrium of the system is assured by a counter-weight and the return of the translation movements by springs.

II.2.1. NEED FOR A REINITIALISATION PROCESS

To allow the operator to move the MA 23 grip throughout its entire field within all possible orientations, whilst still remaining in the zone of manual comfort we use a process called "pumping". This allows the operator to disengage the mechanism (the signals from the handle are no longer taken into account) to place the handle in a comfortable position and to reengage. Such a process associated with the choice of the coupling handgrip of MA 23 which we have made requires a reinitialisation process after each engagement that is to say, the initial syntaxor reference is made to coincide with the initial grip reference at the moment of engagement.

II.2.3. TECHNICAL PROBLEMS OF COUPLING

The calculations required by the coupling of the syntaxor with the MA 23 in that relative reference were made by a Solar 16,65 computer, as were the calculations of transformations of coordinates of the syntaxor itself and the MA 23. The complete calculation is fairly irksome and has required an increase in the sampling period from 40 to 60 ms. In order to evaluate this coupling quantitalively we intend to set the system on a manipulator MAT I where the geometry results in are easier coordinate transformer (uncoupling of the translations and rotations of the handgrip).

CONCLUSION

This paper shows the interest of using relative references linked to the effector, thus essentially allowing the operator to work with a reduced number of control variables adapted to the task. This type of reference is, morever, adapted to telemanipulation in industry when the operator needs to be very mobile and to remote controlled operations overseen by means of TV display.

BIBLIOGRAPHIE

. J. GUITTET, N. QUETIN. Définition d'une méthodologie pour l'étude comparative des modes et moyens de commande d'un robot manipulateur spatial. Convention CNES. Mars 1981.

. J.P. GAILLARD. The effect of electrotactile and visual force feedback on learning and performance in a manipulation task. European annuel conference on human decision and manual control. Delft. Holland, mai 1981.

. V. DUPOURQUE, J.P. GAILLARD, N. QUETIN. Organes de commande manuelle, étude bibliographique et préévaluation d'une poignée à six degrés de liberté. Convention CNES. Août 1981.

. J.P. STROMBONI. Performances du tandem homme-machine en téléopération, 2nd IASTED International Symposia. Davos, mars 1982.

HUMAN DECISION MODEL IN FAILURE COMPENSATION

by

HIROSHI YAJIMA, THOMAS B. SHERIDAN, AHMET BUHARALI

Department of Mechanical Engineering,
Massachusetts Institute of Technology

SUMMARY

The research described here is directed towards developing the structure of a decision model in failure compensation, paticularly a decision model for conflict resolution in failure situation. Failure compensation is one of the important aspects of human intervention to automatic system control. The operator must perform not only tasks under prefixed procedures for failures, but also has the task of problem solving. One of these problem solving activities is the consideration of multiple performance criteria of the process. To the end of assisting the operator in this activity, a conflict resolution model of the operator is studied.

The paper presents the result of a series of experiments about a conflict resolution model. The results suggest that the computer assistance in prediction of control results helps decision making in failure situations, and that the decision model must time-variant. Finally, it appeares that a model that consists of a set of multi-attribute utility function fits the experimental data in failure compensation.

INTRODUCTION

The task of the human operator in complex large scale systems are not only to manage the resources of the system in normal operation, but also to intervene in the automatic control in order to compensate for system Failure. In failure compensation, there seem to be three main aspects: decision making, display interpretation and compensation strategy construction. The first two aspects are studied in this paper.

Until now, there have been many researches concerning failure compensation. Rouse proposed a failure compensation model, and emphasized the coordination of failure detection and failure compensation [1]. On the other hand, many researches have been performed from the application side [2],[3]. But, few researches have focused on the aspect of conflict resolution in failure compensatory decision making. Conflict resolution

between performance criteria is considered to be at the knowledge level of the three level decision structure that has been proposed by Rasmussen [4]. This decision-making problem is classified as a multicriteria optimization problem. Many methods have been developed for multicriteria optimization [5],[6]. However, in failure compensation, some on-line method is required. The multi-attribute utility function is considered most preferable. However, in using the multi-attribute utility, there seem to be many problems such as the utility independence between multiple criteria or parameter setting among the different failure situations.

An experimental paradigm was developed for this study. In this experiment, the operator had to make some trade-off betweet two criteria, namely, (1) the performance index of normal operation and, (2) the abnormal operation criteria which affects the reliability of the system using different types of display.

Some basic questions are:

(1) How does the operator decide on his control strategy in failure compensation ?
(2) How will his preference change with a transition of system failure state ?
(3) What kind of displays are most suitable for the operator ?

EXPERIMENTAL SCENARIO

The experiment is performed under the following conditions: multi-task, forced pace and incomplete information about the task. In this experiment, the subject is required to execute tasks to earn as much score as possible in limited time.

SYSTEM STRUCTURE

The structure for the experiment is illustrated in Fig. 1. It consists of process, display and assistance logic. The subject controls the system using information provided on both the process and it's performance. The process dynamics and computer assistance program are all implemented on a PDP 11/34 and all required information is presented to experimental subjects on a Megatek vector graphic display. The subject enters his control decisions through a digital control box.

PROCESS STRUCTURE

The process consists of multiple Work Areas (WA's) where tasks to be executed appear at random, and proceed towards a deadline. The tasks disappear when the deadline is reached (Fig. 2). One experimental run consists of responding to many tasks, both normal and abnormal. Each task has three attributes: type, score and position as shown in the task of WA 1 in Fig. 2. Corresponding to these two types of task, the process has two different criteria, namely the positive score that corresponds to the process product and the negative score corresponding to process damage. Execution of the normal task increases the normal score and failure to execute the abnormal task increases the negative score.

OPERATOR INTERFACE

The display seen by the operator are illustrated in Fig. 3. The details of each display are shown in Table 1. Display 1 gives the process state information consisted of the value of the task attributes. The task locations are indicated as to whether the task is in the upper half (E1) of a WA or lower half (E2). Display 2 shows the performance that are measured by the scores mentioned above. Display 2 also shows the predictions of these performances as well as time.

PROCESS PARAMETERS

1. TASK PARAMETERS

In each work area, different tasks may be generated throughout the operation of the system. Each task will be characterized with a number of parameters as shown in Table 2. The occurrence rate of the normal task and abnormal task is the only variable of the experiment and three cases of this variable shown in Table 3 are studied. The occurrence rate affects the coordination between the execution of normal task and that of abnormal task. The occurrence rate changes in three steps in the experiment according to the negative score that the subject receives. Task appearance position and speed determine the amount of time the decision maker has in order to respond to the task.

2. EXECUTION PARAMETERS

A manual manipulator (MM) moves between WA's according to command of the subject, and automatically executes any tasks in the WA where it is. Task execution speed of the MM is constant. The rate (score/execution speed) specifies the amount of time the subject has to spend in order to successfully complete the task.

EXPERIMENT

THe basic rules of the experiment are as follows;

(1) Maximize the positive score.
(2) The negative score must be under it's limit.
(3) Available time is limited.

Two kinds of experiments were performed. The first one was in full manual mode to make the effect of the computer assistance clear. The second was in the computer assistance mode and was designed to study the human decision model. In the second experiment, three different variations were used as noted below.

EXPERIMENT 1

In this experiment, the subject watched display 1 and one part of display2 (display 2-1 shown in Fig. 3), and moved the manual manipulator (MM) by pressing the buttons on the control box. The subject did not know the exact position of each task or the presence of a second task in each WA. In this experiment, there was no computer assistance. The parameter values in this experiment were as shown in Table 4.

EXPERIMENT 2

In the second experiment, the computer identified the state of the process and calculated the MM control actions according to the three kinds of strategies. These strategies are shown in Table 5. The subject watched the whole display, and decided the most preferable strategy, looking at the predicted values of positive score, negative score and time remaining for each of the three strategies. When he selected one of these strategies, the computer moved the MM automatically. The prediction values for strategies in display 2-2 were refreshed every 2 seconds. Typical behaviour of the subject was as follows;

(1) Monitoring : during the experiment, the subject monitored the display in these three respects;
   (a) emergence of additional task
   (b) deviation of scores from the predicted values
   (c) change of prediction values during refreshing of display 2

(2) Intervention : the subject decided to change the current strategy from the information of display 2-2 or display 1

(3) Decison making : the subject selected one of three strategies provided by the predicted values

In this experiment, three types of trials were performed.

(a) Type A : maximize positive score under two constraints i.e. negative score and time
(b) Type B : maxamize positive score under one constraint i.e. time
(c) Type C : gain the prefixed positive score under two constraints i.e. negative score and time

Experimental trials were designed to investigate how the constraints affect decision behaviour.

EXPERIMENTAL RESULTS

The score comparisons between experiment 1 and experiment 2 are shown in Fig. 4. The first result seems to suggest that computer assistance makes the positive score increase or negative score decrease in the most cases. The reason of this result can be easily understood , namely, that the computer can treat more information than the human. Actually, the subjects felt the large amount of workload in experiment 1. But the difference between positive scores in two experiments was not so large. This was because the score in experiment 1 depended on the operator skill and the strategies of the subjects, so the effect of computer assistance varied from subject to subject. The histories of negative and positive score for each subject in the three types of trials are illustrated in Figs. 5, 6 and 7. In these trials, the shapes of score histories of each subject are similar. From these data, it may be said that the subject doesn't change his strategy selecting policy according to the type or number of constraints, that is, in all trials YA is risk aversive, MY seems to be neutral, and MI is rather risk prone. From the history of negative score (Fig. 8), it is found that each subject changed his

strategy in selecting the computer strategies during the experiment. When a constraint was added, the gradient of history of negative score became steeper.

The changing points of the history curve in the negative scores were also analyzed using the history of the parameter (predicted negative score in selected strategy / time) as illustrated in Fig. 9. For the first 10 seconds, the subject preferred the most risk aversive strategy, bur after that period he turned to the most risk prone strategy. This changing time corresponds to the timing where the history of the negative score began to increase (Fig. 8). This can be considered as the time when the subject changed his strategy.

In all cases of experiment 2, it was difficult to built-up the utility function in the whole decision space, because subjects could not tell their preference for one of two attribute, independent of time and another attribute. From the experiment above, the following two aspects of the utility function become clear;

(1) The utility function is time-variant.
(2) It is impossible to assume utility independence in the whole decision space.

For most type of multi-attribute utility function, the assumption of utility independence is assumed. To cope with the second problem mentioned above, the next assumption is evaluated.

ASSUMPTION : LOCAL UTILITY INDEPENDENCE

Let p(x,n)=[x0, x(1),....,x(n-1),xe] be a partition of value range [x0,xe] on variable x, Q(i,j,k) be a sub-decision space of the whole decision space P(s,r,t) defined as follows;

$$Q(i,j,k)=[\ s(i-1)< s <s(i),\ r(j-1)< r <r(j),\ t(k-1)< t <t(k)\ ]$$

For the value ranges of positive score(s), negative score(r), and time(t), there exist partition p(s,N), p(r,M), p(t,K) where the multi-attribute utility function $U_{i,j,k}(s,r)$ can be defined as follows for any i (1<i<N), j(1<j<M), k(1<k<K);

$$U_{i,j,k}(s,r)= k_1*v_{i,j,k}(s) + k_2*u_{i,j,k}(r) + k_3*v_{i,j,k}(s)*u_{i,j,k}(r)$$

where $U_{i,j,k}(s,r)$ denotes the multi-attribute utility function in the sub-decision space Q(i,j,k), and $v_{i,j,k}(s)$ and $u_{i,j,k}(r)$ are single utility functions.

This concept means that if the value of each performance criteria is restricted in a small range, utility independence between theses criteria can be satisfied. This concept also tells the fact that though the weight of each performance criteria may be changed in consideration of the whole decision space, it will not change within one small part of the decision space. Following this assumption, one case of constructing multi-attribute utility function was evaluated. Here, the whole decision space was devided into 80 (4x4x5) areas, and many multi-attribute utility functions corresponding to these areas were built. In order to decide the single attribute utility function, a 0.5-0.5 lottery was used. One of the results

is illustrated in Fig. 10. The multi-attribute utility function in this figure seems to fit the experimental data.

EXPERIMENT IN PROGRESS

For the development of a decision model, many issues seem relevant. The first point is that there are some decisions in the experiment that break the consistancy of the utility function approach (point A and B in Fig. 10). This default seems to be understood as the fuzzyness in human judgement. Further research should help to make this point clear.

The second point is that it is rather difficult to built up large number of utility functions of small areas using the questions of lottery to the subject. So, some method to compose utility function of one small area from a small number of utility functions of other small areas developed from questions of lottery will be needed. Here, we must build the factors that decide the shape of utility functions.

The final point concerns the type of display. We are now developing a multi-criteria display whose axes are positive score and negative score. That is the same as Figure 10. In this display, the prediction values are also displayed as graphs. This display will be validated by running detailed experiment.

REFERENCES

[1] W.B.Rouse :"Models of Human Problem Solving : Detection, Diagnosis and Compensation for System Failure" IFAC Symposium on Analysis, Design and Evaluation of Man-Machine System, Baden, FR Germany, Sept. 1982.

[2] W.J.Dellner :"The User's Role in Automated Fault Detection and System Recovery", J.Rasmussen and W. Rouse, editor, " Human Detection and Diagnosis of System Failure" Plenum Press, New York, 1980.

[3] D.A.Lihou :"Aiding Process Plant Operators in Fault Finding and Corrective Action" Ibid.

[4] J.Rasmussen :"The Role of Cognitive Models of Operators in the Design, Operation and Licensing of Nuclear Power Plants", Proc. of Workshop on Cognitive Modeling of Nuclear Plant Control Room Operators, Nurecj/CR-114, ORNL/TM-8614, 1982, pp13-35

[5] Y.Y.Haimes,W.A.Hall,H.T.Freedman :"Multiobjective Optimization in Water Resource Systems: The surrogate Worth Trade-Off Method", Elsevier Scientific Publshing. Amsterdam, 1975

[6] R.L.Keeny :"Utility functions for multi-attributed consequences", 56749494

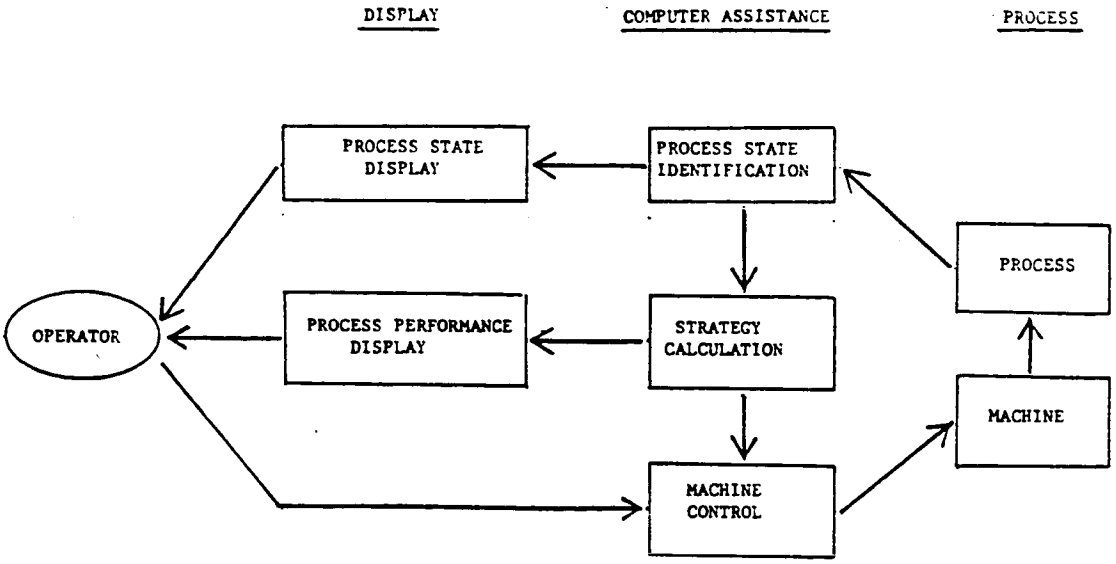


FIGURE 1. SYSTEM STRUCTURE

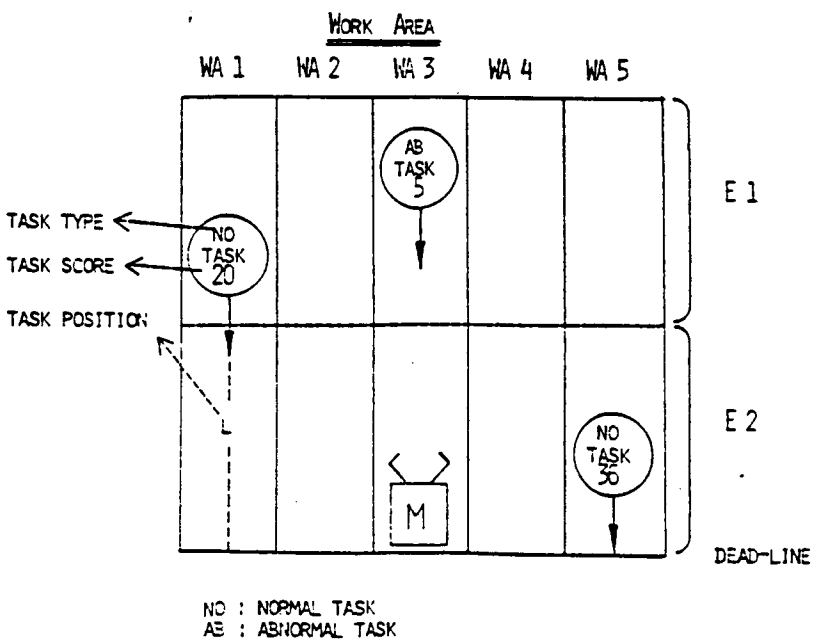
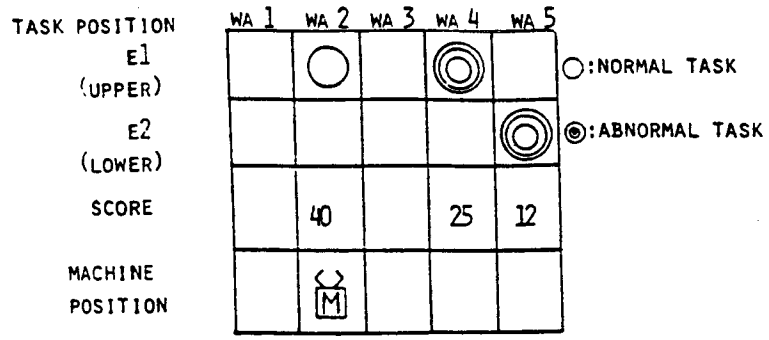


FIGURE 2. PROCESS STRUCTURE

DISPLAY 1. (PROCESS STATE DISPLAY)



DISPLAY 2. (PROCESS PERFORMANCE DISPLAY)

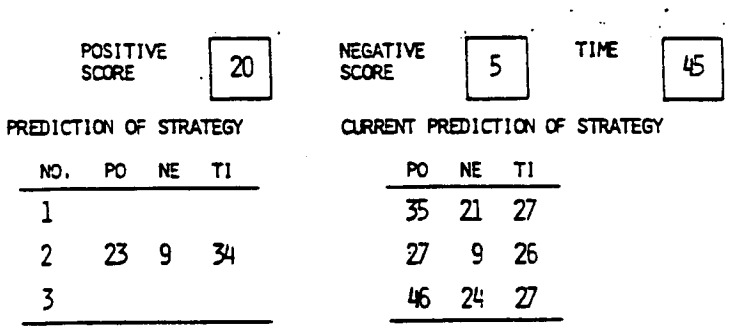


FIGURE 3. DISPLAY IN EXPERIMENT

TABLE 1. DISPLAY SPECIFICATION

DISPLAY	INFORMATION	NOTE
DISPLAY 1	LOCATION OF TASK SCORE OF TASK TYPE OF TASK	EXPERIMENT 1,2
DISPLAY 2	POSITIVE SCORE NEGATIVE SCORE TIME LEFT	EXPERIMENT 1
	PREDICTION VALUES OF THREE OF TASK EXECUTING STRATEGIES. (PREDICTION CONSISTS OF POSITIVE SCORE, NEGATIVE SCORE, TIME LEFT)	EXPERIMENT 1,2

TABLE 2. TASK PARAMETER

PARAMETER		VALUE
TIME OF TASK OCCURRENCE	ABNORMAL TASK	STEP FUNCTION
	NORMAL TASK	CONST
SCORE (W)		RANDOM VARIABLE BETWEEN [1.60]
INITIAL POSITION (1)		CONST
SPEED		CONST

TABLE 3. CASE OF VARIABLE

(SEC)

CASE	NORMAL TASK (Dn)	ABNORMAL TASK (Da)
1	8	12
2	8	10
2	8	8

$$Ta = k(r) * Da$$

$$k(r) = \begin{cases} 1.0 & (0 < r < 20) \\ 0.7 & (20 < r < 50) \\ 0.5 & (50 < r < 80) \end{cases}$$

$$Tn = Da$$

where

Ta: time of abnormal task occurrence
Dn: initial occurrence time of normal task
Da: initial occurrence time of abnormal task
r: negative score

TABLE 4. PARAMETER VALUE

PARAMETER		VALUE
SPEED		1
TASK INITIAL POSITION		100
TASK EXECUTION	MACHINE NUMBER	1
	SPEED	2
CONSTRAINT	TIME	50
	NEGATIVE SCORE	80

TABLE 5. THREE STRATEGIES IN EXPERIMENT

STRATEGY NO.	TSAK EXECUTION PRIORITY
1	NEARNESS TO ANY DEADLINE
2	EXECUTION PRIORITY IS BASED ON NEARNESS TO DEADLINE FOR ABNORMAL TASKS. IF ANY EXITS THEN ON NEARNESS TO DEADLINE FOR NORMAL TASK
3	EXECUTION PRIORITY IS BASED ON NE RNESS TO DEADLINE FOR NORMAL TASKS, IF ANY THEN ON NEARNESS TO DEADLINE FOR ABNORMAL TASKS.

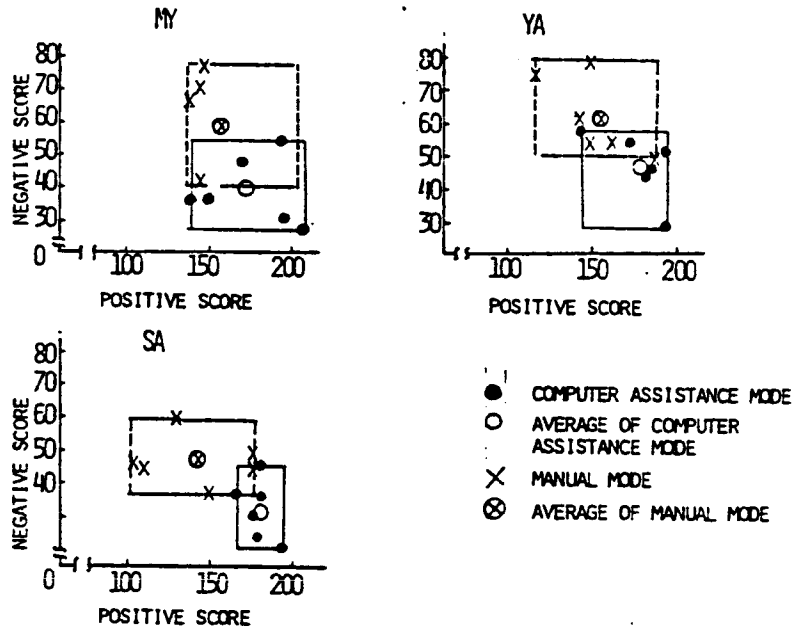


FIGURE 4. COMPARISON BETWEEN MANUAL MODE AND COMPUTER ASSISTANCE MODE

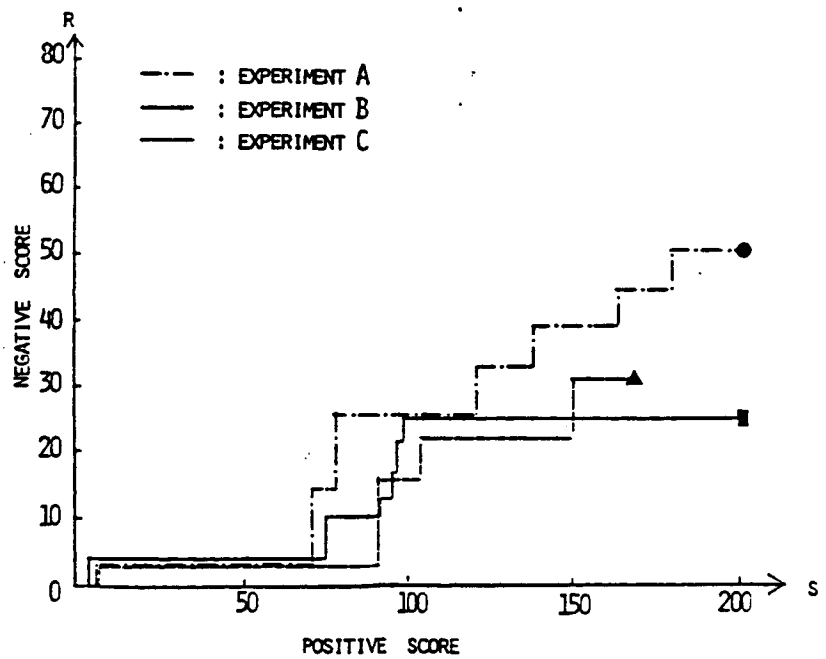


FIGURE 5. SCORE HISTORY OF YA

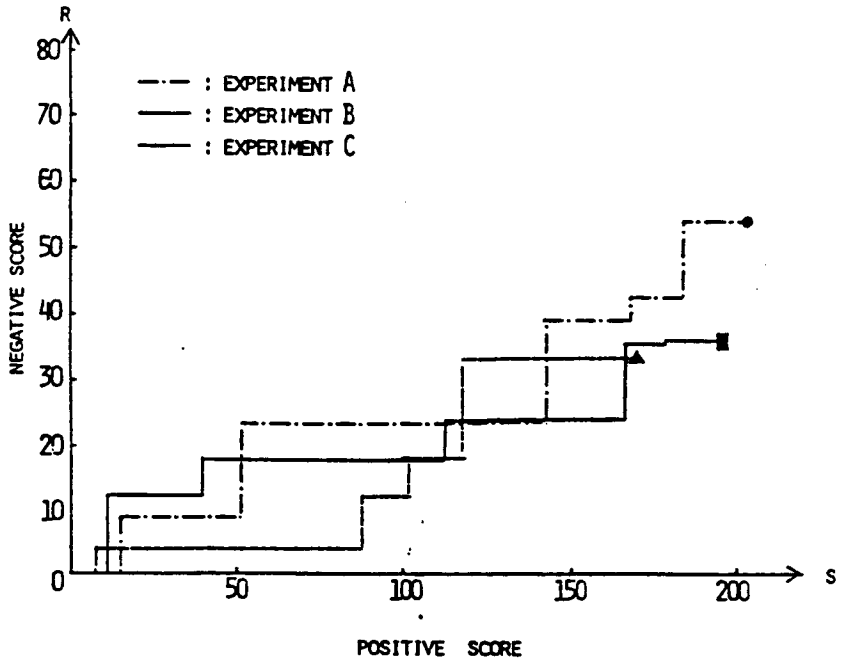


FIGURE 6. SCORE HISTORY OF MY

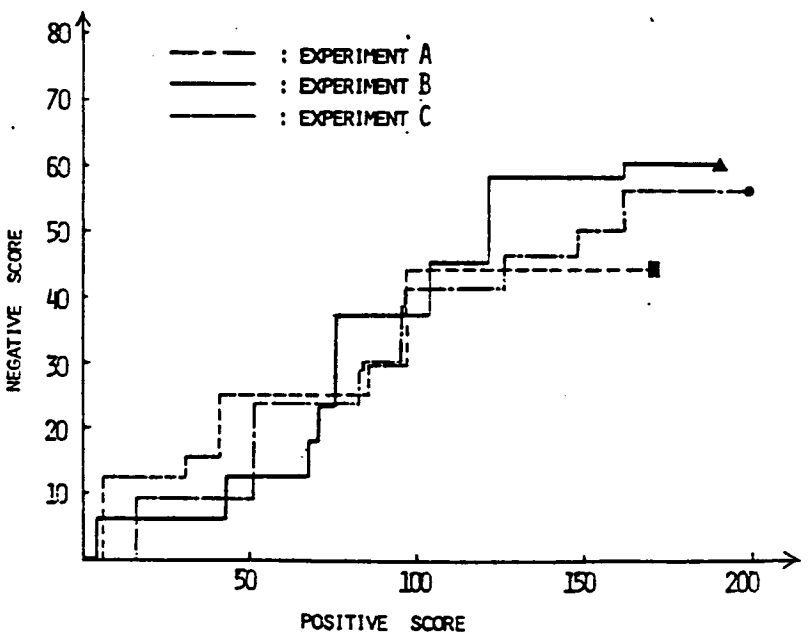


FIGURE 7. SCORE HISTORY OF MI

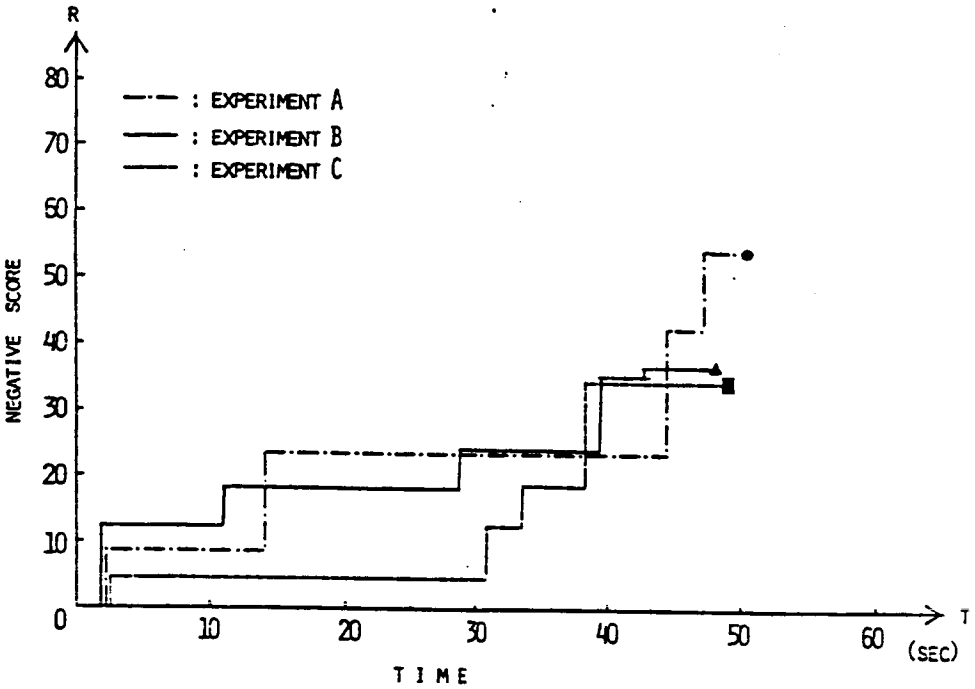


FIGURE 8 NEGATIVE SCORE HISTORY OF MY

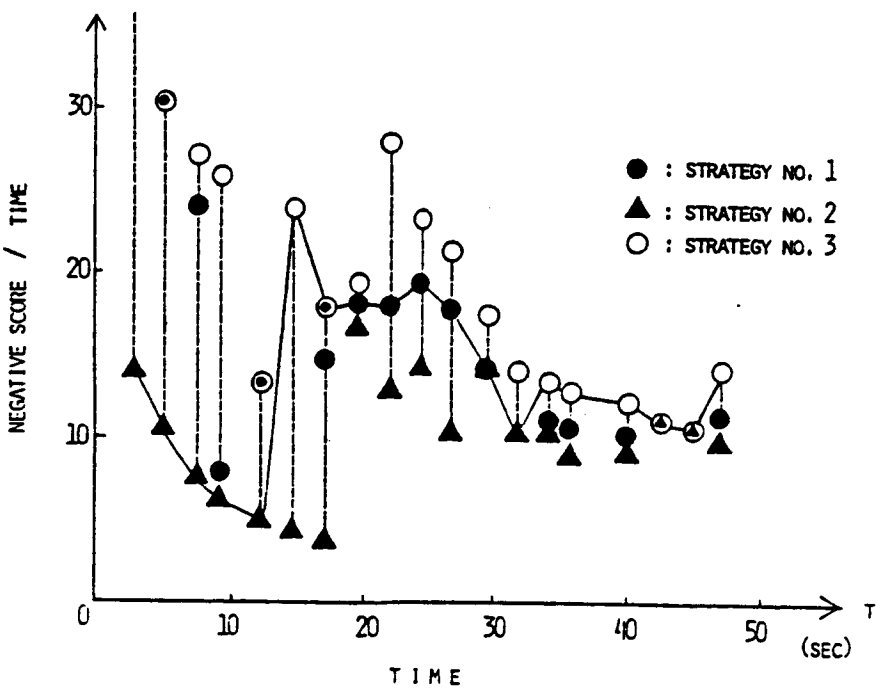


FIGURE 9. TRANSITION OF STRATEGY SELECTION

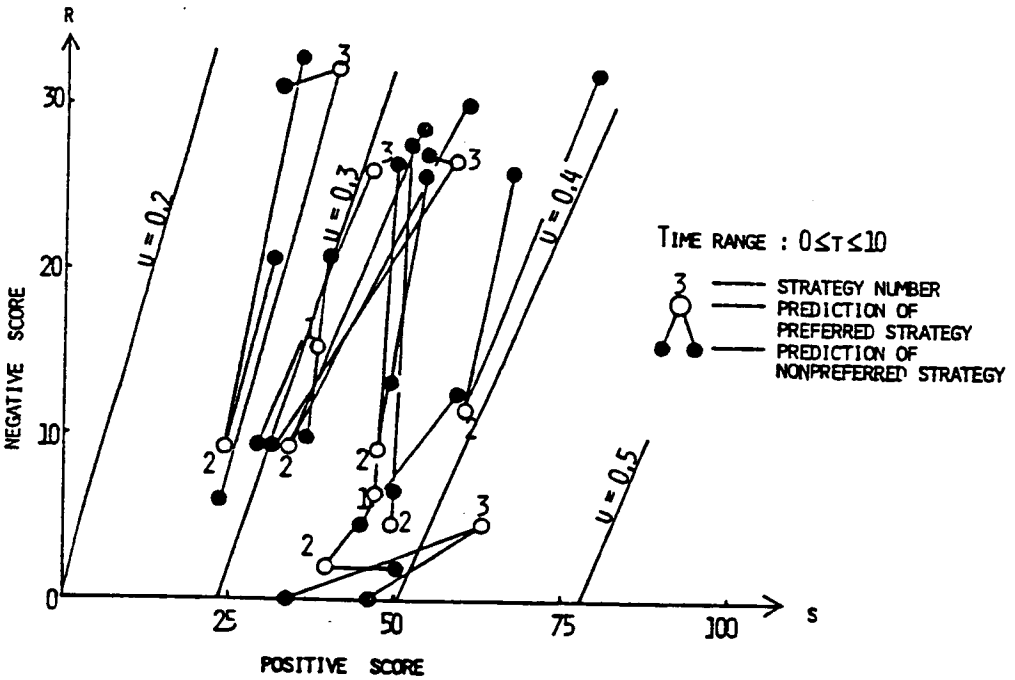


FIGURE 10. VALIDITY OF MULTI-ATTRIBUTE UTILITY FUNCTION MODEL

THEORETICAL AND EXPERIMENTAL ANALYSIS OF PILOT FAILURE DETECTION BEHAVIOUR DURING VARIOUS AUTOMATIC APPROACH CONDITIONS

by

R.C. van de Graaff and P.H. Wewerinke
National Aerospace Laboratory NLR
Amsterdam, The Netherlands

ABSTRACT

This paper deals with an investigation of pilots' failure detection performance in a realistic moving-base flight simulator. The task considered was the monitoring of stabilized automatic approaches. Abnormal conditions (excessive windshears or system failures) were to be detected on the basis of unacceptable deviations of displayed variables from their normal positions.
Four experienced airline pilots participated in the experiment. The independent variables were: failure type (windshear or system failure), the axis in which the failure occurred (longitudinal or lateral), and the prior probability of failure occurrence (0.2 or 0.8). The dependent variables were detection times and the various display deviations at the moment of detection.
Independently of the experimental program a theoretical analysis was carried out, based on a model of the human observer and decision maker which is formulated in terms of linear estimation and classical sequential decision theory. In the paper the experimental results are compared with the corresponding model predictions. In addition, attention is devoted to the extrapolation of the results from relatively probable to rarely occurring failures. The latter condition pertains to most real life situations and is a key factor in theoretical and experimental research in human decision making.

INTRODUCTION

Technological developments in aerospace are attended by more complex and more automated systems. These demand a better insight into human monitoring and decision making behaviour as the human operator's role gradually shifts from manual towards supervisory control. In the context of transport aircraft operation this is very much the case after the introduction of automatic approach and landing systems, and in particular in the future MLS-environment.
The last two decades considerably research effort has been devoted to the study of human information processing in the context of control behaviour. One result is a number of mathematical tools, of which the state-space, time-domain optimal control model has been shown to provide an adequate framework for describing the human processing of information of a dynamic system (Refs. 1-4). Recently, this model has been extended by a decision making model accounting for cognitive functions involved in monitoring an automatic system, detecting system failures, making decisions about its operation, etc. (Ref. 5). This decision making model is formulated in terms of multivariable classical sequential decision theory and accounts for the important effect of correlated information. The model will be briefly discussed in the following section.

In the subsequent section, an experimental program is discussed, which has been performed to investigate pilot decision making behaviour in a realistic simulator experiment, involving the detection of abnormal conditions (windshears or system failures). These situations occurred with a given probability during stabilized automatic approaches. The discussion of the experimental results and the comparison of these results with the model predictions are contained in the last section. Hereby, the effect of prior knowledge about the probability of failure occurrence is also considered, which is of importance for the relevance of extrapolating experimental results to real-life decision situations in aircraft operation.
Finally, the last section contains the principal results.

MODEL OF THE PILOT AS AN OBSERVER AND DECISION MAKER

This section summarizes the basic notions underlying the model of the human operator's functioning in supervisory control tasks, which was used for the description of the pilot's monitoring and decision making characteristics for the various experimental conditions; the model is extensively described in reference 5.
In the model it is assumed that the human observer perceives information of a linear dynamic system which can be described by a Gauss-Markov random sequence. Based on the learned dynamics of this system and the perceived information (i.e. noisy observations), he makes the best estimate of the system state. This is described in standard linear estimation theoretical terms (Kalman filter) and is part of the well documented optimal control model (Refs. 1 and 2).

Now, in the normal mode of operation, the discrepancy between perceived and expected information (the so-called innovation sequence n_k) is a zero mean Gaussian random sequence with covariance N_k. It is assumed that abnormal system operation, as caused by errors in display instruments, malfunctioning of the system, excessive system disturbance levels (e.g. large windshears in aircraft operation) can be represented by a deterministic process, as such unknown to the human observer but detected on the basis of a non-zero mean innovation sequence whose statistic is sufficient to make decisions (test hypotheses) when the system is completely observable.

In terms of classical sequential decision theory (Ref. 6) a so-called generalized likelihood ratio test can be formulated. The test amounts to the comparison of the probability of a non-zero mean with the probability of a zero mean innovation sequence assuming that the human operator makes a short-term estimate of the mean of the innovation sequence on the basis of the sample mean of m past observations ($\tilde{n}_k$).
It can be derived (Ref. 5) that the effect of each observation at stage k on the (log of the) likelihood ratio is given by

$$\Delta L_k = \tfrac{1}{2}\, \tilde{n}_k'\, N_k^{-1} \tilde{n}_k \qquad (1)$$

under the assumption that the sample mean $\tilde{n}_k$ is constant during m observations. The accumulating effect of each observation on the total (log of the) likelihood ratio is given by the recursive expression

$$L_k = L_{k-1} + \Delta L_k \qquad (2)$$

assuming that the innovation sequence is a white noise sequence (independent samples) which is exact in the normal mode of operation. The number

of observations, based on which the decision is made, is chosen such that ΔL_k is, on the average, not decreasing; in other words, using only confirming evidence to make the decision. Otherwise, the positive semi-definite elements ΔL_k would have an accumulating effect on the likelihood ratio.
When the likelihood ratio L_k (representing the total evidence of abnormal system operation) is equal to, or larger than, a decision threshold T, the decision is made that an abnormal condition has occurred. This decision threshold can conveniently be related to the accepted (or assumed) risk according to (Ref. 6) $T = (1-P_M)/P_F$, with P_M the miss probability (i.e. of no response to an abnormal condition) and P_F the false alarm probability.
Referring to references 5 and 6, an expression can be derived for the average number of samples (K) used to make the decision that the system is operating abnormally, which is, for a given sample rate, uniquely (linearly) related to the average detection time

$$K = 2P_F[P(H_o)+TP(H_1)]\ln T/E\{\overline{\tilde{n}'N^{-1}\tilde{n}}\} \qquad (3)$$

where $\overline{(.)}$ indicates the average over the ensemble and $E\{(.)\}$ is the average over the sequence. The equation provides a relationship between the average number of observations and the decision error probabilities (P_F and P_M), for a given innovation covariance N and the non-zero mean failure state sequence which is, however, a given task variable. Thus, the only human decision model parameters are the short-term average sample size m and the innovation covariance N which depends exclusively on the human observation noise covariance (V_y). If an optimal allocation of attention among the visual cues is assumed (f_i^*) utilizing the relationship

$$V_{y_i} = P_o\sigma^2_{y_i}/f_i^*, \qquad (4)$$

the only remaining decision model parameters are the short-term average sample size m and the overall level of attention P_o. The model output of main interest in this study is the average failure detection time corresponding to an assumed error probability for which the values of 0.05 and 0.01 has been chosen (Ref. 5).

Eq. (3) pertains to the situation that there is a given probability that an abnormal condition occurs ($P(H_1)$). This is used to predict the effect of the (prior) failure probability and the related false alarm probability on the detection time. The results will be presented in the section containing the model and experimental results.

EXPERIMENTAL PROGRAM

General

The experiment was conducted on the NLR moving base simulator. The flight simulator was configured to represent the linear equations of motion of a medium weight (29,000 kgf) twin engine jet transport. The task was to monitor a 3 degree automatic approach to touchdown beginning at a range of 5813 meter from the touchdown point (corresponding with an altitude of 1000 ft). The cloud base was set at 300 ft altitude. Furthermore a "moderate" turbulence level was selected for the experiment. The subjects (4 airline pilots) were supposed to respond to the detection of unacceptable approach situations by disengaging the autopilot. Subsequently, they were free to continue the approach manually, or to terminate the run by means of a "hold"-button. The moment of autopilot

disconnection determined the detection time. The subjects were instructed to conceive the task as a realistic approach task and to adopt the corresponding monitoring and decision strategy. Prior to the experimental sessions each subject was extensively trained for 6 to 8 hours to minimize learning effects.

Display information

The display information is shown in figure 1. The information was presented "head-up" to avoid visual transition effects. The visual cues available were: artificial horizon with heading reference, radio altitude (digital presentation in 10 ft steps), indicated airspeed (5 knots per dot), rate of descent (200 ft/min per dash), glideslope and localizer deviations (dots), and further a cross-bar flight director.

Abnormal conditions

Two types of unacceptable conditions were simulated in the experiment. The first type was an excessive longitudinal or lateral windshear which would ultimately result in unacceptable system deviations. The second type was a degraded "ILS"-signal (a second order disturbance superimposed upon the glideslope or localizer indication). The effect was a gradually increasing deviation of the pertinent display signal without effect on other display variables. All disturbances were initiated at 700 ft altitude (which was not known by the subjects).
Finally, two probabilities of failure occurrence, one of 0.8 and one of 0.2 were selected for the experiment. This condition has been chosen as an experimental factor to investigate the effect of the subjects failure expectancy upon the detection performance. Consequently, the validity could be investigated of making extrapolations from experimental results, obtained mostly for frequently occurring situations to rare events. The latter condition pertains to most real life situations and is a key factor in theoretical and experimental research in human decision making.

Experimental design

The subjects in the experiment were 4 experienced airline pilots. The formal experiment consisted of 9 sessions per subject, each of which involved a series of 20 simulated approaches. In 4 of the 9 sessions, the failure probability ($P(H_1)$) was set at 0.8; in the other 5 sessions this probability was set at 0.2. At the beginning of each session the subjects were informed about the relative failure probability for the runs of that session. More specifically, this probability was indicated to the subjects with the terms "relatively high" or "relatively low". The failures were randomly distributed over the subsequent runs so that a randomized block factorial design (with failure type, failure axis and failure probability as experimental factors) was achieved (Ref. 7) to allow a convenient analysis of variance of the results (detection times, and aircraft related parameters at the moment of detection). Consequently, 5 replications per subject per failure were obtained for the low probable failure runs, and 16 with respect to the high probability runs. The formal sessions were preceded by a training period of 6 to 8 hours per subject to minimize learning effects. Finally, pilot ratings and comments were also obtained during the formal sessions for a more complete evaluation of the decision performance.

RESULTS

Model results

The model analysis has been conducted prior to the experiment. The primary model output is the average detection time for a given decision error probability with respect to the different task conditions. The failure magnitudes were chosen in such a way that reasonable values for the detection times were obtained (between 10 and 25 seconds) to avoid experimental artifacts (on one hand the subjects should not realize that the failures were initiated always at a fixed altitude of 700 ft, and on the other hand the detection should take place before cloudbrake at 300 ft to exclude possible effects of additional visual information). More specifically, the longitudinal and lateral windshears chosen, were 8.8 and 5.0 knots per 100 ft, respectively, whereas the system failures selected for the glideslope and the localizer indicator were 0.006 dots/s^2. The corresponding model detection times* are presented in table 1 for the two failure probabilities (0.8 and 0.2) and two false alarm probabilities (0.05 and 0.01). The model predictions indicate only relatively small effects of failure probability and false alarm probability on the detection performance. The high failure probability condition generally yields an improvement in the detection times of less than 1 second in comparison with the low probability case; a decrease in the false alarm probability form0.05 to 0.01 will detoriate the detection performance with circa 1 to 3 seconds. Finally, the model predicts similar detection times for three of the four disturbances; for the fourth condition (the detection of the lateral windshear) the detection time is predicted to be substantially higher.

Experimental results; comparison with model predictions

The experimental detection times for all conditions and subjects are summarized in table 2 (mean values and standard errors of the estimate). The table shows that there is a relatively small intersubject variability for the system failure conditions whereas two different levels are apparent for the windshear conditions. In the following the experimental results for the subjects denoted by A, B and C, D will therefore be distinguished.

The experimental results are compared with the model predictions (both for $P_F = 0.05$ and for $P_F = 0.01$) in figure 2. As is shown, the experimental results of subjects A, B agree well with the model predictions for all experimental conditions. The failure detection times of subjects C, D are generally larger, especially for the windshear conditions. This can be attributed to a more conservative decision strategy than was assumed for the model analysis as reflected by smaller false alarm rates (0.00 and 0.01, while subjects A, B achieved false alarm rates of 0.02 and 0.03) corresponding with larger detection times. The figure indicates further that the two failure probability conditions did not result in a systematic effect upon the detection times (in accordance with the model predictions).

Finally, statistical analysis of the data has also been performed by means of an analysis of variance (a so-called "multiple classification

---

* Assuming for the short-term average sample size a value of 4 seconds, an overall level of attention of - 20 dB, and assuming that perceptual thresholds are negligible.

analysis"; Ref. 8) allowing more general statements about the significance and magnitude of the separate effects of the different experimental conditions. Figure 3 shows the results in terms of the detection times. For comparison the figure also presents the results of this analysis method applied to model "data". It is shown that there is a good relative agreement between the model predictions and the experimental results with respect to the independent experimental factors.
Concerning the experimental results, the only significance was found with respect to the relatively large effects of failure type and failure axis ($P < 0.001$). The small effect of failure probability was not significant ($P < 0.44$). This is an important experimental result (in accordance with the model predictions), which is based on the data of 20 low failure probability sessions (whereby each session included 4 failure runs and 16 "dummy" runs) and 16 sessions with a high failure probability (including 16 failure runs per session).
A further analysis of variance of the magnitudes of other dependent variables at the moments of detection (i.e. the display variables of figure 1) also indicated that the effect of the failure probability condition is not significant. So, in conclusion, the experimental results as well as the model analysis strongly indicate that the effect of failure probability on all dependent variables is negligible small, at least for the range of failure probabilities considered.

CONCLUDING REMARKS

A model analysis and an experimental investigation have been carried out of pilot's failure detection performance during various automatic approach conditions. The tasks considered comprised the detection of excessive longitudinal and lateral windshears, and of system failures on the glideslope and localizer indicators. Furthermore, two failure probability conditions were included to investigate the effect of the subject's failure expectancy on the detection performance.

Generally, a good agreement between the model predictions and the experimental results of two subjects was obtained. The failure detection times of the two other subjects were larger, especially for the windshear conditions, which could be attributed to a more conservative detection strategy.
Specifically, the effect of the various independent task variables on the average detection time was well predicted: the detection of the windshears considered required significantly more time than the detection of the system failures; the lateral failures resulted in significantly larger detection times than the longitudinal failures; finally, a variation in the failure probability from 0.8 to 0.2 did not result in a significant effect on the detection performance (neither in terms of the detection times, nor in terms of display deviations at the moment of detection).

The latter is an important result which supports the validity of extrapolating experimental results obtained for high failure probabilities to real life situations with relatively low failure probabilities. The model, whose predictive capability has been supported experimentally, will be useful to facilitate the extrapolation to rare events.

REFERENCES

1 Kleinman, D.L. and Baron, S.: Manned vehicle system analysis by means of modern control theory, NASA CR-1753, July 1971.

2 Baron, S. and Levison, W.H.: Display analysis with the optimal control model of the human operator, Human Factors, 1977, 19(5).

3 Wewerinke, P.H.: Performance and workload analysis of in-flight helicopter tasks, Paper presented at the 13th Annual Conference on Manual Control, MIT, Cambridge USA, 1977 (also NLR MP 77013 U, April 1977).

4 Wewerinke, P.H.: The effect of visual information on the manual approach and landing, Paper presented at the 16th Annual Conference on Manual Control, MIT, Cambridge USA, 1980 (also NLR MP 80019 U, 1980).

5 Wewerinke, P.H.: A model of the human decision maker observing a dynamic system, NLR TR 81062 U, March 1981.

6 Sage, A.P. and Melsa, J.L.: Estimation theory with applications to communications and control, McGraw-Hill, 1971.

7 Kirk, R.E.: Experimental design: procedures for the behavioural sciences, Brooks/Cole publ.cy. Inc., Belmont USA, 1968.

8 Nie, N.H. et al.: Statistical Package for the Social Sciences (SPSS) Second edition, McGraw-Hill book company, New York, USA, 1975.

TABLE 1

Model predictions* of the detection times (in seconds) for all experimental conditions

		FALSE ALARM PROBABILITY			
		$P_F = 0.05$		$P_F = 0.01$	
DISTURBANCE \ PRIOR PROB. $P(H_1)$		0.8	0.2	0.8	0.2
WIND-SHEAR	longitudinal	12.2	12.8	13.8	13.9
	lateral	21.2	22.4	24.4	24.7
SYSTEM FAILURE	glideslope	14.6	15.5	17.1	17.3
	localizer	14.5	15.5	17.2	17.5

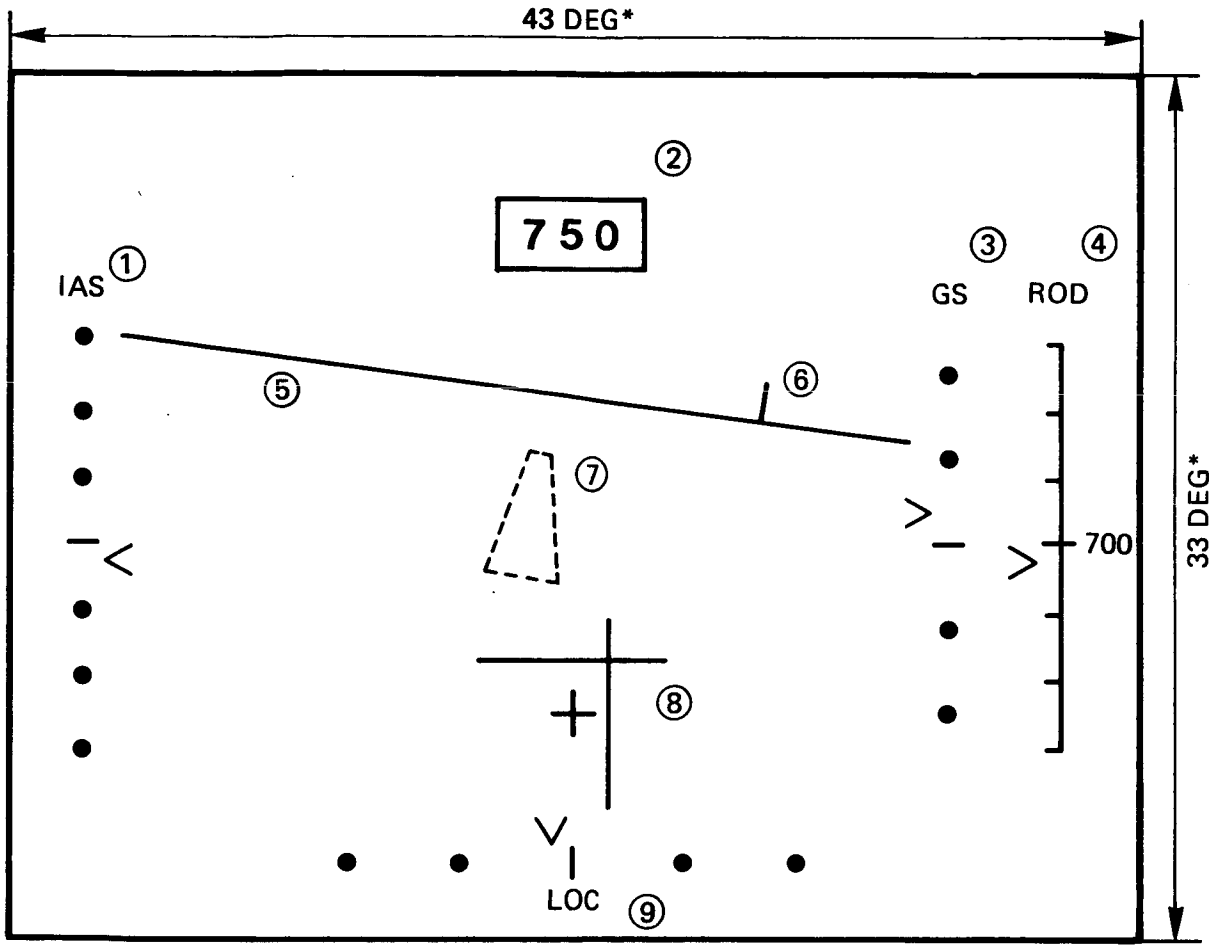
* based upon the assumption that the overall level of attention is -20 dB, the perceptual thresholds are negligible, and the short-term average sample size is 4 seconds.

TABLE 2

Average detection times in seconds for all experimental conditions and subjects

DISTURBANCE		SUBJ.	PRIOR PROBABILITY $P(H_1) = 0.8$ *	$P(H_1) = 0.2$ **
WINDSHEAR	longitudinal	A	18.4 (2.5)	13.7 (1.6)
		B	14.7 (1.4)	20.9 (3.8)
		C	25.2 (1.7)	34.2 (5.5)
		D	24.8 (2.3)	27.4 (3.0)
		Average	20.8 (4.8)	24.1 (8.5)
	lateral	A	22.1 (0.5)	20.8 (1.5)
		B	26.4 (2.4)	29.7 (1.8)
		C	29.6 (0.6)	26.1 (1.1)
		D	30.3 (1.0)	30.4 (1.8)
		Average	27.1 (3.5)	26.8 (4.1)
SYSTEM FAILURE	glideslope	A	17.1 (0.6)	19.6 (0.7)
		B	18.2 (0.7)	18.2 (1.9)
		C	19.6 (1.1)	19.0 (0.1)
		D	21.6 (0.7)	21.5 (1.2)
		Average	19.1 (1.9)	19.6 (1.7)
	localizer	A	15.9 (0.5)	16.8 (0.4)
		B	18.7 (0.4)	17.4 (0.4)
		C	18.1 (0.3)	18.2 (0.3)
		D	18.4 (0.3)	17.7 (0.4)
		Average	17.8 (1.2)	17.5 (0.4)

* based upon 16 replications per subject
** based upon 5 replications per subject
(.) standard error of estimate



① INDICATED AIRSPEED	(5kts/dot)	⑥ HEADING MARKER
② RADIO ALTITUDE	(10ft/step)	⑦ RUNWAY (h<300ft)
③ GLIDESLOPE INDICATOR	(dots)	⑧ FLIGHT DIRECTOR
④ RATE OF DESCENT	(200fpm/dash)	⑨ LOCALIZER (dots)
⑤ HORIZON		* VISUAL ARC

Fig. 1 Visual information presented on the front windshield

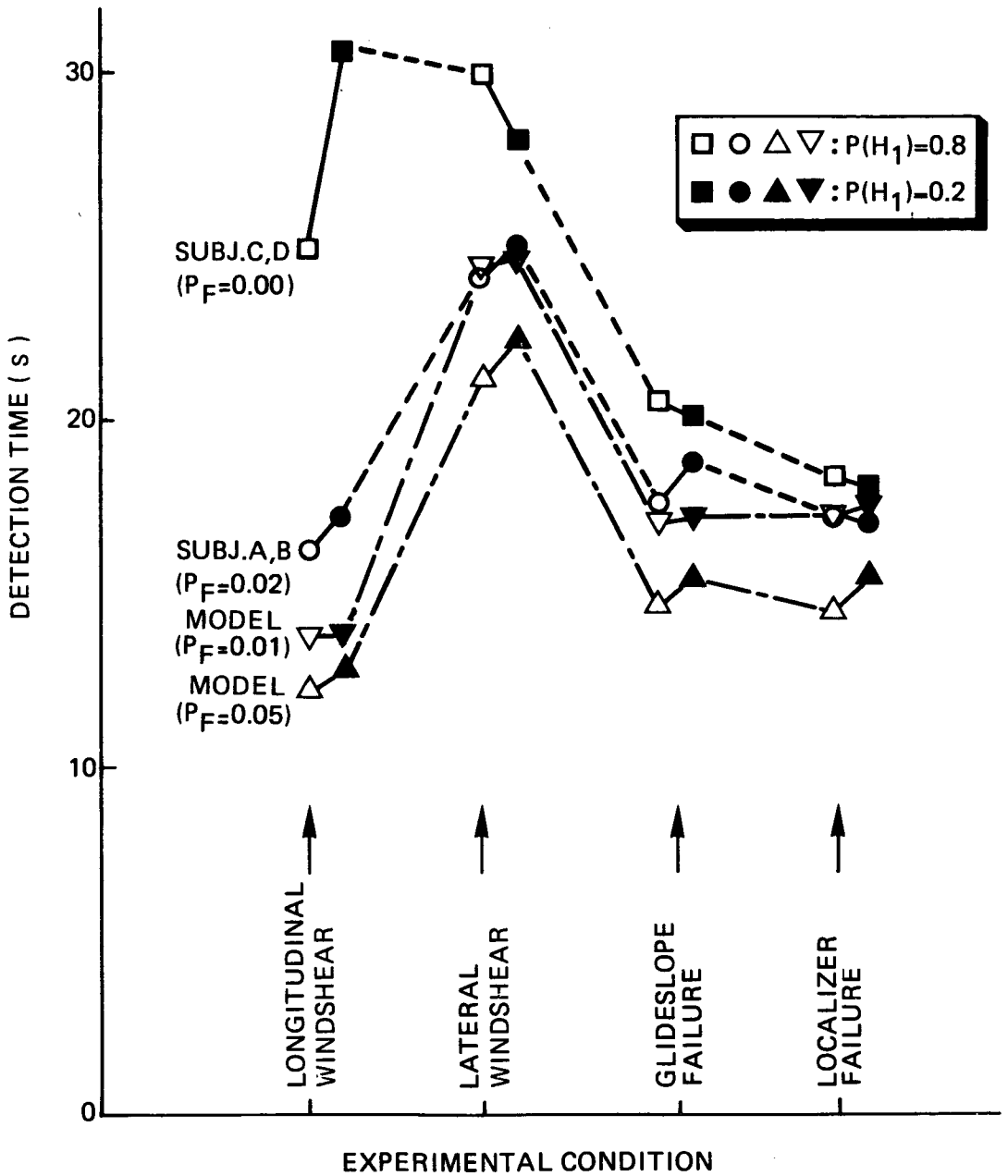


Fig. 2 Comparison of model predictions with experimental results

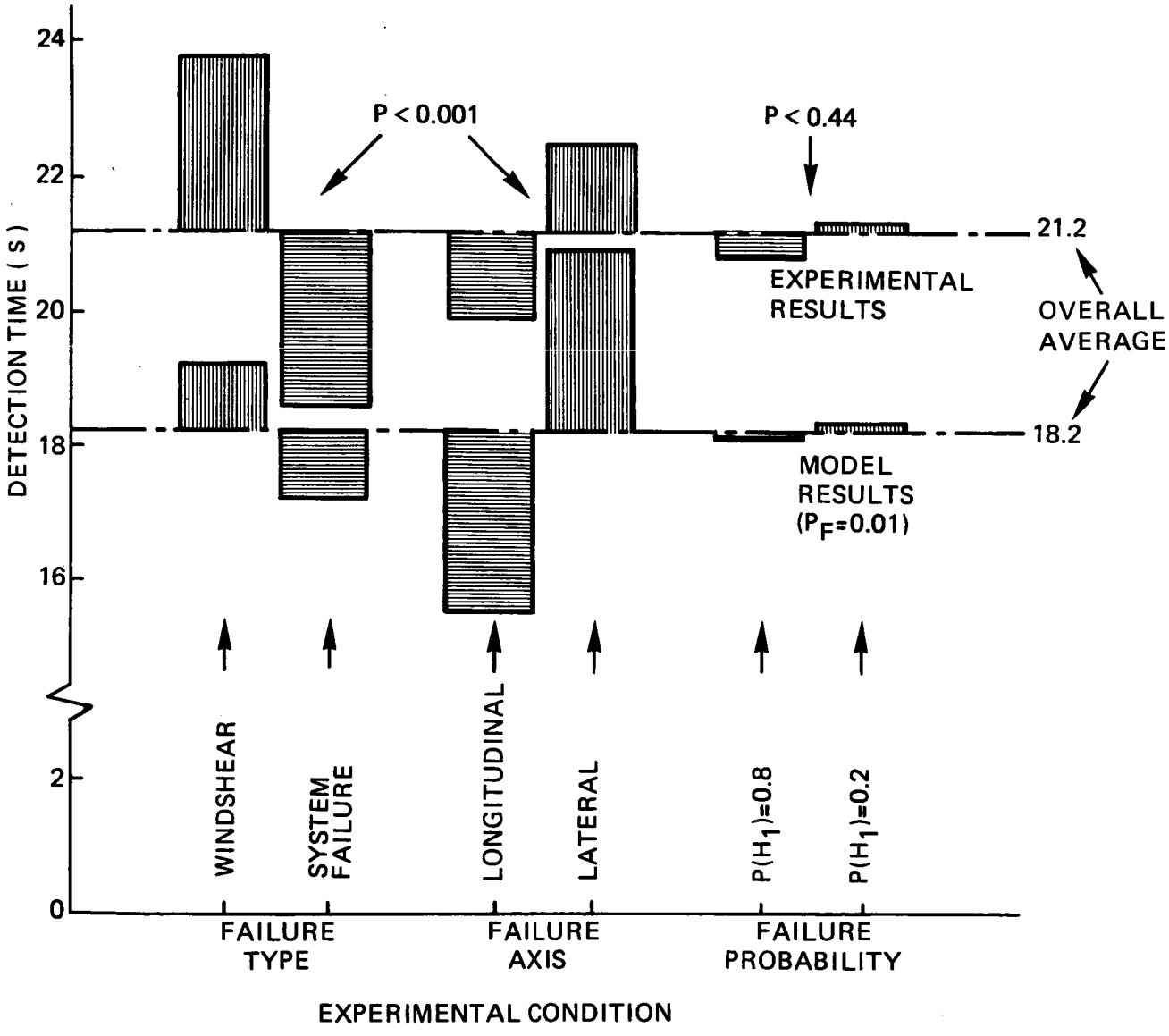


Fig. 3 Comparison of the separate effects of various experimental variables reflected by the model results and the experimental results pooled for all subjects

EVALUATION OF PREDICTIVE, HISTORIC AND STATUS DISPLAYS IN SUPERVISORY CONTROL

R.J. van der Veldt*
Man-Machine Systems Group
Laboratory for Measurement and Control
Department of Mechanical Engineering
Delft University of Technology
Delft
The Netherlands

SUMMARY

A brief overview is given of the current status of the project. This project deals with the evaluation of several sorts of information in a process control situation; especially predictive, historic and status displays will be considered. In this (informal) paper the proposed approach is discussed, and the preparations for the research are described. These preparations mainly consist of the development of two computer simulations of more or less realistic industrial processes.

INTRODUCTION

Due to the rapid developments in electronic devices, control rooms are more and more equipped with advanced computer-driven displays. The great advantage of these advanced techniques is the possibility of presenting the process-information to the operator in all kinds of formats. Moreover, it becomes possible to do all kinds of data-processing and thus showing the operator several *sorts* of information. One of the kinds of information that became possible, is predictive information.
What we try to do in the project is to evaluate the usefulness of predictive displays, trend/ or historical displays and statusdisplays, by making a comparison between these displays [Schneider, Van der Veldt, Stassen, 1982]. Actually, we are then dealing with the question: what information should we present to the operator, in what situations? In this project we mainly consider supervisory control tasks and process tuning tasks. In a parallel project we hope to start, before long, a study dealing with fault management tasks as well.

* This research is partially sponsored by the Foxboro Comp., Foxboro, Mass., U.S.A.

APPROACH

Presently, we have chosen for the following approach, which is twofold: on the one hand it is experimental, on the other hand we hope to develop some more theoretical background. What we try to do in the experiments, is investigating how several process-parameters and display-parameters influence the performance of the operator and of the whole plant. A process-parameter can be the complexity (whatever that may be); a display-parameter can be: with or without a predictive display.
We plan to run experiments with three different scales of processes:
- the first is a very small simulation of linear processes of the 8th order at most,
- the second is a mixed fidelity simulation of a utility plant,
- and the third is a high-fidelity simulation of a nuclear power plant. Experiments with this very realistic simulation will be carried out within an other project.

As for the theoretical approach, we are still in a very preliminary and exploratory phase. One thing we are thinking of, is to derive theoretically the minimal amount of process-information that is required in a given situation. Next, this amount can be compared with the information needed by the human operator to perform his task adequately. This involves some kind of modelling, or, at least, developing a normative model, or reference. This approach, however, is in a too early stage to go into this much further at this moment.

PREPARATIONS

The experimental studies will be carried out, as mentioned above, with computer-simulated processes of several sizes. Experiments consist of the study of subjects' behaviour and performance, while controlling a computer-simulated process in the laboratory. For this purpose two computer simulations are developed in our own lab; the hifi-simulation of a nuclear power plant is built in Halden (Norway). A short description will now be given of the two simulations in Delft: the small simulation and the utilityplant-simulation.

Small simulation

Recently, we have developed a software package with which every possible linear system, of up to the 8th order, can be simulated. Of such a process we can choose up to 8 inputs that can be changed and adjusted by the subject during the experiment. We can also choose up to 8 outputs that can be displayed. Besides, also disturbances acting on the process can be simulated, namely 8 independent noise-signals at most (Fig. 1).
The processes that can be simulated with this package can, in fact, be represented as:

$$x_{k+1} = A\,x_k + B\,u_k + G\,w_k$$

$$y_{k+1} = C\,x_{k+1}$$

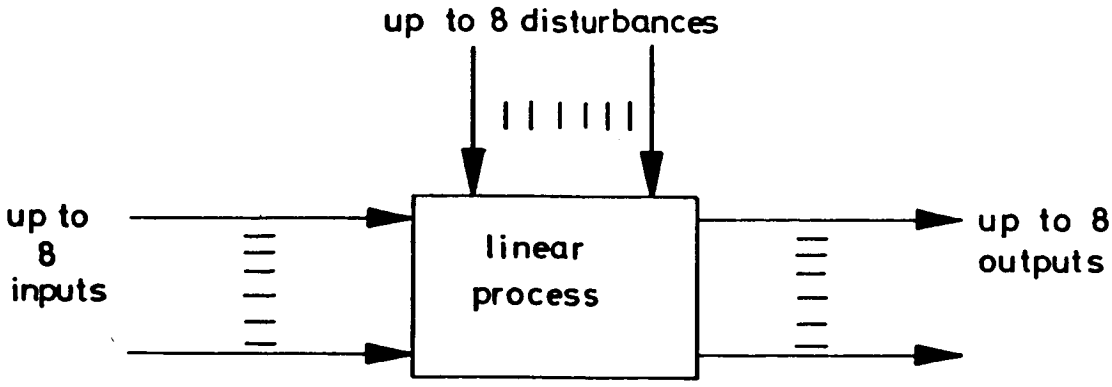


Fig. 1. Schematic representation of the small simulation package.

The software-package also contains provisions like data storage, and a man-machine interface with a semi-graphical CRT.
In fact, this flexibel set-up enables us to do experiments with all kinds of different experimental conditions. In this way we hope to learn something systematically of the human supervisory behaviour and the use of several displays.

Utilityplant simulation

Next to these relatively simple, elementary simulations, we developed a mixed fidelity simulation of a utilityplant [Campbell, Shirley, 1979; Schneider, Van der Veldt, Stassen, 1982]. This plant produces 5 different forms of energy, namely: high pressure steam, intermediate pressure steam, low pressure steam, electricity and compressed air.
The process consists of 4 boilers, producing steam; 1 condensing turbine, producing electricity; 2 backpressure turbines, producing electricity and intermediate press. steam; and 4 aircompressors, producing compressed air. There are also three headers for the produced steam (Fig. 2).
For this simulation, which has some more realistic features than the small simulation, we developed a predictive display. The configuration is as follows: the Fox-computer simulates the utilityplant, and the status-information about the plant is shown to the subject on the operator's

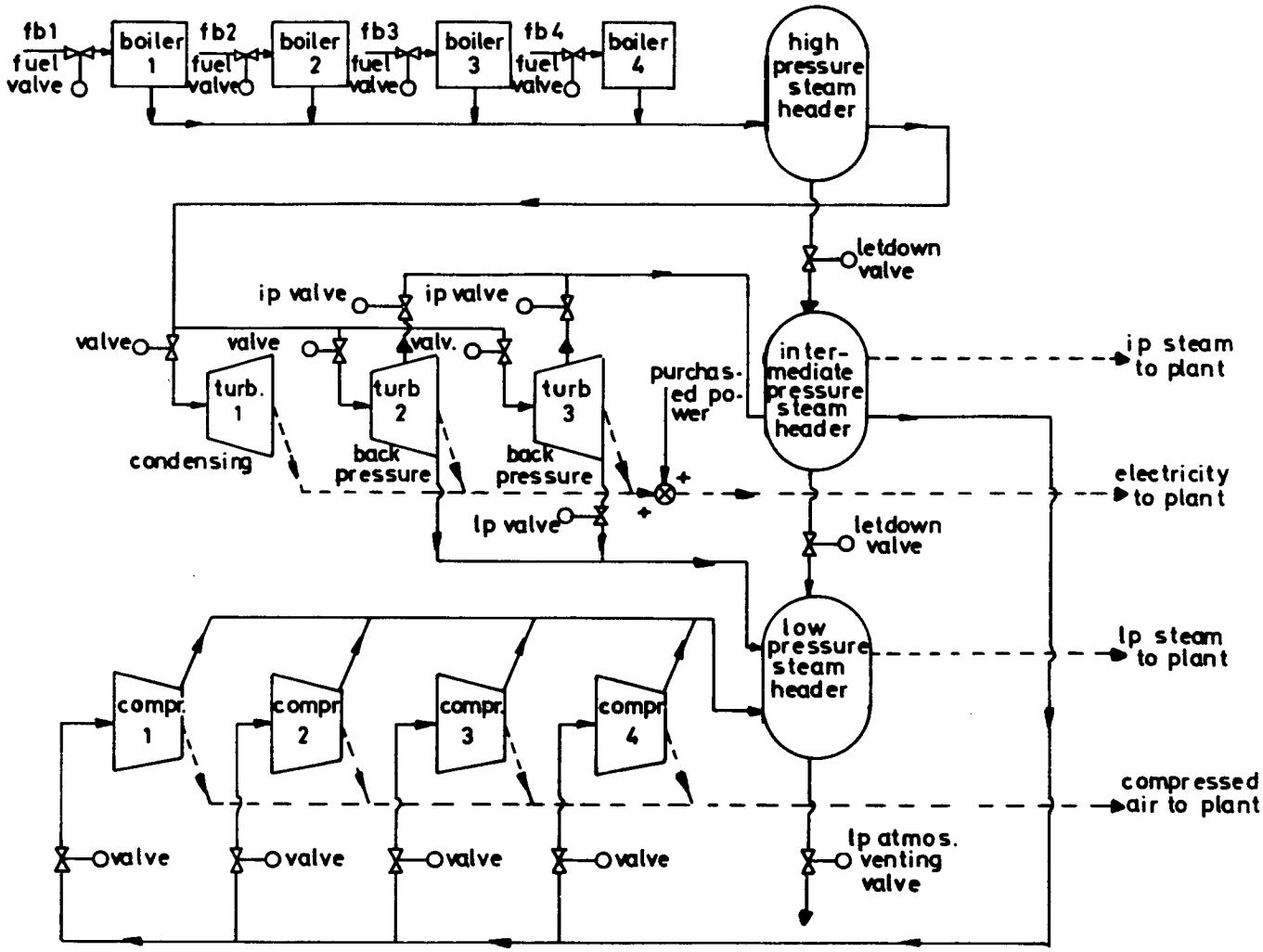


Fig. 2. Simplified Process & Instrumentation diagram of the simulated utilityplant.

terminal. Predictions are calculated with an analogue computer. This set-up gives a reasonable result. However, with the analogue computer we can only simulate a part of the entire utilityplant, and presently we are working on predictions calculated with a digital computer.

Before long the experimental work with the small simulation and with the utilityplant will be started. Some preliminary experiments have already shown the usefulness of both simulations. We expect to obtain, therefore, useful results from the experimental set-up and from the approach that we will follow.

LITERATURE

Campbell, B.D.; R.S. Shirley
An industrial utilities model for research in man-process control. In: Proc. of the 15th Annual Conf. on Manual Control, Wright State Univ., 1979.

Schneider, H.W.; R.J. van der Veldt; H.G. Stassen
The role of overview displays in human supervisory control. In: Proc. of the 2nd European Annual Conf. on Human Dec. Making and Manual Control, Bonn, 1982.

HIGH VALUE ROLES FOR MAN-IN-SPACE (MIS)

By

O. Herbert Lindquist

Honeywell Inc.
Systems and Research Center
2600 Ridgway Parkway, PO Box 312
Minneapolis, Minnesota 55440

Paper presented:

19th Annual Conference on Manual Control
May 23-25, 1983
M.I.T.
Cambridge, Massachusetts 02238

HIGH VALUE ROLES FOR MAN-IN-SPACE

INTRODUCTION

This paper summarizes two studies for establishing high value roles for man-in-space (MIS). The motivation for the study is a growing concern about MIS and the emergence of an issue: What value would man add to space missions if he were located in space? Various national leaders have voiced divergent opinions on this subject.

Current USA operational space capability is limited to unmanned satellites with ground support. The high costs associated with maintaining an effective MIS have prohibited serious systems developments. New needs and threats, and newly available technology capability, necessitate the re-evaluation of whether there is a driving reason to locate man in space. A systems approach to resolving this problem and initiating concept developments is needed.

The objective of Honeywell's study is to identify effective manned space systems and guide the application of capabilities in realizing these systems.

STATEMENT OF THE ISSUE

The value of having MIS is a major space systems issue. The issue has been characterized by two polarizations of thought: the "automatics" who would have man located in a ground facility, and the "white scarves" who advocate MIS (Figure 1). The automatics advocate unmanned satellites, because they believe that automation and advanced technology can be used to alleviate the need for an "expensive" manned space presence. The large expense is associated with placing and maintaining man in the hostile space environment. The white scarves have usually alluded to a capability associated with man's skills, decision-making ability, and motivation.

A key point is the recognition that man <u>is</u> present somewhere in the system. The question then becomes: <u>W</u>here best to locate him/her so as to maximize mission effectiveness or reduce system cost?

The major technology advancements and new military threats illustrated in Figure 1 call for the re-examination of the MIS issue in some organized, quantitative manner. The amount of national resources involved and the negative consequences of a mistake in this decision process make it a "high stakes" issue.

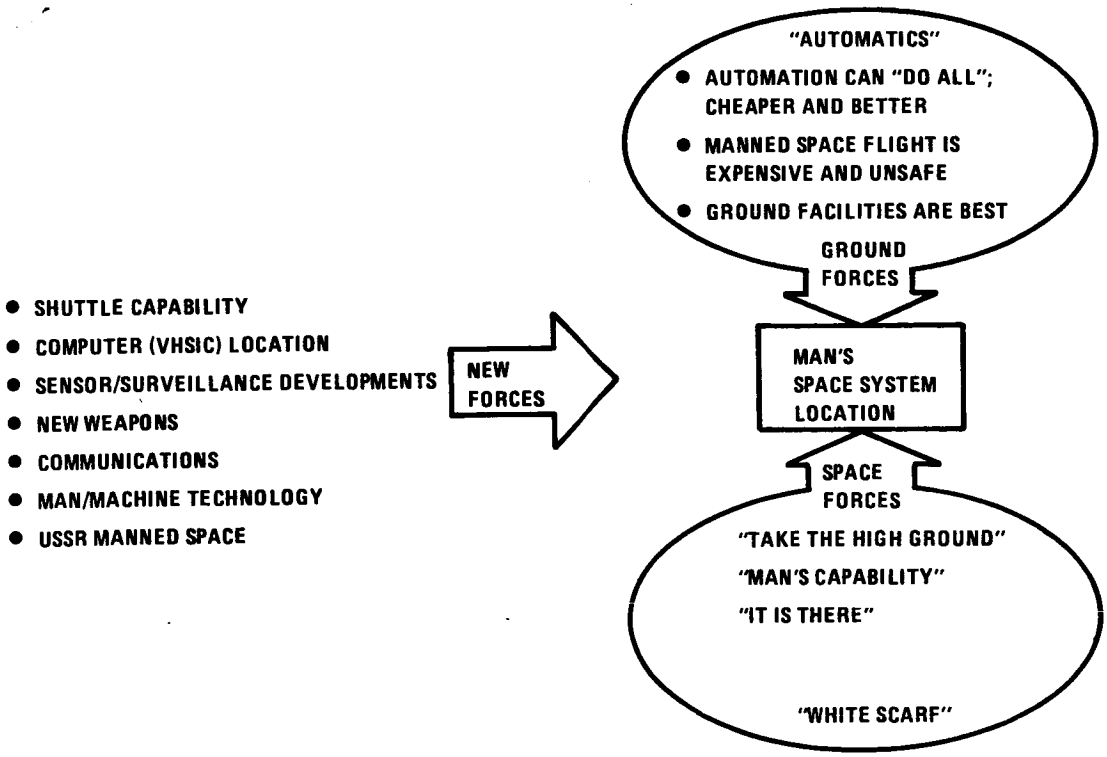


Figure 1. Systems Issue: Where Man?

Problem in Resolving the Issue

Biologist Garrett Hardin[1] has identified a tendency in approaching problems in ecology that has counterparts in the MIS issue. He states that most scientists do not accept the idea of a system. Physicists, the top of the pecking order for scientists, set the habit fashion for all science--and their habits do not include systematic thinking. Narrowly analytical thinking is the mode. Engineering, a child of physics, seldom departs from this pattern. In his summary of this type of thinking, Hardin states that nonsystematic, narrowly scientific (which he characterizes as Single Vision Newtonian) approaches will work to identify separate symptoms of a process, which are then resolved by "buying and building" hardware. Hardin effectively pleads for a systems approach to deal with ecological issues in a total systems context.

In my opinion, this situation also exists in determining the value of locating MIS. The problem exists in three dimensions, related to white scarves, automatics, and decision-makers.

---

(1) Exploring New Ethics for Survival: The Voyage of the Spaceship Beagle. Garrett Hardin, Viking Press, 1968.

In the first dimension, white scarf advocates point out man's general capability and the value he could add, if given the opportunity. Some imply that given the opportunity, man will prove his value, i.e., that because we do not know his value, he should be there. There are, of course, some white scarves who do so because they will provide the hardware needed to have MIS. In any case, the value man adds is an increment which does not seem to have a driving reason for existence.

In the second dimension, automation advocates are strong in their ability to dissect the issue into technology elements that are amenable to separate studies. Given resolution of the various technology issues, via hardware solutions, it will be possible to build the space station. The issues related to man tend toward providing more efficient support systems and better interface hardware. The issue of man's function is currently being delegated to techniques, artificial intelligence for example, of doing more of his functions via software and automation. The automatic advocates request the definition of what one function man is uniquely suited for, so that a software-hardware combination can be postulated which is a "better" solution. The realities of technology limitations and practical systems approaches are not addressed.

The third dimension of the problem resides with the technical decision-makers who formulate the total space program. They prefer to deal with technology readiness, hardware, and program planning realities. Their experience is based upon experience, i.e., if we can make an airplane fly, men will find their role, or if we can orbit a man, he will make significant contributions. The technical decision-makers will partition the problem into manageable technology areas and seek hardware solutions within a given timeframe. The issue of the role of man is one of these problem areas, but it is not viewed in a total context.

Therefore, the issue of the value of man is not being addressed in a systems context.

A SYSTEMS APPROACH

An alternate approach has been formulated for determining the role of man; treat the problem as a systems issue (Figure 2). This approach expresses the mission problem in a broad context, defines various system approaches for meeting the mission requirements, and evaluates each of these approaches in a total context. Mission requirements are to be met both by placing man at an appropriate location within the system and by providing him/her with the appropriate tools to meet performance requirements in a cost-effective manner. Potential system architectures are postulated to provide a means for identifying

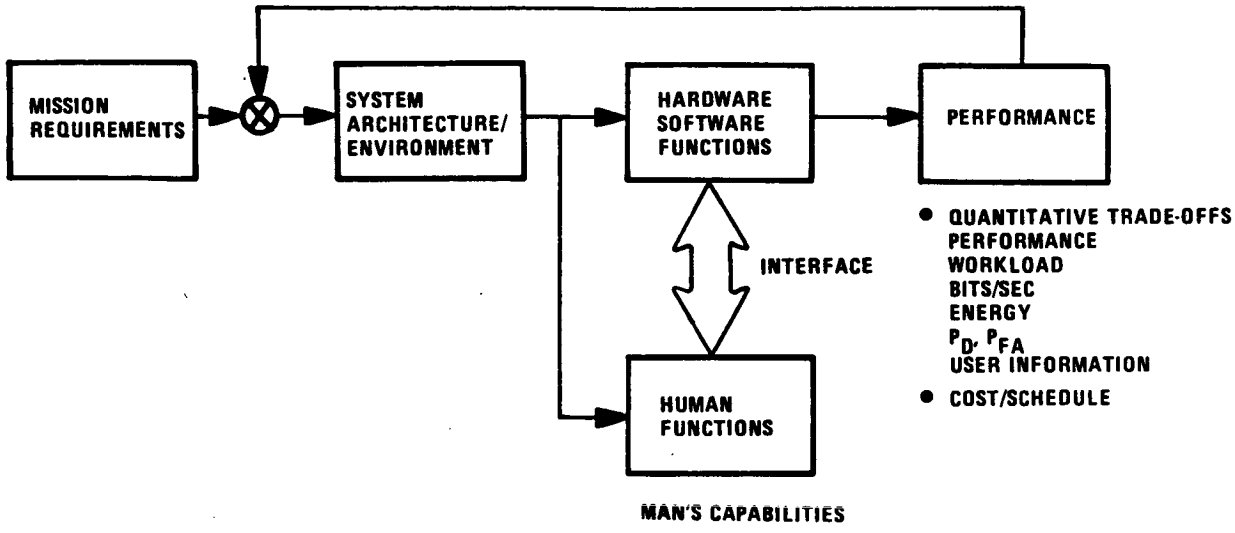


Figure 2. Systems Setting for Roles and Contributions of Man

candidate locations for man in the system. Man-machine specialists will participate in the initial search for system alternatives. Technology specialists will participate in formulating hardware/software system elements to ensure taking advantage of available technology within the realities of schedule and cost constraints. In this way candidate roles for man can be postulated for trade-off studies. The man-machine specialists will define roles that employ the higher order capabilities of man, in addition to his/her manual capabilities. The process of evaluating man's contribution will proceed on a systems performance basis with respect to mission needs. The value of MIS will be evaluated by comparing the system effectiveness of each candidate role for man within the mission under study, i.e., man-in-space and man-on-earth.

Identification of Mission

During 1982 Honeywell examined possible missions by this systems approach. To identify those missions, a review of civilian and military scenarios was made to identify priority areas in which man could be seen in high value roles. Generic classes of space systems were then established, including current and postulated systems, along with a general timeframe.

A multidiscipline study team reviewed this information and identified several of the generic mission areas to determine best candidates for study. This selection has been confirmed in independent USAF studies. Two mission areas were selected for more detailed study:

1. Low-Earth-Orbit Tactical Surveillance (LTS)
2. Transport/Maintenance of Space Assets

These two missions were studied in 1982, and the results are summarized below.

LOW-EARTH-ORBIT TACTICAL SURVEILLANCE MISSION STUDY

As previously stated, the issue of the value of MIS can only be resolved in a systems context. Honeywell selected the specific mission related to the support of Army needs in western Europe. The availability of several years of intensive mission analysis and a detailed scenario simulation provided a comprehensive data base for the elements, tactics, and timelines associated with Army user information needs.

The study was divided into three mutually interactive tasks. We first established user mission requirements based on Honeywell's previous mission analysis. Sensors were conceptually specified to meet the requirements of the user and to provide information under all weather conditions. The total system was then reviewed in terms of processor throughputs, information extraction, and timely communication to the user.

Conclusions

Two major conclusions can be drawn from the study of MIS for low-earth-orbit (LEO) tactical surveillance:

1. Man's Role--It is clear that the best use of man is as an intelligent manager interacting with an extensive sensor/processor/communication suite of equipment. Man is thus the "information manager" in an information management system. As such, he or she controls the disposition of sensors, the acquisition and correlation of data, the extraction of useful information, and the communication of such to the end user. The operator uses the highest human qualities of judgment, decision-making, and the formation and testing of hypotheses to guide the equipment to a timely solution of the user's problem.

2. Man's Location (The Bandwidth Problem)--The second major conclusion evolving from this study is the

realization that a vast amount of data needs to be quickly compressed in order to meet the tactical commander's needs. This is the key issue driving man's location in the space surveillance system.

The data collected by the sensors cannot, even with the data compression techniques envisioned, be transmitted to the ground in a timely manner in a realistic environment.

The key element in this conclusion is the need for providing the users with timely, useful information. This requires near real-time processing of multisensor information. It is not possible to provide this information when the data rates associated with our selected sensors are processed so as to transmit in the expected environment. This means that the human operator, the heart of the interpretation process, should reside in space where access to the preliminary processed data is secure.

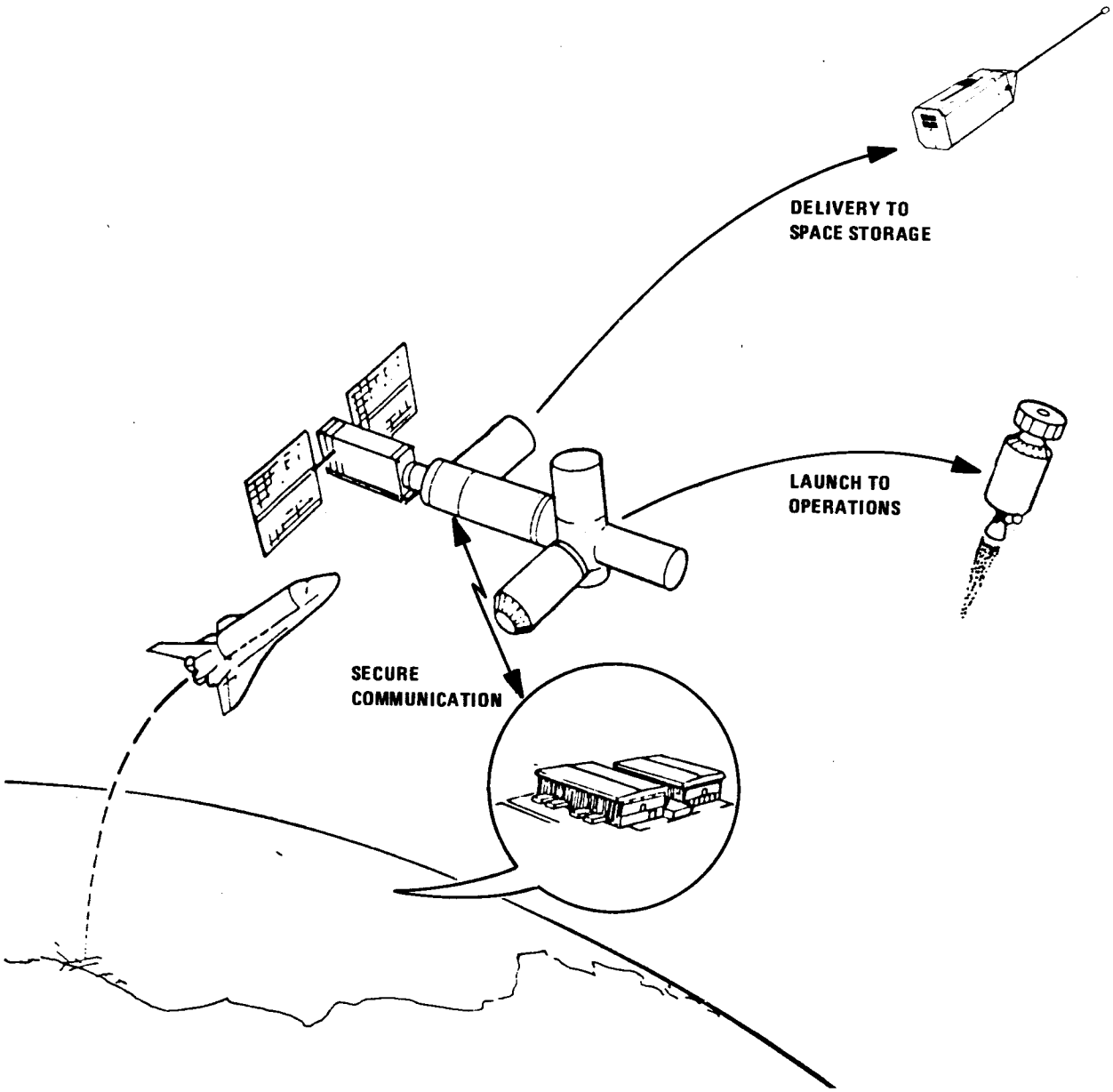
An evaluation of the mission using man in a ground role indicates that the mission could not be accomplished within user requirements, because of communication bandwidth problems and the lack of proven automated concepts.

This analysis has reached the conclusion that MIS equipped with significant computer power is a possible solution to space surveillance for the chosen tactical missions, whereas an unmanned satellite system has serious technical deficiencies.

TRANSPORT/MAINTENANCE OF SPACE ASSETS MISSION STUDY

The space shuttle transport system is planned to deliver satellites to parking orbit for subsequent deployment. A number of problems that are inherent in this procedure can be alleviated by having a manned space station available to supplement the shuttle orbiter. This study was performed jointly with Rockwell International, Inc.

The overview of this mission is shown in Figure 3. The space assets are delivered to the space station via the orbiter. The station receives the payloads and provides services to either launch directly to deployment orbits or to store in protected storage. In this process the station serves as a coordinating function by check-outs and by maintaining assets, and by providing storage, common stores, and assembly. The primary issue relates to determining the added value of having a manned space station.



STATION IS DEPOT/LAUNCH SITE

- CHECK–OUT
- MAINTENANCE
- GARAGE
- LAUNCH
- MAINTAINS COMMON STORES

Figure 3. Logistics and Servicing Mission

Conclusions

The study postulated a space station capability that could provide the function of transporting and maintaining space assets. If this could be implemented, the following benefits could be obtained from a manned space station satellite storage-launch capability:

- Improved system reliability by repair of launch problems and space "burn-in" capability, replacing fragile ground launch systems, and improved communication links.
- Improved system availability by servicing and minimizing repair time, saving time in reconstitution from orbit.
- Lower costs from peacetime launch to parking orbit by shuttle transportation system and reduction in inventory due to repair capability (shuttle launch efficiency).

The result of this study was to identify where man could countribute to the space station having an in-space ability to maintain and reconstitute space assets. An alternate ground-based facility would be more costly and less effective.

The study again indicated the value of having man as an information manager of a major computation facility.

SUMMARY

Two mission studies have identified mission-critical contributions for MIS, which should be studied as alternates to ground-based systems. The two system studies are separately documented in more detail in Honeywell reports currently in publication. A program plan to approach identified issues is also found in these studies. Each study identified an effective use of MIS that would be difficult to duplicate if man's role in the system were confined to the ground. Honeywell's system approach to man-machine system design and the availability of highly applicable technology have progressed to meaningful results and have identified major areas of application.

QUALITATIVE SIMULATION OF COMPLEX DYNAMIC SYSTEMS: AN APPLICATION TO A MARINE POWERPLANT UNDER SUPERVISORY CONTROL

T. Govindaraj

Center for Man-Machine Systems Research
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332
(404) 894-3873

ABSTRACT

An investigation of the possibility and effectiveness of qualitative simulation for training marine engineering personnel to detect, diagnose, and compensate for failures is discussed. Low, moderate, and full fidelity simulators will be used in the training. The moderate fidelity simulator will use qualitative simulation of system dynamics. Details of the qualitative simulation are discussed including a review of several modeling methodologies. Low, moderate, and full fidelity simulators are also described, and an outline of research in progress is provided.

INTRODUCTION

Complex dynamic systems such as marine powerplants are operated in supervisory control mode since most of the functions are automated. Appropriate power levels are maintained by the automatic control systems as long as there are no problems during operation. When a problem arises, the supervisory crew must be able to detect problems quickly, identify the causes and correct the problems. Where full correction is not possible, compensation for problems is desired so that the system can be safely managed until the problems can be fully identified and corrected. While formal procedures can be developed for dealing with many types of possible situation, all possible contingencies cannot be anticipated. Therefore the operation of a large, interconnected dynamic system such as a marine powerplant sometimes depends on the problem solving skills of the crew members.

Various means could be used to develop problem solving skills. One possibility is to provide the operators with on-the-job training via a considerable period of time on the equipment they are expected to use. This is rather expensive, especially for training marine engineering personnel. First, the ships are expensive to operate, and it is therefore difficult to preempt other productive operations for the purpose of training. Second, and perhaps equally important, the real system cannot be made to function in a degraded mode intentionally. This would endanger the ship's operation, as well as the crew's safety. It is obvious that alternative methods are preferable. One possible approach is to let the trainees be apprentices on a real ship, and let them observe how an expert solves problems. This is a very inefficient process, however. Further, faults seldom occur in most automated systems, and cannot be simulated in

such an environment. Hence, training with the real system tends to be both inefficient, and limited in effectiveness. One of the best alternatives for training is through the use of training simulators.

Simulators can be designed at various levels of complexity and realism relative to the systems they represent. The degree of realism and closeness with which the simulators resemble the actual systems can be formalized in terms of the concept of fidelity of simulation. Typically, the higher the fidelity of the simulator, the higher will be the cost. However, for training purposes, it may not be necessary to have a simulator with the "highest" fidelity possible in all aspects. Depending on the skill requirements, certain aspects of the simulation can be provided with high fidelity, with others at low or moderate fidelity levels. In fact, enhancing certain functions, and approximating some others may be useful for training.

The primary concern of the research described here is the design of training simulators to be used to teach the general problem solving skills associated with failure management. The long term goal is to design simulators to enable transferability of training at relatively modest costs. To this end, low and moderate fidelity simulators are being developed for a marine powerplant to train marine engineers in the detection, diagnosis, and compensation of system failures.

SIMULATOR FIDELITY ISSUES

Background

Research issues in simulator fidelity were the topic of a recent workshop in which researchers in the area of training participated [1]. Various simulator and training techniques in use were discussed, and fidelity issues were debated. Baum [2] presented hypothetical relationships between "physical" and "functional" fidelity versus the degree of simulator effectiveness. Even though these relationships have been used for a number of years, they appear to lack a firm experimental or theoretical basis.

A good discussion of simulator fidelity and its relevance for training is discussed by Montague [3]. Examples are provided to show that physical fidelity ("physical isomorphism") is not always necessary, and at times counterproductive for training purposes. The inability to specify fidelity in terms of actual training requirements is attributed to the lack of good mental models. Mental models such as those developed by researchers in cognitive science are discussed at some length. How these could possibly be used is described. Montague recommends that specifications of knowledge and experience required to attain specific performance levels be used to define training task fidelity.

Orlansky [4] maintains that fidelity is only one of many features that influence the effectiveness of a training device. He suggests that the effectiveness of a given fidelity level be measured in terms of the performance on the job rather than at school. While fidelity itself is not explicitly defined, the possible relationship of different amounts of fidelity to training effectiveness is discussed.

As the discussion above shows, no consensus seems to exist as to what constitutes fidelity or how it can be measured. One possibility is to characterize fidelity in terms of training effectiveness and transfer of training. This would reduce the concept of fidelity to be mainly an intervening variable which need not be defined rigorously.

Transfer of training and measurement issues are discussed in papers by Valverde [5], and Blaiwes and colleagues [6]. However, these measurement techniques do not appear to provide a means of measuring simulator fidelity. Adams [7] presents detailed arguments to point out the difficulty of evaluating training devices including simulators. Measuring the transfer of training, as well as the rating methods in which persons experienced on the actual equipment rate the simulator, are shown to have problems. He suggests that these evaluation procedures are unnecessary if the simulators are designed on the basis of scientific principles. An example of such a principle is that human learning is dependent on knowledge of results. Designing simulators on a firm scientific basis would bypass the question of simulator fidelity altogether, since it ceases to be an issue any longer.

Dimensions of fidelity

The issue of defining fidelity and studying its implications can perhaps be clarified if one begins with the premise that fidelity is a multi-dimensional concept. This allows a partitioning of the problem of defining fidelity into more manageable subproblems, from both a conceptual and measurement perspective. This approach is advocated in this paper.

Traditionally fidelity in simulators was synonymous with the physical similarity of the simulator to the actual system. A "good" simulator looked, and felt like the real equipment it was meant to simulate. This implies that the simulator had, at the least, the same number and types of displays and controls as the real system. Physical fidelity is concerned with the variables that are presented, the form in which they are presented (pictorial, verbal, etc.), and in how much detail they are given. If this requirement is only partially satisfied, the simulator lacks full physical fidelity. Another important consideration is the environment. For instance, noise, vibration, and thermal conditions similar to the real life environment should be provided where high physical fidelity is a requirement.

These requirements imply that simulators with high physical fidelity are usually rather expensive. However, the need for this high level of face-validity is often difficult to demonstrate. Even though a good correspondence between the simulator and the actual system provides a realistic environment with little deviation from the actual work environment, its contributions may not be clear. Possibly, a less detailed representation that provides only a limited number of essential controls and displays could provide an adequate training device. (Montague [3] provides some evidence in support of this.) This would imply that low or moderate physical fidelity levels may sometimes be acceptable.

Structural fidelity refers to the nature of relationships between various subsystems that make up the system. Feedback and feedforward relationships, and coupling through system states, inputs, and outputs are

relevant. Hierarchical relationships including the different levels of abstraction, and the details within a given level are important factors to be considered. Structural fidelity also deals with the aggregation of the detailed system elements and the levels of abstraction required to represent the system.

Dynamic fidelity refers primarily to the evolution of the system states over time, and their presentation to the operator. Dynamic fidelity is affected by the state representation and the level of approximation of the relationships among state variables as well as provision of the ability to control the system, and the feasibility for system reconfiguration. Dynamic fidelity issues are the primary focus of this paper. Other issues are included for the purpose of placing the problem in the proper perspective.

It is worthwhile noting that the three dimensions of fidelity defined here (i.e., physical, structural, and dynamic) offer a rather different perspective than the more traditional notion of engineering versus psychological fidelity. Basically, the point of view advocated in this paper is that psychological fidelity is multidimensional in and of itself, and that engineering fidelity is a related but different concept, one that is not pursued in this paper.

Measuring fidelity

Physical fidelity could possibly be quantified by using some simple pattern matching techniques. Controls and displays that form the man-machine interface could be compared at the simulator and at the actual system. Number, type, relative distances and configurations, are some of the primary factors that could be used in the determination of similarity measures. Techniques from syntactic pattern recognition could be useful [8,9].

Graph theoretic methods may be useful for measuring structural fidelity [10,11]. Suitable hierarchic representations may be necessary for this. Distance measures for nodes and arcs in the simulator and the system could be used to quantify the relationships between various parts of the systems. Since graph theoretic methods are widely used in pattern recognition [12], measures for structural and physical fidelity may be related and similar to each other.

Control theoretic measures could be useful for measuring dynamic fidelity. Among others, these could include measures of errors for important variables, response times, and the operational effort required. Analytical measures such as controllability and observability may also be useful. Methodologies from optimal control theory [13] and system identification [14] appear to be relevant for dynamic fidelity measures. Appropriate techniques are being explored.

An important consideration in defining and measuring fidelity is the impact of context and purpose. The environment and the purpose for which the simulator is to be used have a strong influence on fidelity. Also, fidelity of simulation appears to be closely determined by the context in which it is used. Therefore, context free measures of fidelity may be difficult to derive.

For example, even though it may seem desirable to have good "physical" fidelity in a simulator used for problem solving training, physical fidelity may not influence problem solving performance greatly; availability of the required information may be more crucial than the format and face-validity. In contrast, in training for another aspect of training in the same environment (e.g., rapid execution of emergency procedures), high physical fidelity may be absolutely essential.

Summary

Features that affect fidelity may be viewed from a man-centered and a machine-centered viewpoint. The former would include the various human abilities and limitations that affect fidelity, including sensory, perceptual, and cognitive aspects. The latter aspect would include factors such as the level of detail, and nature of structural representation, hierarchical relationships among the different subsystems, and accuracy of dynamical representation. The human and the machine aspects are interdependent and both have psychological implications. Therefore, these aspects should not be considered independently. The machine aspects greatly affect the cognitive factors. For example, if certain subsystems and their interconnections are not explicitly provided by the simulator, the trainee may have to compensate for these by inferring their states. Hence, a lack of these functions could reduce the fidelity.

Apart from the issues which are based mostly on system characteristics, another method of studying fidelity is in terms of its use and effectiveness. Since the purpose of simulators is in training, one might say that a simulator has high fidelity if the training on the simulator is transferred well to the actual system. The level to which the training can be transferred can be a measure of fidelity. Unfortunately, the effectiveness of transfer of training can be difficult to assess, particularly if one attempts to isolate long term effects independent of other variables.

APPROACHES TO DYNAMIC FIDELITY

The primary concerns of this paper are the dynamic fidelity issues when a large, interconnected system is simulated. Such a system has traditionally been simulated in high fidelity. This could possibly be due to the influence of engineers who were responsible for the design of the actual system as well as the simulator. Nothing short of the complete dynamics was considered acceptable for the simulator. Also, since manual control was still prevalent in most systems, physical motions including noise and vibration were provided in the simulation. However, these requirements may be unnecessary for problem solving in supervisory control where information such as time histories of primary variables may be sufficient. It appears that concepts from two rather different domains could prove useful for modeling the dynamics for the simulators. One is from control theory, based on aggregation, decoupling, and order reduction techniques. The other is the use of mental models in cognitive science. The latter has been proposed in artificial intelligence research for representing systems and devices. Details of these two approaches are discussed below.

Aggregation, decoupling, and order reduction

In general, various components of a large scale system interact with each other via inputs, states, and outputs [15]. When two systems in series are connected to each other only via inputs and outputs, the outputs of the predecessor become the inputs to the successor. The individual systems in such a case are decoupled. Otherwise the systems are coupled, whether they are in series or in parallel, since the states of one system are also affected by the other system. This coupling makes it impossible to study the different systems separately. Differential equations of fairly high order are required to represent coupled systems. Large matrices required for storage and manipulation of system equations make the computations slow and difficult. These considerations also impose constraints on the size of the system that can be modeled on a computer. In addition, high order coupled systems are rather difficult to visualize for modeling and manual control. Hence simpler representations are desirable.

Complexity of a large scale system can be reduced if the system equations are approximated through aggregation [16,17]. Components are aggregated so that the total number of subsystems is reduced and the overall structure is simplified. Minor components are lumped together to obtain a smaller number of possibly loosely coupled subsystems. After aggregation, the individual subsystems are coupled primarily through inputs and outputs. Coupling through the states is weakened or eliminated entirely.

In weakly coupled systems, individual subsystems can be decoupled so that they can be studied independently. In physical systems, the decoupling can be done so that the subsystems represent components which are only loosely coupled with the rest of the system. The system representation is thus simplified, and the computational load is reduced.

Order reduction methods can be used to further simplify high order decoupled systems [18,19,20]. In reduced order models of dynamic systems, the order of the differential equations is reduced so that only dominant modes are retained. For example, in systems where fast response is important, relatively slow modes are simply discarded. (In other cases it may be appropriate to retain the slow modes and eliminate the fast modes.) In a complex, interconnected system with time constants at different orders of magnitude, this type of simplification may be rather difficult to implement. In such cases, a combination of suitable aggregation, decoupling, and model reduction techniques can be used.

When human operators control a system other types of simplifications and approximations may be appropriate. This arises due to the inability of the human to fully utilize the state and output variables. It is desired to obtain reduced order models that are compatible with possible internal models of the system that the human operator has in a supervisory control situation. Manual control models such as the crossover model or the optimal control model cannot be used since we are not concerned with predicting human performance [21]. The possibility of developing a suitable simulation of the physical system that may also aid in the development of cognitive models for a dynamic system is being investigated. A simple model structure suitable for representing a wide variety of

complex dynamic systems is desired.

"Cognitive" models

Investigators in artificial intelligence, cognitive science, and psychology have done research in mental models of simple systems. Representative examples of various attempts at modeling are papers by Kieras [22], Rieger [23], de Kleer and Brown [24], Stevens [25], Johnson-Laird [26], and Gentner and Stevens [27]. The cognitive models described in these papers deal with models of devices. Usually no dynamics is involved. Only the sequence of operations are described without regard to their exact time relationships. The following discussion will illustrate the nature of these modeling approaches.

Kieras [22] describes mental models that people have about devices, based on descriptions of functions and components. He provides a survey, where experts and non-experts were asked to describe simple electronic devices such as radios. He concludes that people's knowledge of devices consists of major categories such as what the device is for, how it is used, its structure in terms of its subdevices, its physical layout, how it works, and its behavior.

Rieger [23] proposed a "common sense algorithm" (CSA), that provides a data structure for modeling cognition and information processing. The structure is defined in terms of five "building blocks", Wants, Actions, States, Statechanges, and Tendencies. These five are used to construct 25 primitive links. The goal of this approach was to express the "dynamics of just about everything" using the primitives. The CSA was used for modeling the operation of a reverse-trap toilet, and the sawing of a board in half.

Mental models of physical mechanisms are developed by de Kleer and Brown [24] using certain "esthetic principles". They employ qualitative simulations, referred to as "envisionments", as a means of generating mental models. Envisionment follows from the device structure, and results in "causal attribution". The mental model is described as a combination of envisionment and causal attribution. Esthetic principles are then identified that are useful in evaluating mental models of physical devices as well as for help in forming the mental models. A qualitative model is given for a buzzer. Detailed description of a buzzer mechanism is given, and a plausible mental model of its functions is discussed.

Both the CSA approach and the models based on the esthetic principles appear to have a number of drawbacks for modeling dynamic systems. These models could be possible means of explaining what appears to go on in the mind. However, even for fairly simple devices, these model descriptions become enormously complicated and large. Coupled with the growth of the model size is the problem of whether or not the humans are actually capable of managing such complex, detailed models. At this point in time, experimental evidence in favor of these models appears rather limited.

Attempts at using such "qualitative" models for describing dynamic systems have been made in connection with the design of training simulators. Such a simulation is described in connection with a steam powerplant [25]. This is helpful for understanding device structures. Simulation models of components have been built. The simulation appears to

be very useful for training, by taking the trainee step by step through the dynamical behavior.

Summary

The inherent "robustness" and the formal structure of the cognitive models are attractive. However, they appear to complicate the problem and greatly increase the number of states and, consequently, computational requirements. Possibly, these models provide a means of representing a physical device, though not necessarily a good method for representing a complex, interconnected dynamic system with continuous states. Real time simulation of these models would be extremely difficult, if not impossible.

On the other hand, control theoretic approaches provide a means of simulation without compromising rigor and computational convenience and efficiency. No qualitative behavior is available, however. What is needed is a modeling approach that incorporates the desirable features of both of the above methods for real time simulation.

PROPOSED MODELING APPROACH

Basic requirements

The issue of primary interest at this point is the simulation of the dynamics so that a simple representation is possible. The simplified model may not be appropriate for control purposes or for predicting the numerical values of any of the system variables as a function of time. However, it should be sufficient to explain the qualitative behavior of the entire system or any of the subsystems.

When a particular component or element fails, it may take some time for the problems to be felt by the entire system. The problems propogate through the system in a sequential order in the beginning. When a number of components have been affected, the system could degrade at a faster rate. Usually the operator uses the sequence in which the events happen at various parts of the system as a clue to detect and isolate the problem. Hence the dynamic simulation must preserve the order in which events occur when a problem begins. The sequence of actions required to isolate and identify the event should be the same for both the simulator and the actual system. Also, the response for a given input must be the same in both cases. Since the detailed state information may not be necessary for comparison, or for producing the same outputs, reduced order representation may be acceptable.

All the important variables that could aid in the detection of the fault must be provided in the simulation. The level of approximation will determine the extent and depth to which the fault isolation and detection can be carried out. With an appropriately simplified dynamics it should be possible to simulate the overall, gross qualitative nature of changes.

Basic methods

Linearized dynamics about some steady state will be used for modeling a dynamic system. When transient operations are possible, as in startup and shutdown phases of complex systems, a number of operating points can be established about which the system equations are linear.

Large dynamic systems such as those considered here have a number of interacting loops. Within any loop, there could be coupling between subsystems and modes, characterized by various degrees of interdependence. For modeling purposes, large systems will be decoupled so that interactions between loops is minimized without compromising the "dynamic fidelity". Within a loop, various modes will be decoupled to the extent possible.

Approximations to the actual dynamics are to be achieved by the elemental combinations of first and second order systems. A first order system can be described in terms of a time constant and a gain. In a discrete time representation, the time constant can be expressed simply as the number of time intervals required for steady state after a step input has been applied. A second order system can similarly be explained by the number of cycles before a suitably defined steady state is reached. This would combine the damping and the natural frequency. For a second order system, it seems plausible that the number of cycles before steady state can describe the response appropriately. Higher order systems can be approximated by various combinations of these "elemental" systems.

The possibility of modeling the system using a combination of dynamical and symbolic models is being explored. An appropriate scheme would be to use conventional dynamic system models for subsystems that require frequent state changes. Subsystems where changes in variables are small and infrequent could be modeled using symbolic structures. The higher levels will have some form of "knowledge representation", where semantic, verbal representations of dynamics will be used. The idea is to obtain a mix of algebraic and symbolic operations that result in reduced storage and computational requirements.

Yet another possibility is the use of fuzzy set theory. Fuzzy set theory has been found useful for designing controllers [28,29,30,31]. Under manual control, operators develop control laws from experience with the systems over a period of time. Heuristics are necessary since such systems tend to be rather slow and complex. The fuzzy controllers link the control laws expressed using linguistic expressions and numerical values of control required to operate the system. Such a link between qualitative representation of control laws and quantitative values of control suggests that fuzzy set theory may be useful for modeling of dynamic systems. A model based on fuzzy set theory can also provide system states and outputs in a form compatible with the mental models that the human operators use to control the system.

Overall scheme

For simulating the dynamics, the basic structure of a subsystem will have three parts. The first part will be used to convert and transform the states and other system variables for use in the qualitative representation. This is a "pre-processor" that quantizes the states

roughly so that the generic model of dynamics can handle it. The central, second part will be "generic" dynamics made up of simple combinations of first and second order systems. The third part will transform the results of the qualitative dynamics back to quantitative state and system variable values as necessary. This is a "post-processor" that transforms the qualitative dynamic states into a suitable quantitative form. It is seen that the first and the third parts form the interface of the subsystem with the dynamics part. The quantitative values of the states are perhaps rough quantizations that preserve the order and sequence relationships among various system variables. The pre-, and post-processors are in some sense inverses or complements of one another.

An application

An application of these methods to modeling the dynamics of a marine powerplant is being pursued. A typical marine powerplant uses an oil-fired boiler to produce steam. The steam drives a series of turbines. The turbines are used to derive mechanical energy to drive the propellors, pumps, and compressors, and for electricity generation. Water for the boilers is distilled from the sea. The water-steam system forms a closed loop, with any leaks being compensated for by the addition of distilled water. Air is taken from the atmosphere, pre-heated by exhaust gases, and used to burn the fuel oil in the burner. The hot gases are then used in the boiler to generate steam, as well as to superheat the steam coming out of the boiler. Before exhausting into the atmosphere, the hot gases pre-heat the water entering the boiler. Fuel oil is preheated by auxiliary steam before coming into the burner. Since there are a number of mechanical subsystems such as turbines and pumps, a lube oil loop is used. This is a closed loop, where the lube oil goes through the bearings and other elements, picks up heat, and gets cooled by water. Failures in any element of the four major systems (water, air, fuel, and lube) could affect the operation of the powerplant.

A simplified version of the powerplant comprising of the the water/steam loop and the air loop will be simulated using the modeling concepts discussed above. The major components of the reduced system are shown in Figure 1. Each of the individual subsystems shown will be replaced by modules with approximate dynamics.

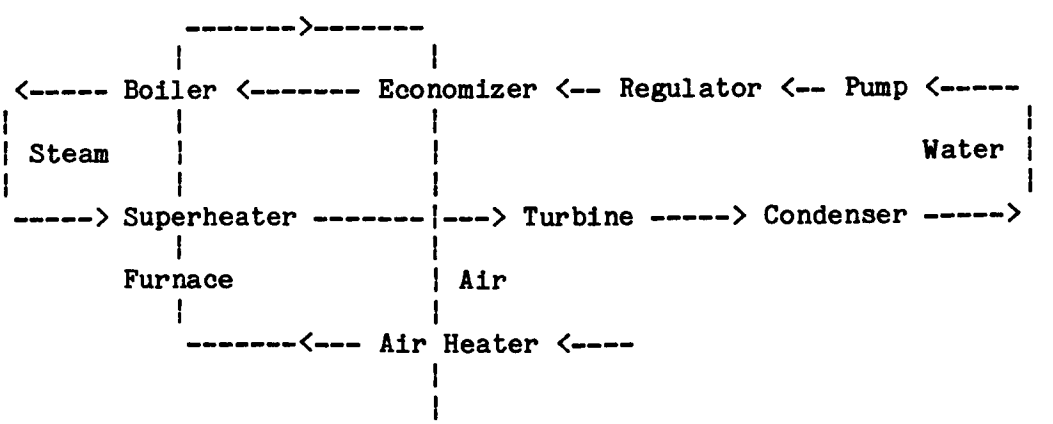


Figure 1. Air and Water/Steam Loop of a Marine Powerplant

A qualitative model for the dynamics will be used in a moderate fidelity simulator of the marine powerplant. This moderate fidelity simulator will be compared to a high fidelity simulator and a low fidelity simulator in transfer of training studies. These studies will focus on the problem solving performance of marine engineering personnel.

The high fidelity simulator consists of a replica of an engine and control room with the associated control and display panels. The engine dynamics is simulated on a mini-computer to a high level of detail. Simulated equipment such as burners, boilers, and plumbing are seen from the engine room as it would be on a real ship. Environmental conditions such as heat, humidity, and vibration are also simulated. Hence it is apparent that this has high fidelity along any reasonable dimension that one uses.

The low fidelity simulator will have a limited amount of information displayed, in a format somewhat different from the real equipment [32]. Information will be provided using system schematics. Symptoms corresponding to the failures or problems will be presented on a video display terminal. No dynamics is involved. The subject can check for the status of gauges and other displays through the computer interface. Display status stored in advance for each of the failure situations will be provided. The subject cannot control the system. Hence it is seen that this has no dynamic fidelity, low physical fidelity, and moderate structural fidelity. The latter is given through the system schematics.

Comparing the moderate-fidelity simulator to the low and high fidelity simulators, the simplified dynamics and qualitative response to operator's control inputs of the moderate-fidelity simulator provide an intermediate level of dynamic fidelity. Structural fidelity is maintained at the moderate level of the low-fidelity simulator. Physical fidelity is improved by providing a limited number of displays.

SUMMARY AND CONCLUSIONS

This paper has been concerned with the development of qualitative models of the dynamics for large, interconnected systems. Such models will be incorporated in the design of moderate fidelity simulators used for training in problem solving. The qualitative models of dynamic systems are also expected to be helpful in representing the mental models that supervisory controllers have of systems.

Preliminary work has been completed in deriving the appropriate equations for the dynamics. Approximate equations for dynamics will be derived for the major components of the marine powerplant. The suitability of the approximations will be evaluated by comparision of the simulation results between two sets of dynamics. The first set will use the actual system equations for the perturbed states. The second set will involve the approximate models. For various events that are characteristic of the subsystems, the overall system response will be presented to human operators. They will be asked to use these system responses to correctly detect and diagnose faults. If the approximation is functionally correct, the fault diagnosis performance on either of the systems will not differ from each other. The approximation will be modified if necessary until a reasonable performance agreement is achieved between the two models of

dynamics. Transfer of training studies will then be conducted using the simulators at low, moderate, and high fidelity levels.

ACKNOWLEDGEMENT

This research is supported by the Office of Naval Research under contract N00014-82-K-0487, Henry M. Halff, Contract Monitor. Discussions with William B. Rouse helped improve the paper substantially.

REFERENCES

[1] Hays, R. T., ed. Research Issues in the Determination of Simulator Fidelity: Proceedings of the ARI Sponsored Workshop, U. S. Army Research Institute for the Behavioral and Social Sciences, Technical Report 547, November 1981.

[2] Baum, D. R., "A Framework and Topics for Empirical Research on Training Simulator Effectiveness", in [1], pp.25-38.

[3] Montague, W. E., "Is Simulation Fidelity the Question?", in [1], pp.93-109.

[4] Orlansky, J., "Research on the Fidelity of Simulators", in [1], pp.143-151.

[5] Valverde, H. H., "A Review of Flight Simulator Transfer of Training Studies", Human Factors, vol. 15, no. 6, pp. 510-523, 1973.

[6] Blaiwes, A. S., Puig, J. A., and Regan, J. J., "Transfer of Training and the Measurement of Training Effectiveness", Human Factors, vol. 15, no. 6, pp. 523-533, 1973.

[7] Adams, J. A., "On the Evaluation of Training Devices", Human Factors, vol. 21, no. 6, pp. 711-720, December 1979.

[8] Gonzales, R. C., and Thomason, M. G., Syntactic Pattern Recognition, Addison-Wesley, 1978.

[9] Fu, K. S., Syntactic Pattern Recognition and Applications, Prentice-Hall, 1981.

[10] Henley, E. J., and Williams, R. A., Graph Theory in Modern Engineering: Computer Aided Design, Control, Optimization, Reliability Analysis, Academic, 1973.

[11] Chen, W.-K, Applied Graph Theory: Graphs and Electrical Networks, North-Holland, 1976.

[12] Bunke, H., "Attributed Programmed Graph Grammars and Their Application to Schematic Diagram interpretation", IEEE Trans. Pattern Anal. Machine Intell., vol.PAMI-4, no.6, pp. 574-582, November 1982.

[13] Kwakernaak, H., and Sivan, R., Linear Optimal Control Systems,

Wiley-Interscience, 1972.

[14] Eykhoff, P., System Identification: Parameter and State Estimation, Wiley, 1974.

[15] Mahmoud, M. S., and Singh, M. G., Large Scale Systems Modelling, Pergamon, 1981.

[16] Aoki, M., "On Aggregation of Similar Dynamic Systems", in M. Singh, A. Titli, eds., Large Scale Systems Engineering Applications, North-Holland, pp. 7-19, 1979.

[17] Siret, J. -M., Michailesco, M., and Bertrand, P., "On the Use of Aggregation Techniques", in M. Singh, A. Titli, eds., Large Scale Systems Engineering Applications, North-Holland, pp. 20-37, 1979.

[18] Genesio, R., and Milanese, M., "A Note on the Derivation and Use of Reduced-Order Models", IEEE Trans. Automat. Contr.= vol.AC-21= no. 1, pp. 118-122, February 1976.

[19] Kokotovic, P. V., O 'Malley, Jr. R. E., and Sannuti, P., "Singular Perturbations and Order Reduction in Control Theory - An Overview", Automatica, vol. 12, pp. 123-132, 1976.

[20] Moore, B. C., "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction", IEEE Trans. Automat. Contr., vol. AC-26, no. 1, pp. 17-32, February 1981.

[21] Sheridan, T. B., and Ferrel, W. R., Man-Machine Systems: Information, Control, and Decision Models of Human Performance, The MIT Press, 1974.

[22] Kieras, D. E., What People Know About Electronic Devices: A Descriptive Study, University of Arizona Technical Report No. 12 (UARZ/DP/TR-82/ONR-12), 1 October 1982.

[23] Rieger, C., "The Commonsense Algorithm as a Basis for Computer Models of Human Memory, Inference, Belief and Contextual Language Comprehension", in Proceedings of Theoretical Issues in Natural Language Processing, Cambridge, MA, June 1975, pp. 180-195.

[24] de Kleer, J., and Brown, J. S., "Mental Models of Physical Systems and Their Acquisition", in Cognitive Skills and Their Acquisition, John R. Anderson, ed., Lawrence Erlbaum Associates, pp.285-309, 1981.

[25] Stevens, A., "Quantitative and Qualitative Simulation in Portable Training Devices", June 1982, Bolt Beranek and Newman Inc., for National Academy of Sciences.

[26] Johnson-Laird, P. N., "Mental Models in Cognitive Science", in Perspectives on Cognitive Science, Donald A. Norman, ed., Ablex/Lawrence Erlbaum Associates, pp.147-191, 1981.

[27] Gentner, D., and Stevens, A. L., eds., Mental Models, Lawrence Erlbaum Associates, 1983.

[28] Mamdani, E. H., and Assilian, S., "An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller", Int. J. Man-Machine Studies, vol. 7 pp. 1-13, 1975.

[29] Mamdani, E. H., "Advances in the Linguistic Systhesis of Fuzzy Controllers", Int. J. Man-Machine Studies, vol. 8 pp. 669-678, 1976.

[30] King, P. J., and Mamdani, E. H., "The Application of Fuzzy Control Systems to Industrial Processes", Automatica, vol. 13 pp. 235-242, 1977.

[31] Tong, R. M., "A Control Engineering Review of Fuzzy Systems", Automatica, vol. 13 pp. 559-569, 1977.

[32] Govindaraj, T., and Su, Y.-L., "Training Simulator Design for a Marine Engine Room at Different Levels of Fidelity", to be presented at the 27th Annual Meeting of the Human Factors Society, Norfolk, VA, 10-14 October 1983.

PILOT CONTROL DURING AIR COMBAT MANEUVERING

By
R. Lynn York
Center for Systems Development
SRI International
Menlo Park, CA 94025

Today's high-performance tactical aircraft constitute a complete weapon system. The major components are the aircraft, the pilot, and the aircraft's armament. Successful employment of these sophisticated, state-of-the-art "systems" requires that each major component perform satisfactorily, individually and collectively. Integral to successful employment is the ability of the pilot to control both the aircraft and its armament so that a mission is executed satisfactorily. How do you measure mission performance of such a complex system with so many elements? One way is to monitor the four elements shown in Figure 1: (1) the aircraft, (2) the pilot (physical), (3) the pilot (mental), and (4) the armament. All four of the elements contribute to mission success or failure. By monitoring all four elements it would be possible to isolate problem areas. Was mission failure due to the aircraft, the pilot, or its armament?

In this paper, a concept is proposed that will allow monitoring the various elements during operational training exercises, with the focus placed on pilot interaction with his aircraft and its armament. The proposed data collection system is a combination of the Tactical Aircrew Combat Training System/Air Combat Maneuvering Instrumentation (TACTS/ACMI) and the Ambulatory Physiological Monitoring System (APMS). The expected output is identification of pilot-peculiar problems that may be associated with inadequate or improper control of his system.

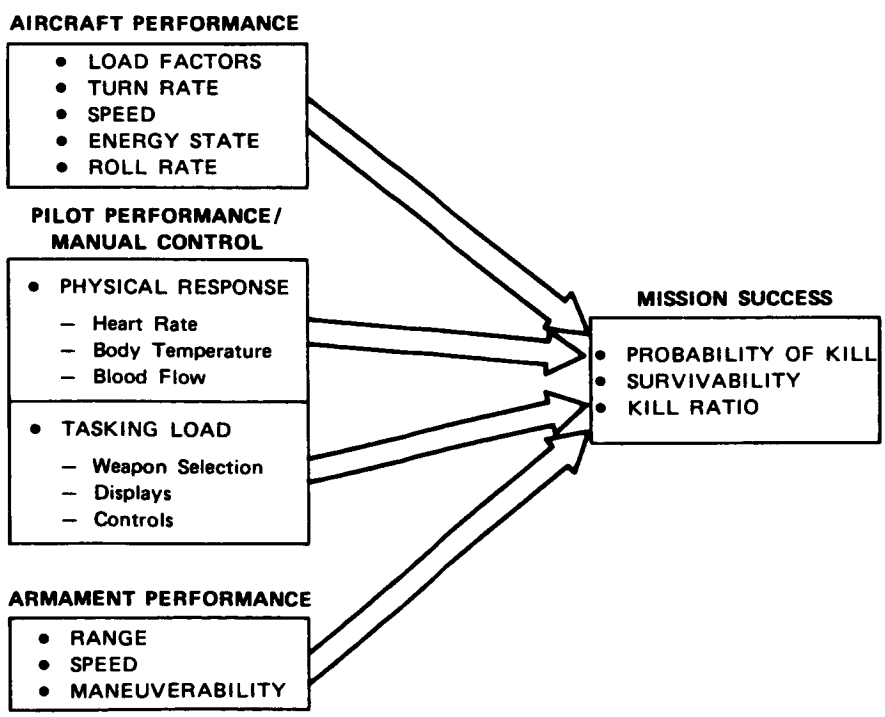


FIGURE 1

A PSYCHOPHYSICAL EVALUATION OF AN AREA OF INTEREST (AOI) DISPLAY

Stephen Stober, Andy Lippay, Murdoch McKinnon, Brian Welch,

CAE Electronics Ltd.

Thomas Longridge
HRL, Willams AFB

ABSTRACT

In order to provide a low cost, wide field of view visual system for aircraft simulation, techniques are being developed for displaying a high resolution area of interest (AOI) within a larger low resolution field of view. One project being developed by CAE Electronics Ltd., and the Human Resources Laboratory (Williams AFB) is the Helmet Mounted Desplay (HMD), in which an AOI is slaved to the pilots eye movements. The following paper describes psychophysical experiments which are being set up in an attempt to define the specifications for such a display system.

INTRODUCTION

Aircraft combat simulators have become an invaluable part of fighter pilot training. These trainers can give the pilot simulated experience in take-offs, landings, air refueling, carrier landings, low level flight and basic aerial combat maneuvering. Simulated combat sorties can provide the fighter pilot with some of the skills normally acquired during wartime engagements.

The fidelity of the simulation, especially the visual aspect, is a critical factor in determining whether a pilot acquires appropriate combat skills. The visual system, for instance, must have a resolution which is high enough to simulate apparent targets at a distance, such that the image size approaches threshold dimensions. Visual systems that are presently available however, have limited resolution characteristics. For instance, a typical Computer Generated Imagery (CGI) visual system used in combat trainers, has a resolution of approximately 8 arc minutes per pixel. This limited resolution makes it impossible to depict an accurately scaled enemy fighter beyond 2 - 3 miles, since smaller images would disappear between the raster lines of the CRT display.

Using conventional Computer Generated Imagery (CGI) - CRT display techniques, one could increase the display resolution by increasing the number of displays generating the total field of view (FOV), thereby increasing the concentration of available imagery.

Spooner (1982) estimates that in order to attain a resolution of 2 arc minutes per pixel, it would require approximately 24 CRTs to cover a hemisphere with imagery. This type of display would not only be prohibitive in cost, but impractical to install and maintain. Another approach would be to employ CRTs with higher resolution, approaching approximately 5,000 TV lines per picture height. Technologically however, a CRT display with this resolution is presently unavailable and in the event that it is, its cost would undoubtedly be prohibitive.

A current approach to the problem of generating high resolution visual imagery, is to employ an Area of Interest (ADI) display, which takes advantage of the acuity properties of the human visual system. Although visual acuity is approximately 1 minute of arc in the fovea, resolution drops off abruptly to about 10% of that, 20 degrees into the periphery. As a consequence of this acuity profile, it may not be necessary to present a highly detailed scene in the area peripheral to the viewing direction. In fact, the principle of an AOI display is to present high-detailed visual information in the central field of view and lower resolution information in the periphery.

CAE Electronics Ltd., and the Human Resources Laboratory at Williams AFB are presently developing a Helmet Mounted Display (HMD) using the AOI display principle. The HMD consists of a visual display having three nested regions, with progressively higher resolution and higher level of detail approaching the central area of view. The high and medium resolution areas will be slaved to eye movements relative to the head, while the peripheral low resolution FOV will be slaved to the angular position of the head. In this way, as the pilot moves his head, the overall CGI scene will change with variations in the point of regard. As well, eye movements measured by an eye tracker mounted on the helmet, will be taken into account such that the high and medium resolution fields will always be displayed in the central viewing area.

The visual images are produced by GE light valves projecting CGI imagery through fibre optic cables onto two miniature Pancake windows mounted on a flight helmet. Each eye views an 80° horizontal X 60° vertical FOV. With a 25° overlap area, the total instantaneous field of view is approximately 135° horizontal X 60° vertical. The high and medium resolution fields will be inset into the larger low resolution area and slaved to movements of the eye. The resolution of the central high resolution field will be approximately 2 arc minutes while that of the peripheral field will be approximately 8 - 10 arc minutes.

The ultimate goal of this visual system is to produce an AOI display which is indistinguishable from the real world in terms of pilot performance for a required task. In order to approach this goal, parameters for an AOI display such as inset size, shape, resolution differential, and edge blending must first be determined. Baldwin (1981) reported an experiment which investigated these parameters with an AOI display being developed at the Naval Training Equipment Center (NTEC). Using magnitude estimates, Baldwin's observers rated an AOI display on a scale of 1 - 4 basing their judgements on the quality of the image. Additional subjective comments by the observers augmented the data collection. The results of the NTEC experiment suggested that a 25° AOI inset with soft, blended edges and an AOI/peripheral display resolution drop of approximately 5/1, produced the best stimulus configuration.

The purpose of the present experiment is to extend the investigation initiated by Baldwin. Although Baldwin's experiments reflected observers' judgements on the perceptual quality of the various AOI displays, it is believed that a study investigating how the performance of an observer is affected by varying characteristics of a display would best determine the required parameters. Specifically, in the proposed study using detection and recognition tasks, the manner in which AOI display characteristics affect performance will be assessed, by comparing the results produced with various mixed resolution AOI displays and a homogeneous high resolution display. In this manner, an ojbective assessment of the parameters required in an AOI visual system will be made, using tasks not unlike those performed in an aircraft combat mission trainer.

EXPERIMENTAL PROCEDURES

Apparatus

The apparatus that will be used in the proposed set of experiments is an Eye-Slave Projection (ESP) system developed by CAE Electronics Ltd., (Fig. 1). It consists of three Kodak Carousel lamps, optics and projection lenses mounted on a optical bench such that the images from the three channels may be superimposed on a rear projection screen. Stimulus slides may be placed in each of the three holders situated between the projection and condenser lenses. The mask holders located in front of the stimuli are mounted on an aluminum plate (Fig. 2) that moves horizontally along linear bearings. The mask plate is fixed onto another aluminum plate that moves vertically. In this way, the movement of the two plates allows the mask to travel in two dimensions. Light passing through the slides and masks produces a projected image of a high resolution area inset into a low resolution field.

The motion of the plates is controlled by hydraulic actuators which receive input from the CAE designed control circuitry. The circuitry accepts signals, having a magnitude of approximately 300 mV/degree, from the Gulf and Western (Model 200) eye tracker. The slightly underdamped system results in the mask holder travelling 1 inch in 10 msec (with a 5% overshoot). This travel is equivalent to the projected image moving approximately 6000 degrees/ second.

The proposed set of experiments will initially be investigating eye slaved movements in the horizontal dimension. Thus, a rear projection screen will be masked such that a single horizontal line is presented to the observer. Due to the limited range of the eyetracker ($\pm$ 20°), the field will subtend 50° in length by approximately 45' in height. As shown in Figure one, two parallel front surface mirrors are located in front of the projection lenses. Light passing through the two mirrors is reflected onto the rear projection screen viewed by the observer. One of the mirrors rotates along its horizontal axis such that the projected image may be vertically displaced. A computer controlled stepping motor located at the axis of rotation actuates the motion of the mirror.

STIMULI

The stimuli are constructed by photographing arrays of dot matrix generated characters, producing positive slides. Slides of identical character

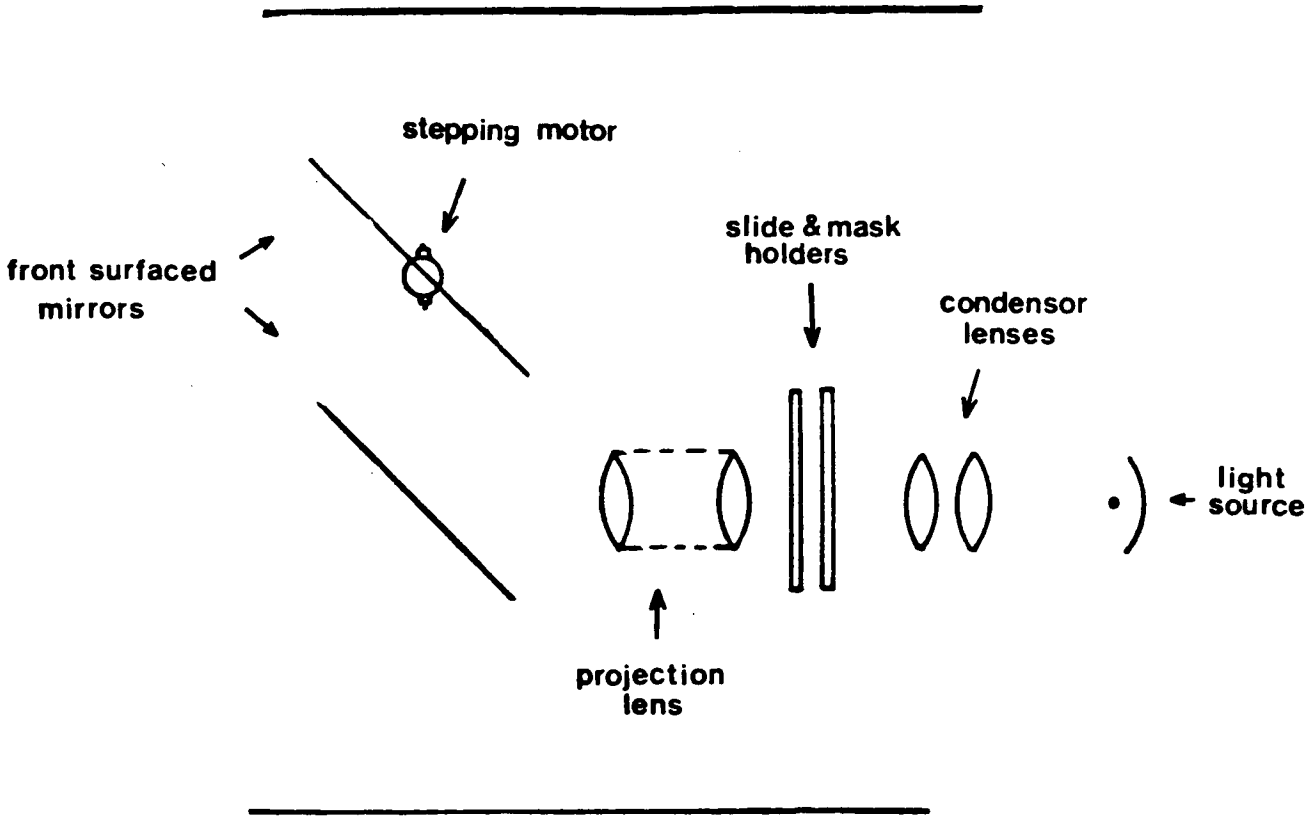


Figure 1. A schematic of the eye-slave projection optics.

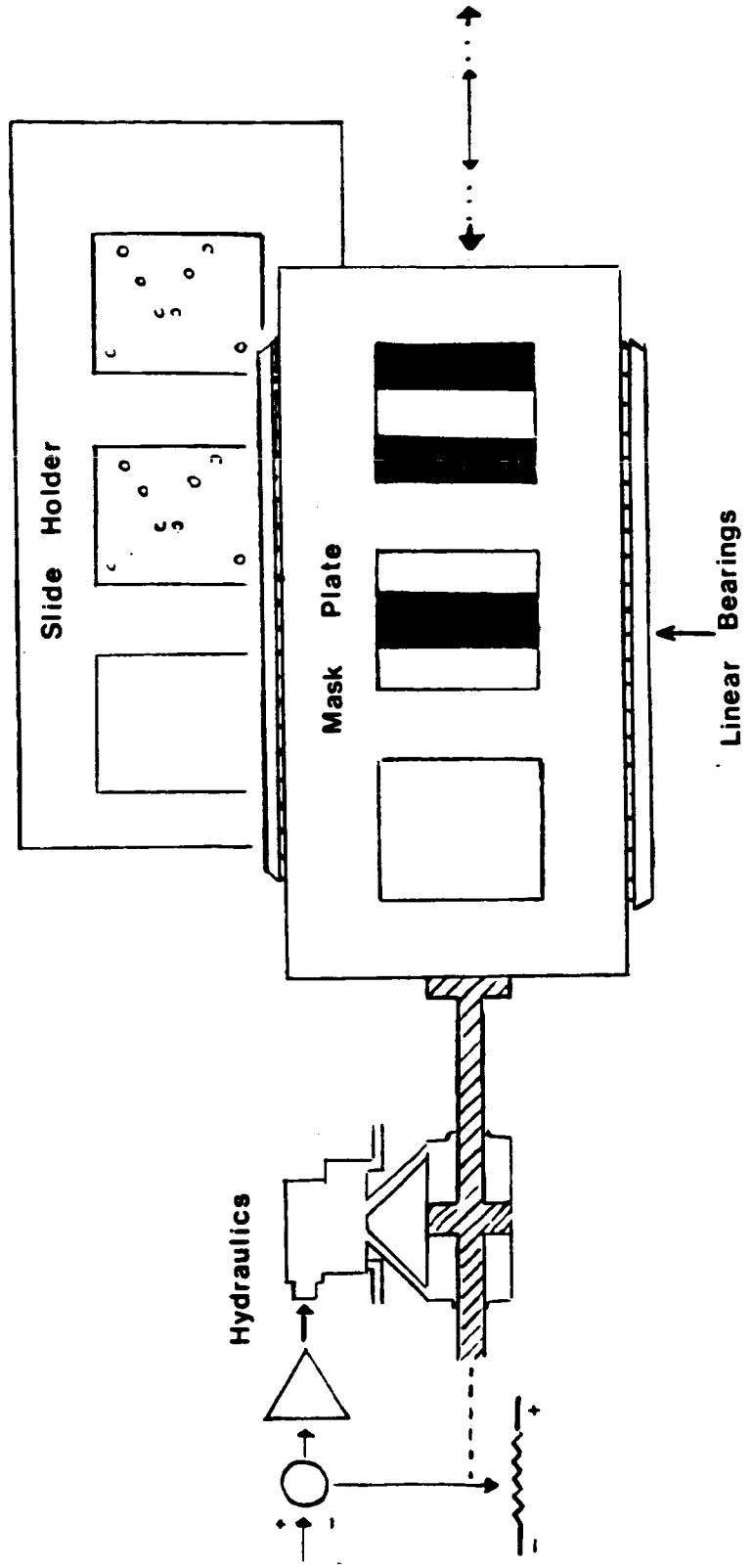


Figure 2. The hydraulic actuator controlling the mask plate in the X dimension

arrays will be placed in two channels of the ESP such that their images are superimposed at the rear projection screen. The 'low resolution' slides are created by defocussing the camera and changes in character size resulting from this procedure, are compensated for in the photographic process. Different 'low resultion slides' will be produced ranging in resolution from 6 arc minutes per line pair - 20 arc minutes. The 'high resolution' slides will have a resolution of 2 arc minutes. Characters will consist of landolt C's with the following dimensions; width = 13', height = 17', landolt C gap = approximately 10'.

Method

Table 1 illustrates the various AOI inset sizes, FOV, and AOI/FOV resolution relationships that will be used in the experiment. For these initial experiments, a blurred transition between the high and low resolution areas will be used. For each randomly presented condition, the observer will be required to indicate via an appropriate lever press, whether a gap in the landolt 'C' is oriented to the right or left. For approximately 50 spatial locations across the FOV, landolt 'C's oriented in both directions will be available for presentation (see Figure 3a). The computer will actuate the stepper motor such that on each trial, the location and orientation of the character will be randomly determined. Reaction time to accomplish the task, frequency of saccades, and average fixation duration will be compared across conditions. A microcompuer will control the experimental events as well as measure the responses.

Experiment 2 will be conducted in order to explore the influence that background texture has on this task. The same conditions and procedures outlined above will be used; the only difference will be found in the stimulus array. Instead of the characters being presented in a blank field the target character will be embedded within a string of O's (see Figure 3b).

Summary

The results generated by this study will begin to describe those variables that are critical in the production of an AOI display. In future experiments, other performance measures and AOI display variables will be investigated using the eye slave projector. These studies will look at such variables as: AOI vertical size, character type, AOI/peripheral resolution gradient, luminance, and AOI movements in two dimensions. Ultimately, the results from these studies will be used to generate a helmet mounted visual display that best exploits the characteristics of the human visual system.

References

Baldwin, D.M. Area of Interest - Instantaneous Field of View Vision Model. Proceedings of the Image Generation/Display Conference II, 1981.

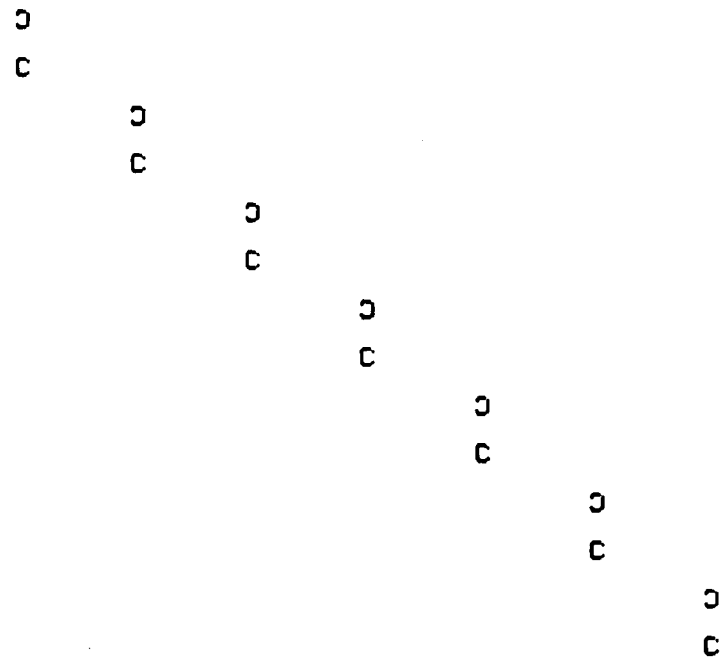
Spooner, A.M. The trend towards Area of Interest in Visual Simulation Technology. Proceedings of the Interservice/Industry Training Equipment Conference, 1982.

Table 1.

Experimental Conditions

FOV	AOI size (degrees)	Resolution (arc min./line pair)	
		High	Low
50	10	2	6
50	15	"	8
50	20	"	10
50	25	"	12
50	30	"	14
		"	16
		"	18
		"	20

(A)



(B)

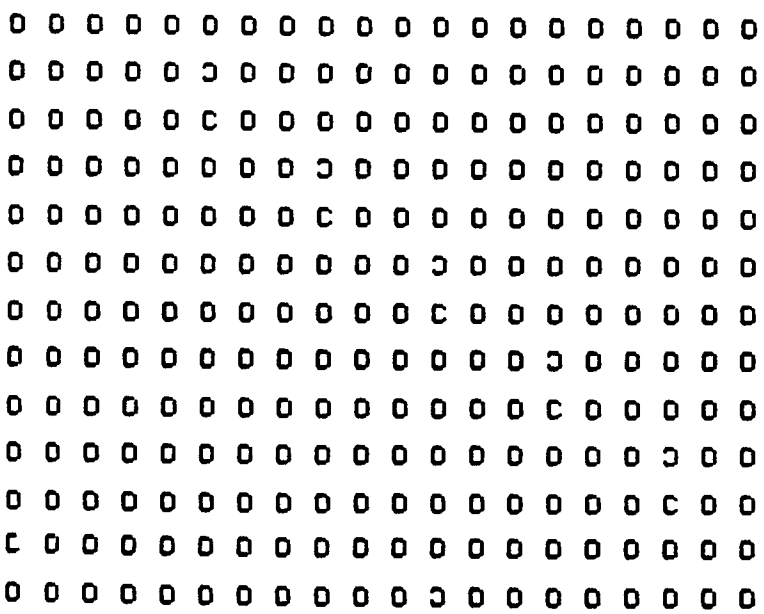


Figure 3. The stimulus configurations for experiment 1 (A) and 2 (B).

PERCEPTION OF ROLL RATE FROM AN ARTIFICIAL HORIZON AND PERIPHERAL DISPLAYS

R.J.A.W. Hosman and J.C. van der Vaart

Delft University of Technology, Department of Aerospace Engineering, The Netherlands

SUMMARY

It was shown earlier that peripheral side displays can help to improve the performance of subjects in tracking tasks. The work reported in the present paper was undertaken to find out whether this improvement was due to an increase in 'arousal level', a more accurate or a faster perception of roll rate.
The experiment described investigates the perception accuracy of roll rate by subjects from a central CRT-display (simulated artificial horizon), peripheral side displays (moving chekerboard pattern) and both displays combined. Discrete values of roll rate were presented to the subjects during exposure times between 0.1 and 0.8 second.
Immediately after stimulus exposure the displays were either plainly blanked or masked by a dithering line on the central display or dithering checkerboard patterns on the peripheral displays.

Results show that the roll rate perception process from peripheral displays is more accurate and up to 0.1 second faster than from the central display. Masking shows to have a different influence on the central visual perception than on the peripheral visual perception.

1. INTRODUCTION

The tests described in this paper are a part of a larger research program in the field of pilot's motion perception, presently in progress at the Department of Aerospace Engineering of the Delft University of Technology.
It is well known that simulator or airplane motion visible in the peripheral visual field have an influence on pilot's control behaviour. See Refs. 1, 2, 3 and 4.
In Refs. 3 and 4 an experiment was described where all combinations of a central CRT-display (simulating an artificial horizon) two peripheral side displays and simulator cockpit motion were used to establish the influence of the separate and combined ways of motion perception on subject's tracking performance and control behaviour.
The results of that experiment showed a strong performance improvement due to the addition of peripheral visual cues and motion cues. Further a considerable change of subject's control behaviour in terms of increased cross-over frequency and/or phase margin, was demonstrated.
Due to the fact that these improvements were achieved only by changing the display configuration (i.e. adding peripheral visual cues and or motion cues) it was assumed that the above mentioned improvements resulted only from improvements of the first step of the subject's information processing i.e. the perception process. Due to the lack of information on the accuracy of perception and the time needed to observe and process the different sensory stimuli in a typical piloting situation, a more detailed investigation on motion perception was started.
It will be obvious that visual stimuli presented on a CRT-display or on a TV-monitor are much easier manipulated than the motion stimuli generated by a simulator motion system. Therefore it was decided to study motion perception

from central and peripheral visual displays first.
In Ref. 5 first results of two experiments on roll attitude and roll rate perception were presented. A number of important differences between roll attitude perception and roll rate perception emerged from these first attemps, the most important being perhaps that attitude perception is more accurate and is done quicker than rate perception.
The present paper describes an extended study on roll rate perception only in which the same central and peripheral displays were used. Just as in the previous experiment, discrete roll rate stimuli were presented to the subject. Subjects were requested to make an accurate estimation of the rates and to answer by pressing the appropriate key of a keyboard. One important aspect of a closed loop task was retained by displaying the subject the error magnitude immediately after the response. In the present experiment the range of exposure times was shifted to exposure times shorter (down to 0.1 sec.) than the ones in the experiment reported in Ref. 5. In one experimental configuration the displays were blanked after stimulus exposure. In the other the displays were masked. This was done in order to investigate, if as for attitude perception, there is a so-called 'iconic memory' for rate perception.
With these two extensions a wider range of perception accuracy could be assessed. The influence of redundant observations, resulting from the combination of central and peripheral displays, on the roll rate perception accuracy, were studied into some detail.
Apart from the extensions to the tests of Ref. 5, the number of subjects was extended from two to three in such a way that all three test subjects completed all tests of the extended experiment.
The main output variables of the experiment are the perceived roll rate and the response time. The question has to be raised whether changes in response time due to changes in display configuration and exposure time should be attributed to the perception process alone. The further steps of the information processing however are not changed. It has been argued in Ref. 6 that the effects of disturbances of parts of the information processing on the response time are additive, if the disturbances each act on a different step of the information processing.
Thus if the perception process in the present experiment needs a longer or shorter time interval the response time will correspondingly be longer or shorter and there is no need to expect any interaction on the decision and response steps of the information processing.

2. TEST FACILITY AND DATA REDUCTION

All measurements were performed in a low noise room where, in front of the subject' seat, a central (foveal) CRT-display (Tektronix 604 monitor), was mounted in an instrument panel. Peripheral visual cues were provided by two TV-monitors (Bosch Fernseh Monitor) placed on each side of the subject's seat, see Fig. 1. Subjects gave their responses via a digital keyboard, see Fig. 2.
The relative positions of central and peripheral displays and the subject's eye reference point are shown in Fig. 3.
Also shown in Fig. 3 is the image on the central display, simulating an artificial horizon. The repetition rate was 250 times/sec. and the position of the horizon line was updated at a rate of 50 times/sec.
The peripheral displays showed a movable checkerboard pattern with squares of 5 x 5 cm generated by a moving pattern generator (developed at the Delft University) at a repetition rate of 30 frames per second.
All experimental runs were controlled by a hybrid computer (EAI Pacer 100).
A rather complete description of the test facility hardware and the computer programs is given in Ref. 7.
During the experiments, each run consisted of 100 stimuli at fixed intervals, the sequence during one interval was as follows, see Fig. 4.

At the beginning of an interval a new value e_{n_1} of roll rate was presented and this event was marked by an audiotone in the subject's headset. After observation the subject was required to respond by pressing the appropriate key of the keyboard. The response magnitude is designated by $\Delta 0_n$. Immediately after the response the error value

$$e_{n_2} = e_{n_1} - \Delta 0_n$$

was shown on the display, thus giving the subject an immediate knowledge of the result. The exposure time Δt_{exp} could be varied and was set at a constant value by the experimentor prior to each run. In one particular experimental condition, the exposure was retained until the subject's keyboard response. In all other conditions the stimulus e_{n_1} was made to disappear at the end of the preset exposure time. In that case subjects were required to give responses only after termination of the exposure, responses during exposure time were neglected by the computer program in all test runs with limited exposure time.
Termination of the exposure could be programmed in either of two ways. Apart from simply 'blanking' the display after exposure, an additional feature of the test facility was that immediately after exposure, the central and the peripheral displays could be masked.
The central display was masked by a horizontal line randomly moving up and down. This was controlled by a wide band (5 Hz) zero mean noise signal. The peripheral displays were masked by making the checkerboard pattern move randomly up and down by the same noise signal.

At the start of a next interval a new value of roll rate was generated by adding a quantized sample i_n of a noise signal to the latest value of the error value e_{n_2}:

$$e_{n_1} = e_{(n-1)_2} + i_n$$

The standard deviation of the noise signal from which the sample i_n was taken and the scaling of the display were chosen such that the number of discrete values and the practical range of i_n, e_{n_1}, e_{n_2} and $\Delta 0_n$ during the experiment corresponded to the range and the number of keys used on the keyboard device ($\pm$ 10 keys).
During each run, the variables i_n, e_{n_1}, e_{n_2}, $\Delta 0_n$ and the subject's reaction time RT were recorded.
After a run the following values were calculated by the computer:
the number of correct keyboard responses, n_c,
the number of times that a display level j was shown, $n\ s_j$,
the sum of observation errors per display level j, $n\ e_j$,
the mean value and variance of i_n, e_{n_1}, e_{n_2}, $\Delta 0_n$, the reaction time RT and a score parameter S_c, defined by

$$S_c = \sigma^2_{e_{n_2}} / \sigma^2_{e_{n_1}}$$

3. EXPERIMENT

As mentioned in the introduction the present experiment is an extension of the one on roll rate perception described in Ref. 5.
As before three display configurations were used: central display only (C), peripheral display only (P) and central and peripheral displays combined (CP).
The stimulus interval time was set at 2 seconds as in the earlier experiment.
The range of the roll rates presented was to be representative of routine airline flight ($\dot{\varphi}_{max} = \pm$ 25 deg./sec.), so the standard deviation was set at $\sigma_{\dot{\varphi}} = 8$ deg./sec. The range of keys to be used on the keyboard was set again

at ± 10 corresponding with ± 25 deg./sec. on the displays.
Based on previous experience the range of exposure times used in the present experiment was chosen such that the score parameter S_c, i.e. subject's performance, was most strongly influenced.
The minimum exposure time was imposed by the hardware of the test facility. The peripheral displays, having a repetition rate of 30 frames/sec., can only display 3 frames during an exposure time of 0.1 sec., which was, after extensive evaluation, considered as a minimum for the present tests.
In the previous experiment the minimum exposure time was varied between 0.2 and the response time RT (± 0.8 sec). In that exposure time range, the score parameter S_c for the central display configuration (C) was influenced, but in the case of the two other configurations (P and CP) the score parameter S_c did hardly change.
The pre-experimental evaluation showed that using 0.1, 0.15, 0.2, 0.3 and 0.4 sec as exposure times, the roll rate perception process and thus the score parameter S_c could be influenced over a wide range for all three displays configurations.
From the previous experiment on roll attitude perception and from Sperlings work Ref. 8 it is known that the perception of roll attitude or characters deteriorates if immediately after a short exposure, a random image is exposed.
This phenomenon called 'masking' is due to the fact that the 'noisy' picture partly destroyes the short visual storage of images (iconic memory). The question raised was whether for rate perception there would also exist a short visual storage or iconic memory. Therefore the experiment was extended with the masking feature for the exposure times mentioned above. In addition to the above mentioned experimental conditions, the exposure time was extended to the subject's response. In that case the displays could not be blanked or masked. With the above mentioned experimental conditions a total number of 33 different experimental runs were obtained. Each subject replicated all 33 experimental runs 5 times.

4. SUBJECTS AND TEST PROCEDURE

Three subjects, two University staffmembers and one Dutch Airworthiness Authority staffmember, all three qualified jet transport pilots, volunteered in the experiment. They were instructed to respond primarily as accurate and secondly as quickly as possible to the presented stimuli. They were not required to continually fixate there eyes on the central display but were free to look at the keyboard device when giving responses. If only peripheral displays were used, subjects were instructed to fixate their eyes after responding, on the blank central display, until the next response. Apart from the immediate feedback of the error after each keyboard response, subjects were informed of the score parameter S_c and the number of correct keyboard responses after each run.
For the preliminary evaluations and for the purpose of training a total number of 220 runs was made.
After an adequeate steady level of performance was obtained in this way, the experiment was carried out during a number of morning sessions. This number of sessions differed over the subjects due to the large number of experimental runs and depending on subjects availability. A total of 33 x 3 x 5 = 495 runs were made for the experiment.

5. RESULTS

The results of the score parameter in the present experiment are shown in Figs. 5 and 6.
Rate perception from the central display only with the exposure time Δt_{exp} equal to the response time RT will be referred to as the standard condition. There is a clear difference between the score parameters for roll rate perception from the central display (C) only on the one hand and the peripheral displays (P) and the combined central and peripheral display (CP) on the other.

From the peripheral displays a more accurate roll rate perception is achieved, especially at short exposure times.
It is apparent that the roll rate perception needs a much longer exposure time to reach the final accuracy, in the case of a central display than in the case of the peripheral displays.
Masking (see Fig. 6) influences the roll rate perception at the small exposure times (Δt_{exp}, 0.2, 0.15 and 0.1 sec.). In the case of the combined central and peripheral displays, subjects apparently are not able to derive any additional information from the peripheral displays at $\Delta t_{exp} = 0.1$ sec. This causes the score parameter S_c to increase from 0.158 to 0.475 due to masking at $\Delta t_{exp} = 0.1$ sec.
When considering the decrease of the score parameter with increasing exposure time the following analytical function can be used to model this phenomenon:

$$S_c = A + B\, e^{-\frac{\Delta t_{exp}}{\tau}} \qquad (5.1)$$

Equation (5.1) is in fact a tentative model of perception and its subsequent processing based on the reasonable hypothesis that a certain time duration is needed to obtain a reasonable estimate of a noisy observation.
This simple expression enables us to express the decrease of the score parameter with exposure time by the time constant τ. See Fig. 7.
The following table gives the different values obtained with eq. (5.1).

Display configuration	Blanking τ in sec.	Masking τ in sec.
Central	0.070	0.068
Peripheral	0.041	0.072
Central and Peripheral	0.041	0.061

In this way the shorter exposure time needed to perceive the roll rate if peripheral visual cues are available, are expressed in terms of smaller time constants in eq. (5.1).
In table 1 the results of analyses of variance on the score parameter S_c and response time RT are given.
Subjects, display configurations and exposure time are seen to have a significant influence on the score. Masking itself has no influence at all, but the interactions between masking and exposure time and masking and display configurations do have a significant influence.

In Fig. 8 the response times RT are shown as a function of exposure time. The mean values of the response times are shorter than the response time for the standard condition. The response time for rate perception from the central display is longer than for the conditions with peripheral and central and peripheral displays, see table below.

Display configuration	Blanking RT in sec.	Masking RT in sec.
Central	0.83	0.85
Peripheral	0.77	0.76
Central and Peripheral	0.80	0.82

The results of the analysis of variance in Table 1 show a significant influence on the response time of the subjects, the display configuration and the interaction between exposure time and masking.
In Figs. 9 and 10 some examples of the mean errors of the perceived roll rate $\Delta\dot{\varphi}$, plotted as a function of displayed roll rate $\dot{\varphi}$ for the display configurations C and CP and exposure times Δt_{exp} = 0.1 and 0.4 sec. are shown.
The perception error is increasing with increasing absolute magnitude of the perceived roll rate, resulting in a typical V-shaped curve.
Decreasing exposure times tend to shift these curves upwards to larger values of $\Delta\dot{\varphi}$.

6. DISCUSSION AND CONCLUSIONS

Rate perception, blanking and masking

From the results it appears that roll rate can be perceived from the central as well as from the peripheral displays. However, responses for the peripheral displays are faster than for the central display. This fact emerges from the shorter exposure time Δt_{exp} needed for accurate roll rate perception. In more general terms this can be expressed by the shorter time constant in the expression of chapter 5. In addition the response times RT_P and RT_{CP} are shorter than the response time RT_C.
Based on Ref. 6 it is concluded that the changes in response time result from the changes in the perception stage of the information processing due to the changes in the display configuration.
The influence of masking on the perception process seems to be rather complicated. The score parameter increases for the shortest exposure times. Thus roll rate perception becomes less accurate. The response times are hardly influenced. The time constant τ as defined in eq. (5.1) for the peripheral display- and central- and peripheral display configuration are increasing due to masking.
The relative increase of the score parameter for the peripheral displays S_{cp} is larger than for the central display. The increase of the score and its standard deviation due to masking corresponds with the greater uncertainty in the perception process. Based on these results an 'iconic memory' for rate perception seems to be possible.
Due to the differences in minimum exposure time needed, the shorter response time and larger influence of masking, it can be concluded that the perception of roll rate from peripheral displays and central display are to be considered as different processes.

Do subjects combine redundant sources of information?

From the foregoing it is obvious that subjects use and combine information from both kinds of displays. It is easily derived (see Refs. 9 and 10) that if σ_z^2 is the variance of an optimal estimation of z derived from different and independent observations z_1 and z_2, characterized by $\sigma_{z_1}^2$ and $\sigma_{z_2}^2$ respectively, that the following relation holds:

$$\frac{1}{\sigma_z^2} = \frac{1}{\sigma_{z_1}^2} + \frac{1}{\sigma_{z_2}^2} \tag{6.1}$$

Using the score parameter, the corresponding expression would become:

$$\frac{1}{(\hat{S}_c)_{CP}} = \frac{1}{(S_c)_C} + \frac{1}{(S_c)_P} \tag{6.2}$$

In Fig. 11 the score $\hat{S}_c$ calculated according to (6.2) has been plotted as a function of the measurement score $S_{c_{CP}}$. The result of the masking condition with exposure times 0.1 and 0.15 sec. were not used, due to the fact that the score parameters $S_{c_{CP}}$ and S_{c_C} were almost equal and the response times RT_{CP} and RT_C were almost equal. This indicates that the subjects did not use peripheral information.
It appears that:

$$\hat{S}_{c_{CP}} = 0.0061 + 0.70\ S_{c_{CP}}$$

From the results now available it is apparent that subjects indeed combine the available information to a more accurate perceived roll rate.
That the experimental results do not fit the relation given by eq. (6.1) may be due to the fact that subjects either do not 'optimally' combine independent sources of information or that the information derived from the different displays is not strictly independent.

7. REFERENCES

1. Moriarty, T.E., Junker, A.M. and Price D.R.
   Roll axis tracking improvement resulting from peripheral vision motion cues.
   Proceedings of the Twelfth Annual Conference on Manual Control, NASA TM-X-73, 170, 1976.

2. Stapleford,R.L., Peters, R.A. and Alex, F.R.
   Experiments and a model for pilot dynamics with visual and motion inputs.
   NASA CR-1325, 1969.

3. Hosman, R.J.A.W. and Van der Vaart, J.C.
   Effects of visual and vestibular motion perception on control task performance.
   Proceedings of the First European Annual Conference on Human Decision Making and Manual Control, Delft, 1981.

4. Hosman, R.J.A.W. and Van der Vaart, J.C.
   Effects of vestibular and visual motion perception on task performance.
   Acta Psychologica 48 (1981) 271-287.

5. Hosman, R.J.A.W. and Van der Vaart, J.C.
   Accuracy of visually perceived roll angle and roll rate using an artificial horizon and peripheral displays.
   Proceedings of the Second European Annual Conference on Human Decision Making and Manual Control, Bonn, 1982.
   Also in: Delft University of Technology, Department of Aerospace Engineering, Report LR-393, 1983.

6. Frowijn, H.W.
   Selective drug effects on information processing. Thesis.
   Institute of Perception TNO Soesterberg, The Netherlands, 1981.

7. Kuil, R.J. and Reid, L.D.
   Description of a test facility to investigate human internal noise sources using a digital display and keyboard.
   Delft University of Technology, Department of Aerospace Engineering, Memorandum M-399, 1981.

8. Sperling, G.A.
   A model for visual memory tasks.
   Human Factors Vol. 5, p. 19-31, 1963.

9. Maybeck, P.S.
Stochastic models, estimation and control. Volume I.
Academic Press, New York, 1982.

10. Howard, I.P.
Human Visual Orientation.
John Wiley and Sons. New York, 1982.

Table 1: Results of the analyses of variance, on the score parameter S_c and the response time RT.

	S_c	RT
Subjects S	****	****
Exposure time Δt	****	*
Display configuration C	***	***
Blanking/Masking M	*	-
Interactions:		
Δt,S	****	****
C,S	****	****
M,S	****	****
Δt,C	****	*
Δt,M	***	***
C,M	***	**
Δt,C,S	****	****
Δt,M,S	****	**
C,M,S	-	-
Δt,C,M	****	*
Δt,C,M,S	-	***
Replications	***	-

$\alpha < 0.25$ *
$\alpha < 0.10$ **
$\alpha < 0.05$ ***
$\alpha < 0.01$ ****

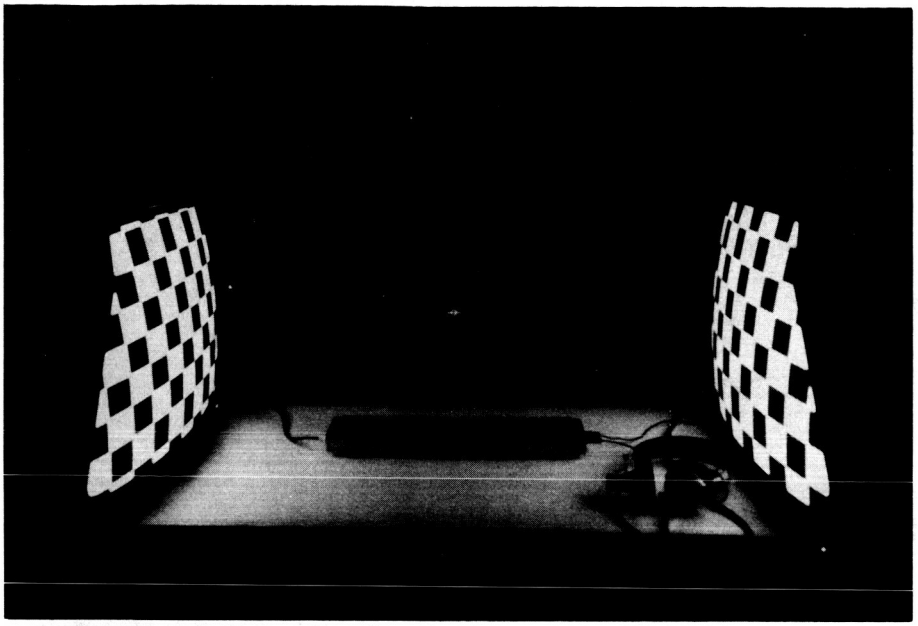


Fig. 1: Overview of the test facility showing central display, the peripheral displays and the digital keyboard.

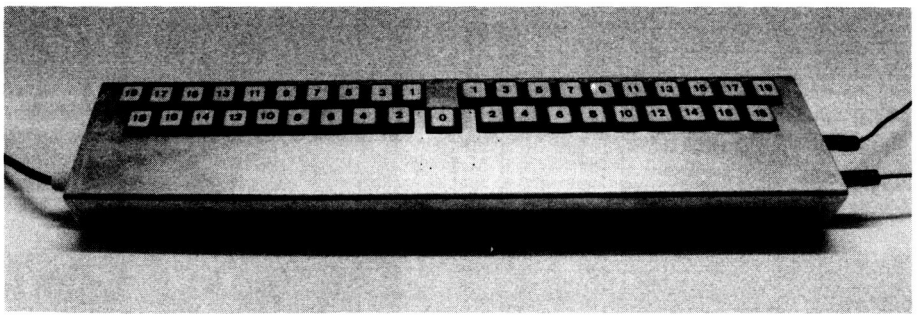


Fig. 2: Digital keyboard device.

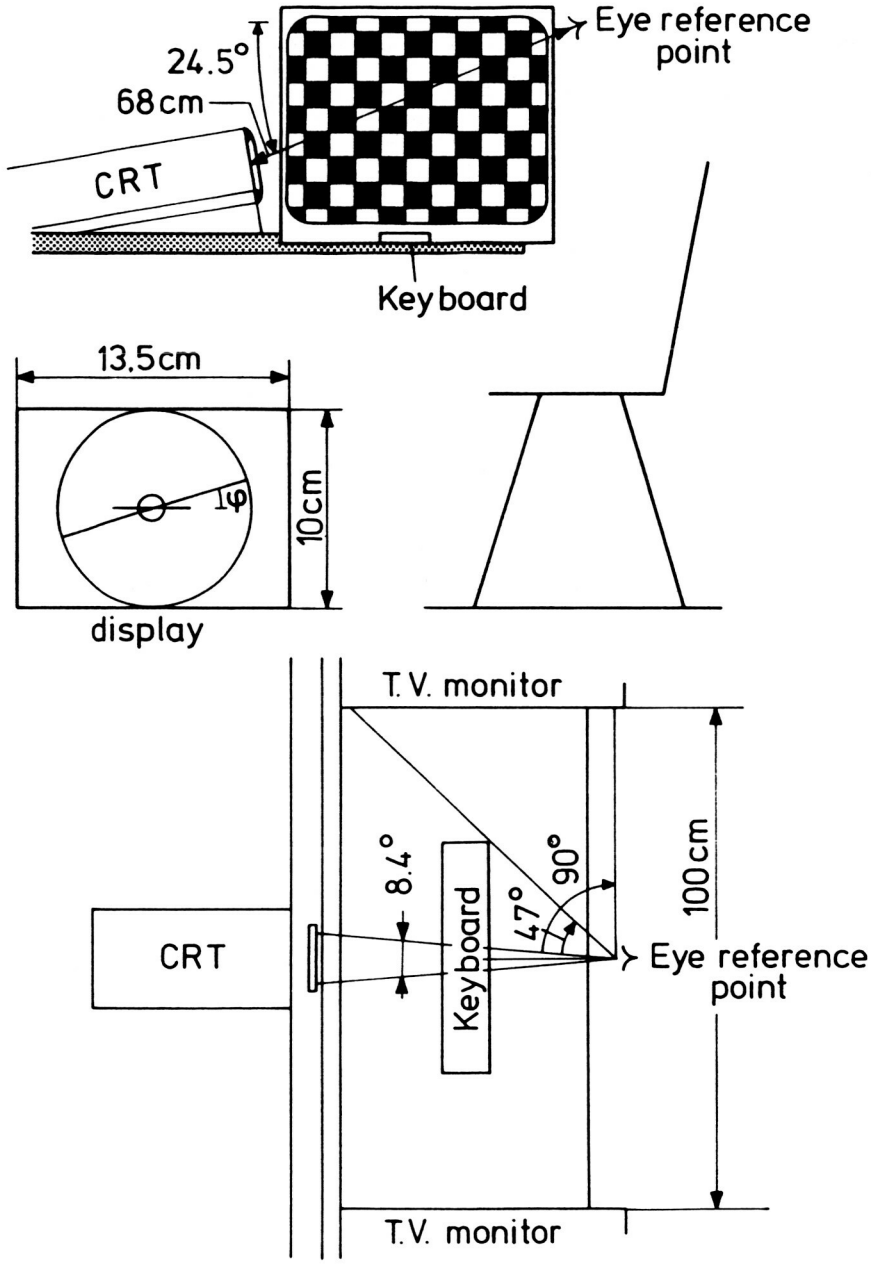


Fig. 3: Positions of displays relative to the subject's eye reference point. Central display image and dimensions.

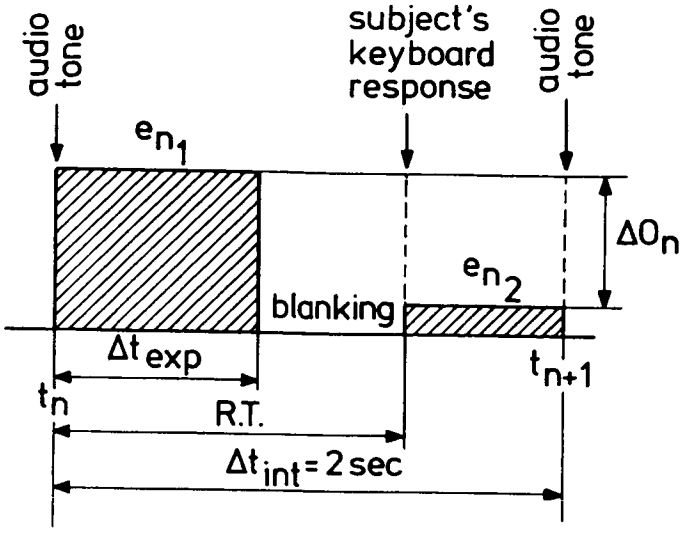


Fig. 4: Sequence during one interval of a test run.

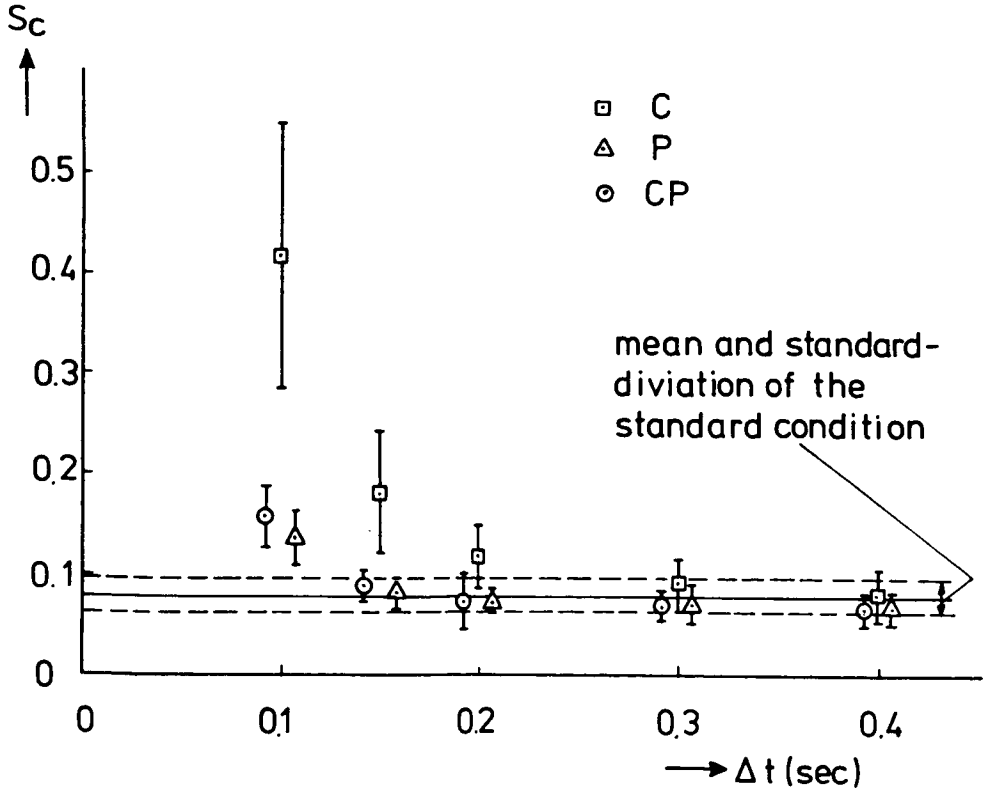


Fig. 5: Score parameter S_c for blanking as a function of exposure time Δt_{exp} and display configuration.

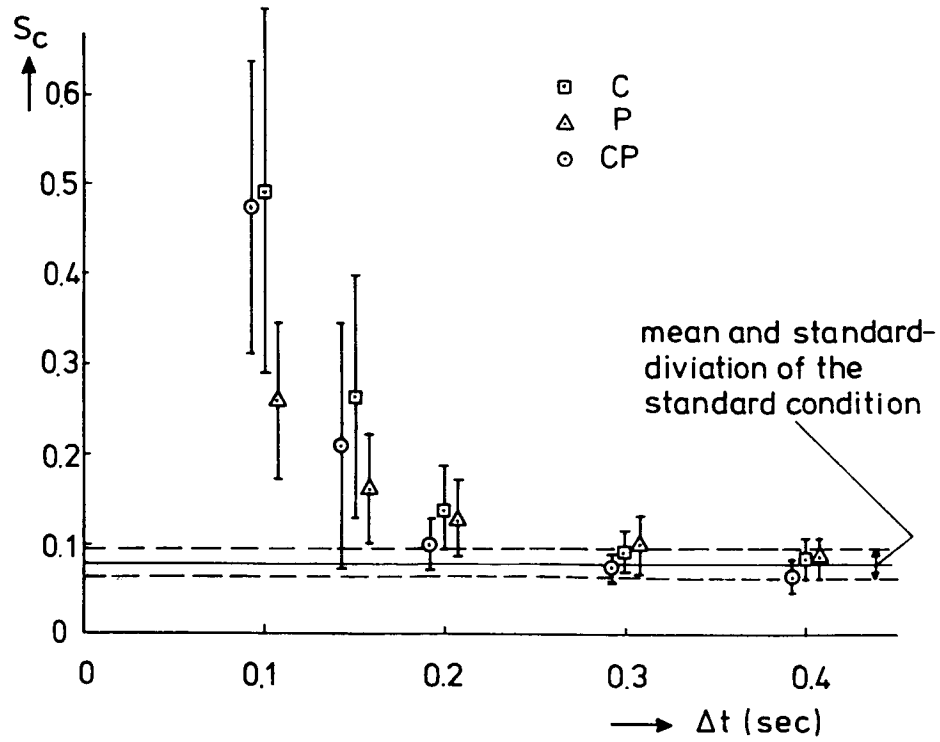


Fig. 5: Score parameter S_c for masking as a function of exposure time Δt_{exp} and display configuration.

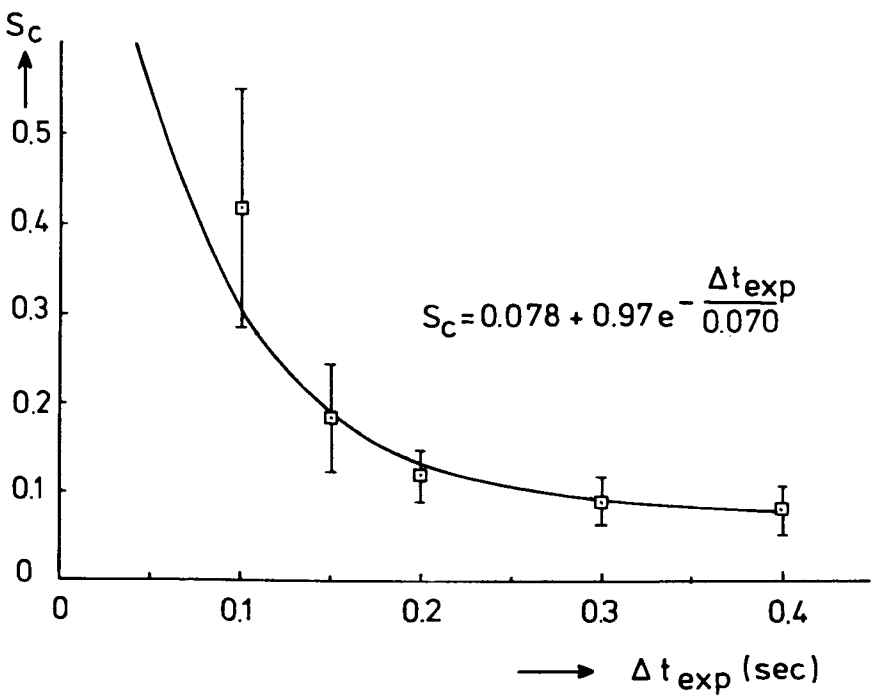


Fig. 7: The model of eq. (5.1) fitted to the data of the display configuration with centrol display only and blanking after stimulus exposure.

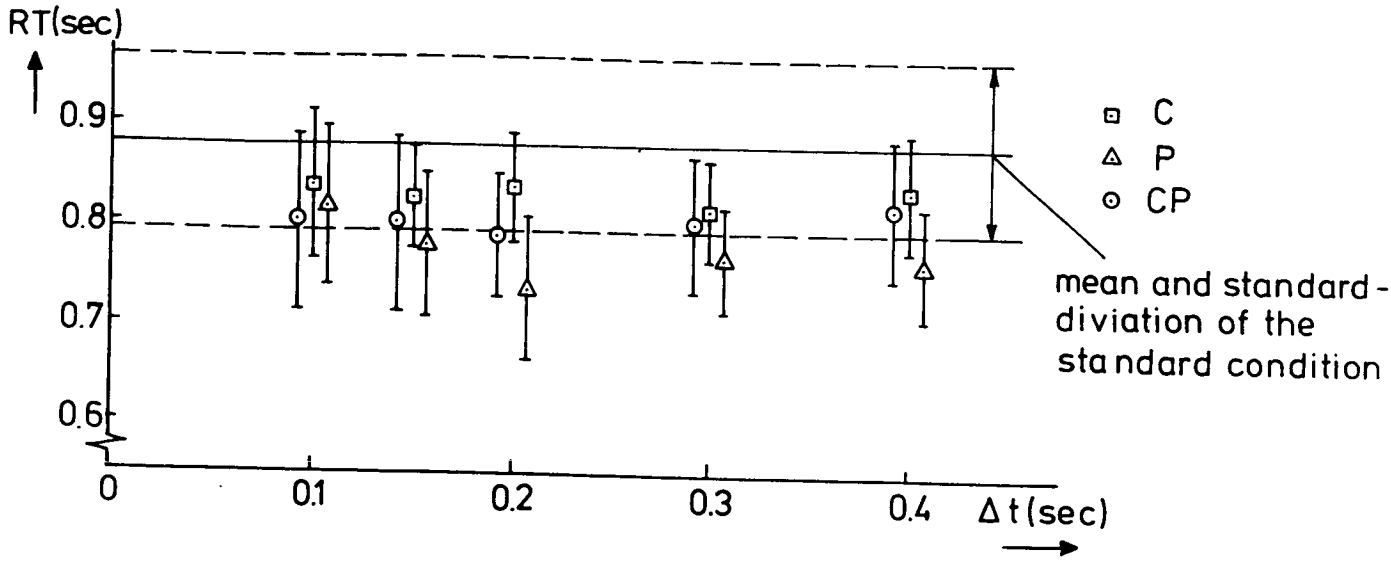


Fig. 8a: Response time to roll rate stimuli as a function of exposure time and display configuration and blanking after stimulus exposure.

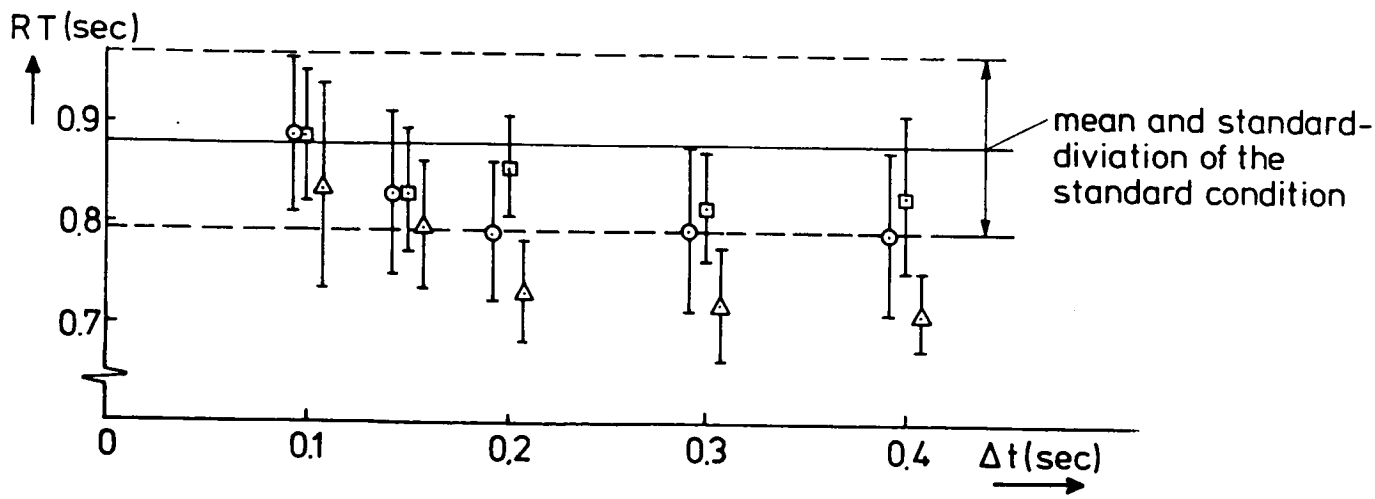
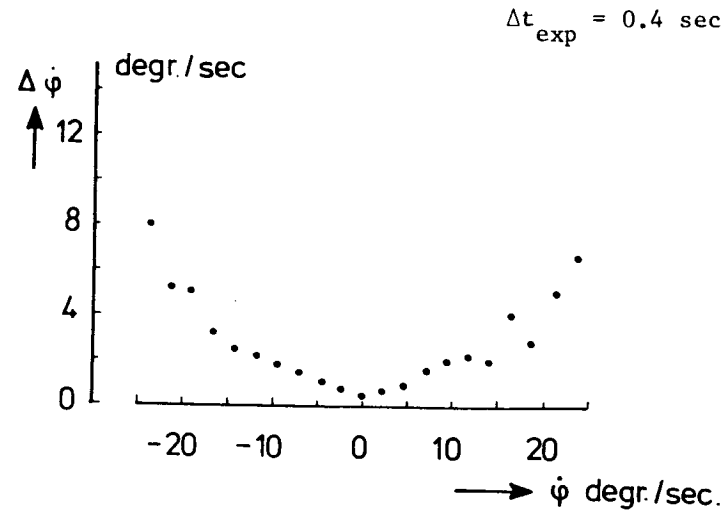


Fig. 8b: Response times to roll rate stimuli as a function of exposure time and display configuration and masking after stimulus exposure.



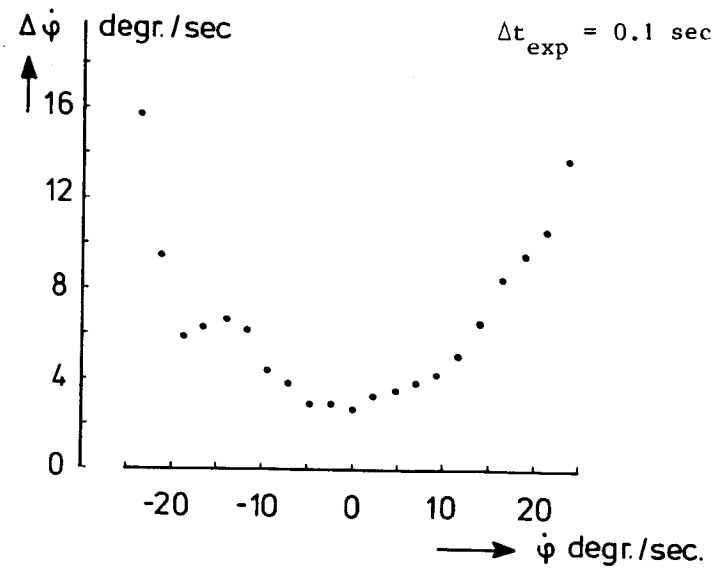
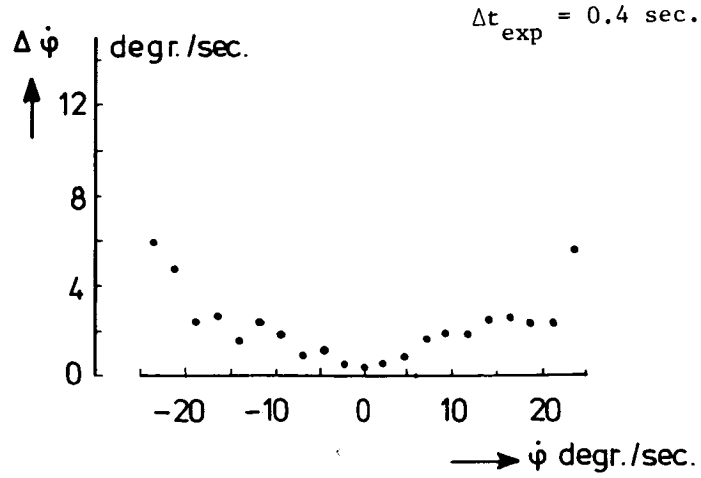


Fig. 9: Mean error $\Delta\dot{\varphi}$ of perceived roll rate as a function of displayed roll rate $\dot{\varphi}$. Central display only, blanking.



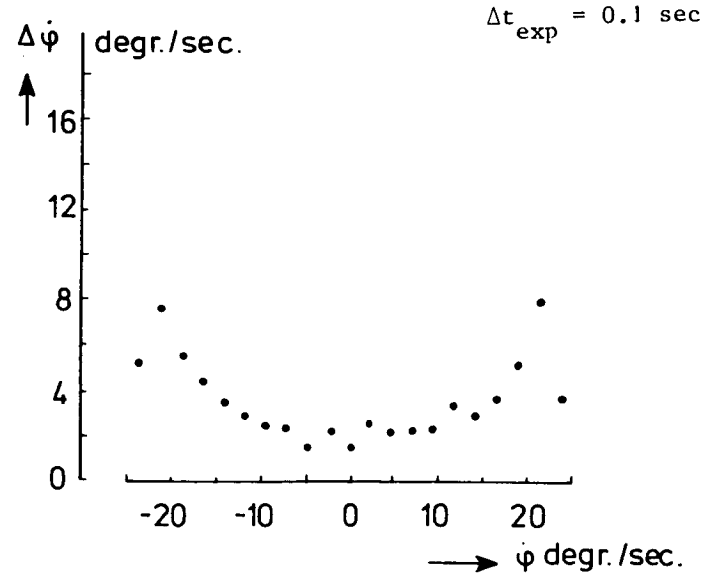


Fig. 10: Mean error $\Delta\dot{\varphi}$ of perceived roll rate as a function of displayed roll rate $\dot{\varphi}$.
Central and peripheral displays, blanking.

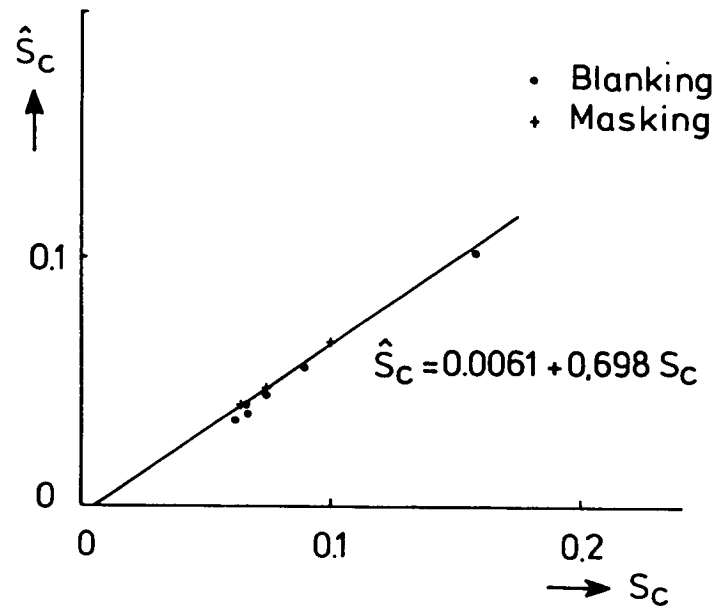


Fig. 11: The estimated score parameter $\hat{S}_c$ versus the actual score parameter S_c for the central and peripheral displays combination.

Extended abstract of paper presented in:
19th Annual Conference on Manual Control
May 1983, MIT Cambridge Mass.

Design of Optimal Motion for Flight Simulators

J. Ish-Shalom, B.C. Levy, W.M. Siebert and L.R. Young

An abstract simulator design problem is formulated as follows: given a dynamic system, Sa, called the actual system and another dynamic system, Ss, called a simulator for Sa, and given a function which drives the system Sa, the problem is to find an operator, properly constrained, which will generate the input to Ss on the basis of the input to Sa, such that the discrepancy between the outputs of Sa and Ss will be as small as possible. This abstract simulator design problem is formulated as an optimal control problem and in the linear-quadratic case this problem is decomposed into two separately solvable subproblems: (i) deterministic and (ii) stochastic. Both subproblems are solved; the stochastic one for the Gaussian case only. An examination of the properties of the solution puts in evidence a parallel decomposition theorem and provides an interpretation of the dependence of the simulator design on the given parameters. These and other properties simplify the solution to enable the extension of the solution to include several nonlinear effects.

The study of the nonlinear effects includes three topics. The first topic is an extension of the deterministic-stochastic decomposition to include nonlinear dynamic system equations using a quadratic cost function. This decomposition shows a general method of how to separate and then combine the "open-loop" (deterministic) and the "closed-loop" (stochastic) solutions for the abstract simulator design problem. This is also true for the general control problem appearing in robot control design.

The second topic is the development of a Pseudo Linear Quadratic controller PLQ for linear and nonlinear dynamic systems. The PLQ controller is derived from the standard Linear Quadratic (LQ) optimal control solution by solving for a quasi-quadratic cost and a quasi-linear system for each value of the state. This results in a feedback with a leading linear term, i.e. using feedback gains that are functions of the system state rather than constants. Though PLQ is not a solution to any known formulated optimization it is an extension of the standard LQ control. Furthermore, in the cases tested it has properties that match those of known optimal nonlinear controllers derived for linear dynamic systems using a nonquadratic cost. In contrast with these nonlinear optimal controllers, PLQ is easier to compute and easier to implement, because of its "linear" form. Many of the PLQ properties still need to be developed including conditions for global stability for the multi

dimension dynamic system case. It is expected that the resulting PLQ controller would show similar robustness properties as the LQ controller.

The third nonlinear effect explored is a sign sensitive cost formulation and solution. The cost function is put into a form that includes a correlation function term that is evaluated between the outputs of the systems Sa and Ss. It is shown that, any anti-symmetric compressive memoryless output function, cascaded to the linear dynamics of both Sa and Ss, would lead to a cost function that should include a sign sensitive term. This problem is put into an LQ form which no longer has a positive definite cost. It is shown that a unique solution exists for the abstract simulator design problem. Finally, putting all these elements together, allows one to develop a method for the design of abstract optimal simulators.

Next, the solution and properties of the abstract simulator problem are applied to the design of motion generation for moving-base flight simulators. The optimization criterion selected is a quadratic norm of the difference between the physiological outputs of the vestibular organs of a pilot in an imaginary reference airplane and those of a pilot in the simulator. Vestibular models based on physiological and psychophysical experiments were used, including consideration of vestibular sensor saturation and the multiplicative nature of the physiological noise in the nervous system, modeled by an anti-symmetric compressive memoryless output nonlinearity.

The LQ abstract simulator properties imply a 2-2-1-1 physical axis decoupling theorem for the feedback gains, i.e. Pitch-Surge, Roll-Sway, Yaw, Heave axis group decoupling. The 2 axis coupling is because of gravity. This 2-2-1-1 theorem is well known to designers of simulator motion systems. What is usually overlooked is that the feed-forward gains do not decouple the same way because of the effect of the airplane dynamics coupling. When axis transformations are included in the motion system implementation, coupling between all six physical axes is obtained -- a property not existing in current designs. An example of this effect is the proper motion generation for the falsely called "Coriolis motion sensation" which usually requires a simulator with full 360 degree rotation capabilities. The axis system handling is further extended to include measured head rotation which further improves the simulator motion sensation.

All in all, a new design method is suggested for the design of the motion of moving-base flight simulators. This design method permits to includes six new design elements: (i) combining a deterministic to the stochastic solution; (ii) a time-varying system; (iii) nonlinear system arising from a higher than quadratic cost function; (iv) a sign-sensitive cost; (v) axis transformation effect; (vi) head rotation effect.

As a demonstration of this methodology, several design examples were solved and simulated. The results conform to the set of empirically found design rules used by experienced engineers: to determine the structure (2-2-1-1 theorem), the initial setting of the pole locations, the expected

lower motion fidelity as the poles' frequency increases, cross coupling gain between the linear and rotation motion input (called g tilt).

Twenty pilots tested several of these designs which were implemented for the pitch and surge axes on a Link GAT-1 General Aviation flight simulator Trainer. These tests confirm the suggested design method, including equal weighting for the normalized vestibular linear and rotation components. Furthermore, these tests show the effects of motion on the pilot's control for a sudden unexpected flaps-down transition during level flight. Lastly, PLQ was used to build a nonlinear-motion-generation system for the Link GAT-1 simulator. This PLQ nonlinear design can be implemented easily for the full six-degrees-of-freedom case.

The examples and pilot tests which were performed are preliminary investigations into the feasibility of the optimal motion design approach for flight simulators. The results so far are promising. The causal, linear, time-invariant "optimal" motion system derived has parameters of the same order of magnitude as the conventional motion systems in use today. However, unlike these systems, the "optimal" motion system can be "tuned" by a non-expert using this computer design method to satisfy a variety of additional conditions such as: different travel lengths of the simulator, different flight trajectories, and different emphasis on motion cues. Furthermore, it uses expected future airplane motions, accounts better for hard limits by use of PLQ and takes into account axis transformations and head movements. It is simpler to carry out and as a bonus gives the control system design for the motion-base itself.

It is recommended that this design method be converted into an optimal motion system design compiler that is capable of transforming a simple minded, non-expert specification of the required motion system into a flight simulator motion-generation system. The design method obtained can also be used for model-following or robot motion design.

Reference

Ish-Shalom, J., "Design of Optimal Motion for Flight Simulators," Ph.D. Thesis, MIT, Cambridge Mass., Dec. 1982.

ABSTRACT
FOR
19th ANNUAL CONFERENCE ON MANUAL CONTROL

Title of Paper: Validation of a G-seat Roll-Axis Drive Algorithm

Authors: Edward A. Martin (ASD/ENETS, W-PAFB OH 45433, (513) 255-7325)
and Grant R. McMillan (AFAMRL/HEF, W-PAFB OH 45433, (513) 255-3325)

The results of a pilot study to evaluate the effectiveness of the G-seat drive algorithm reported by Martin and McMillan (1982) are presented. The algorithm was designed by matching buttock pressure data measured in the Roll-Axis Tracking Simulator (RATS; a whole-body motion simulator) and the g-seat, and contained both roll position and roll acceleration components. The drive equation was evaluated in a pilot study using a disturbance regulation task. Human performance was initially compared for two conditions:

(1) G-seat driven to produce scaled buttock pressures mimicking those which would be experienced in the RATS, and

(2) G-seat static.

A statistically significant improvement in human performance was found with the pressure matching algorithm (Figure 1). The fact that this difference was not large enough to suggest operational significance, however, prompted further investigation.

The pilot study was expanded to assess, independently, the contribution of the vehicle position and acceleration components of the 'pressure matching' algorithm (Figure 1). Additionally, the utility of a drive algorithm consisting solely of scaled vehicle velocity was evaluated (Figure 2). A marked improvement in performance was seen for the pure position and pure velocity algorithms with respect to the 'pressure matching' algorithm. No improvement was seen for the pure acceleration algorithm. Transfer of training to the RATS whole-body motion environment was evaluated for subjects trained with the 'pressure matching,' position, and velocity algorithms. The subjects trained with the position and velocity algorithms showed better initial performance than those trained with the 'pressure matching' algorithm.

The results of the pilot study indicate that, for the task setting employed, motion information can be effectively displayed in a g-seat using either pure position or pure velocity drive algorithms. This is evident in superior human performance both during g-seat training and after transfer to the RATS motion environment.

Reference: Martin, E. A. and McMillan, G. R., "Development of a G-seat Roll-Axis Drive Algorithm." Paper presented at the 18th Annual Conference on Manual Control, 1982.

PRESSURE MATCHING ALGORITHM RESULTS

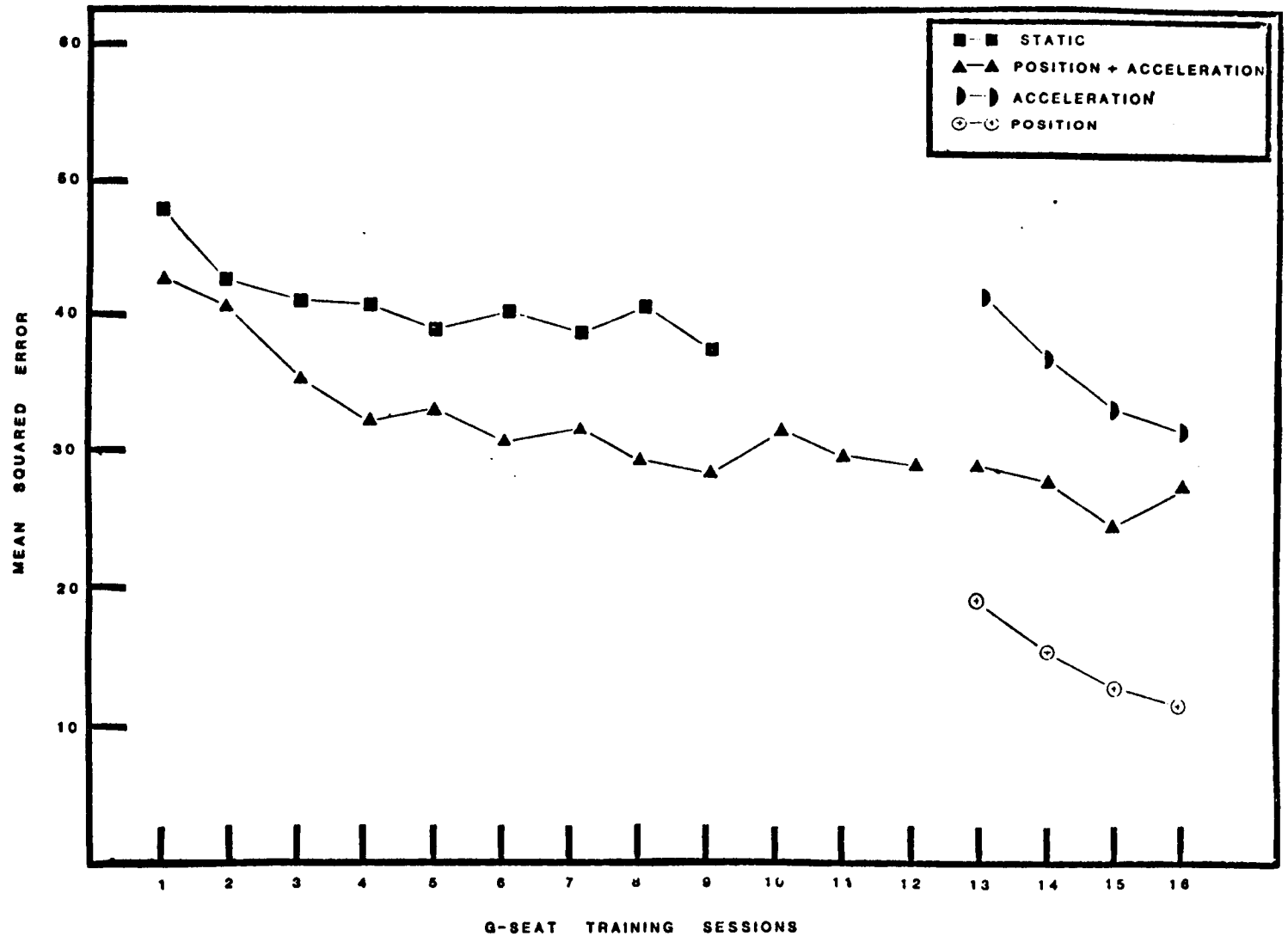


Figure 1. Tracking Error during G-seat Training with Four Levels of G-seat Cuing.

The 'Position + Acceleration' Drive Law was designed to provide a scaled match of buttocks pressures. The 'Acceleration' and 'Position' Drive Laws were each derived from the 'Position + Acceleration' Drive Law by setting the position or acceleration component, respectively, to zero.

POSITION AND VELOCITY ALGORITHM RESULTS

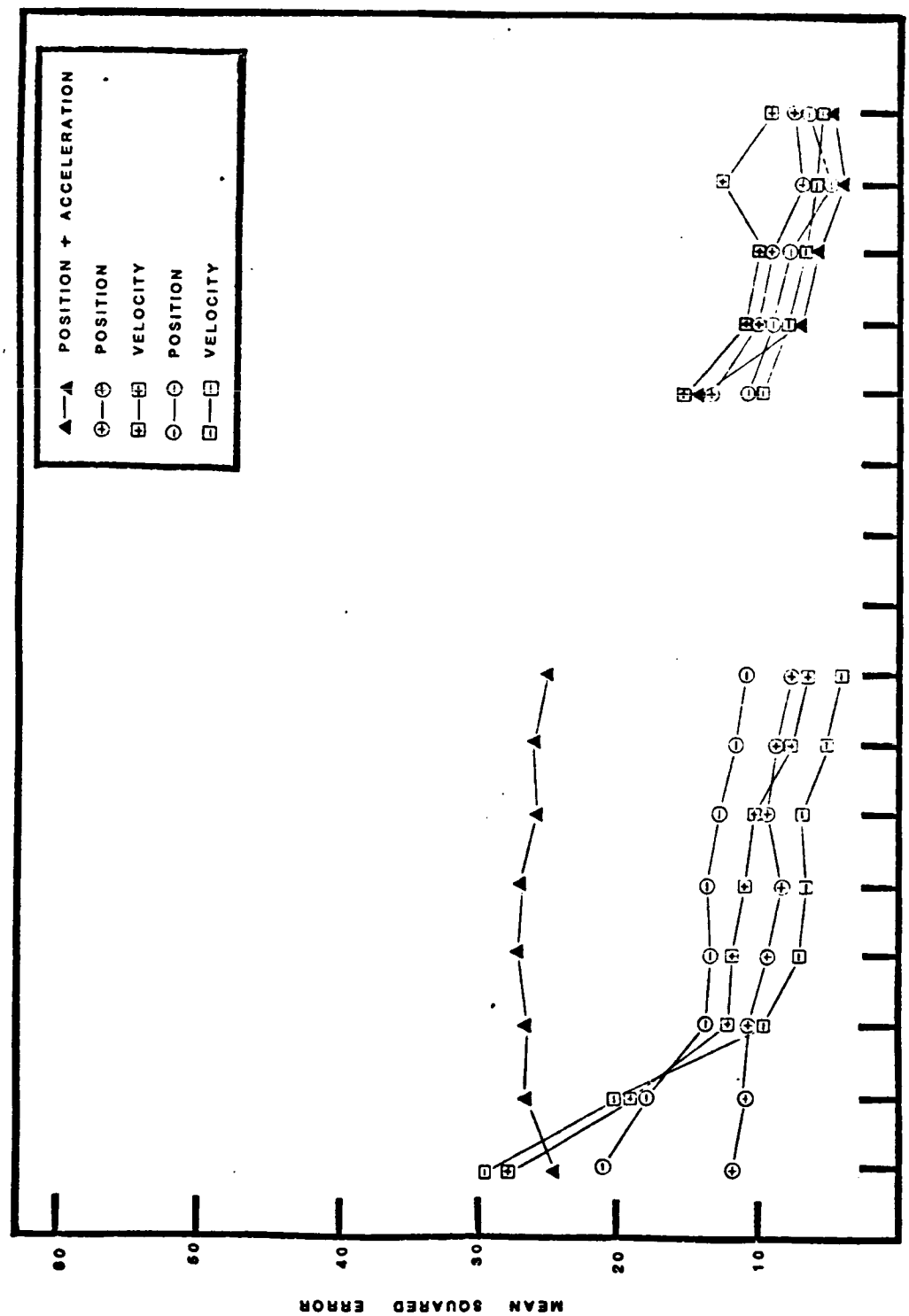


Figure 2. Tracking Error during G-seat Training and in the Whole-body Motion Simulator. Three levels of G-seat cuing were employed during training. The 'Position + Acceleration' Drive Law was designed to provide a scaled match of buttocks pressures. The seat was driven both with and opposite simulated vehicle position or velocity for the 'Position' and 'Velocity' Drive Laws.

DEVELOPMENT OF A MODEL FOR HUMAN OPERATOR LEARNING: A PROGRESS REPORT

William H. Levison
Bolt Beranek and Newman, Inc.
10 Moulton Street
Cambridge, MA 02238

ABSTRACT

The research reported in this paper was 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 flight management. Manual control data obtained in previous AFAMRL laboratory studies was analyzed with regard to learning behavior. This analysis consisted of three steps: (a) model analysis with the optimal control pilot model (OCM) to determine the relations between stages of training and independent "pilot-related" model parameters; (b) tests of some hypotheses concerning the underlying effects of training on control-strategy development; and (c) preliminary analysis to explore relationships between the perceptual cueing environment and the pilot's internalized representation ("internal model") of the task situation. The results of the analysis suggest that continued practice on the tracking task leads to a more precise, consistent, and linear (i.e., less "noisy") type of response behavior, and to an improved internal model. Analytic results further suggest that, if the OCM is modified to account for the pilot's ability to construct his internal model, we should be able to predict the effects of the task structure (including plant dynamics, input spectra, and cueing environment) on the rate at which (and degree to which) the human operator develops his estimation and control strategies.

INTRODUCTION

Objectives

Considerable savings in operator training costs can usually be realized by conducting a substantial portion of the training in a simulator, rather than in the fully operational environment. If the simulator design is focused on the task of inducing an appropriate set of operator responses to a task-relevant set of perceptual cues, and not on faithfully mimicing all aspects of the stimulus environment, further savings can be expected through the use of well-designed part-task and low-cost training devices. At present, partly because of the lack of validated, quantitative models for learning, the design of such training devices remains very much an art.

The research summarized in this study was directed toward the long-term objective of developing 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 flight management. A more specific goal was to extend the optimal control model (OCM) for pilot/vehicle systems into a predictive tool that relates the acquisition of continuous estimation and control strategies to the perceptual cueing environment. The work reported here extends the results reported previously [1], and is described more fully in [2].

Full-scale development of such a model extension was beyond the scope of this study. Rather, as a first step in this direction, a small data base concerning control-strategy learning is described and offered for use in validating models for learning, and an approach to further model development is suggested.

Criteria for Model Development

The following criteria have been adopted to guide the development of a model for control-strategy learning. Clearly, the model must account for important trends found in the relevant data base. This includes matching the effects of various aspects of the task environment on both learning rates and asymptotic performance (i.e., how fast the subject learns, and how well he learns).

Second, since the long-term goal of this study is to develop a predictive model for learning behavior, the model should have a structure and parameterization that facilitate extrapolation to new situations. Such extrapolation is enhanced if we can (1) determine a minimal set of independent model parameters that applies over the tasks of interest, and (2) separate independent parameters into task-related and operator-related categories. (Model parameterization is discussed in the following section).

Finally, model structure should be compatible with what is known about human capabilities and limitations. That is, it should "make sense".

BACKGROUND

Review of the Model

We review briefly certain key aspects of the OCM; the reader is directed to references [3,4] for further details on model structure and application.

The OCM, as used in this study, is an analytic tool that predicts statistics of operator and closed-loop system performance for a given task environment. The "input" to this model, then, consists of a full statistical description of the problem. For convenience, we categorize the input space into

parameters related to the external task environment (presumably, under the control of and/or known by the experimenter), and parameters related to human limitations (determined empirically or assumed on the basis of previous studies).

Task-related parameters include a linearized description of the tracking ("vehicle" or "plant") dynamics, an analytic representation of the tracking input spectrum, and a statement of the performance objective in terms of a weighted sum-squared error criterion. For laboratory tracking tasks in which control and display gains are idealized, the "pilot-related" parameter set usually includes (1) a white Gaussian "observation noise" process associated with each perceptual input, to account for operator response randomness (i.e., "pilot remnant"); (2) a "motor noise" process to account primarily for certain imperfections in the pilot's knowledge of system response properties; (3) a "time delay" to account for neural conduction times and other aspects of information processing that may be represented as a pure transport delay, and (4) a "motor time constant" to account for both subjective and physiologic constraints on pilot response bandwidth.

Although the observation noise processes are considered linearly independent for each perceptual quantity used by the pilot to control his vehicle, observation noise covariances appear to vary in proportion to corresponding signal variances in simple laboratory tracking tasks with idealized displays [5,6]. Thus, it is often convenient to treat the effective observation noise/signal ratio as a model parameter, rather than the individual noise covariances.

Model "outputs" are of two types: predictions of operator response strategy, and predictions of the statistics of system time histories (e.g., mean-squared error, etc.). The model structure for the operator's response strategy incorporated in the existing OCM implementations is shown in Figure 1. The "adjustable", or adaptive, portion of the response strategy is partitioned into three elements: (1) an optimal estimator to yield a least-squared-error estimate of the (delayed) system states, given the noisy delayed perceptual inputs assumed available to the operator; (2) an optimal predictor to best compensate for the operator's time delay; and (3) an optimal control or decision law to generate the appropriate response, given the operator's estimate of the system state.

Imbedded in the model structure is an "internal model" of the task environment which, in effect, accounts for all the correlations among system variables. Such an internal model is required for optimal estimation and control, and may be considered to reflect the human operator's knowledge of the task situation. For the model applications reviewed in this paper (and for most previous work as well), the internal model has been

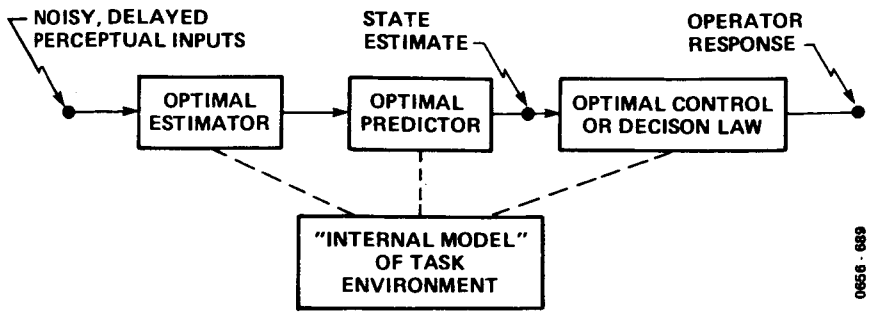


FIG. 1. CURRENT MODEL STRUCTURE FOR OPERATOR RESPONSE STRATEGY

assumed to faithfully reflect the statistics of the task environment. While this assumption appears justified for wide-band laboratory tracking tasks, we shall, during the course of this discussion, challenge its validity for tasks in which vehicle dynamics include significant lags, and/or subjects are incompletely trained.

Figure 2 shows that a fixed set of pilot-related model parameters can yield a good qualitative match to frequency response measures (pilot gain, phase, and remnant) for a significant range of laboratory tracking dynamics. (The match can be improved, of course, by "tuning" the parameters for each condition.) The reader is referred to Levison [2] and to citations contained therein for further details concerning the data base and model analysis.

If we consider a wider range of tasks, especially tasks employing dynamics of higher order and/or greater delay, we begin to see a degradation in certain "pilot" parameters with increasing task difficulty. In particular, Figure 3 shows a trend toward larger motor time constant and larger time delay with increasing effective vehicle phase lag (computed at 4 rad/sec). Detailed model analysis indicates that this trend cannot be ascribed to a lack of sensivitity of the parameter identification scheme, nor can it be ascribed to an insensitivity of tracking performance to these particular information-processing limitations [2]. Rather, such analysis suggests that the task-related parameter changes reflect limitations on the operator's information-processing capabilities that have not yet been explicitly identified (for example, the degree to which an accurate internal model of the task environment can be constructed).

Thus, the desired separation of operator- and task-related

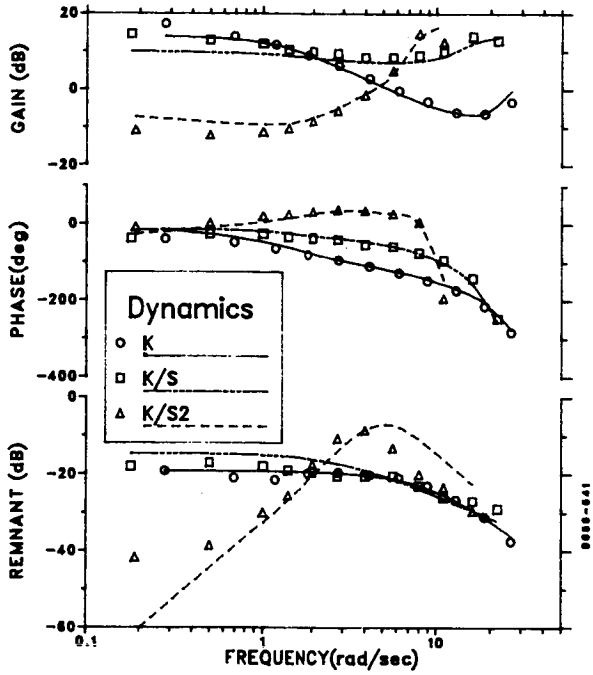


FIG. 2. FIXED-PARAMETER MATCH TO THREE TRACKING TASKS

parameters has not been fully realized. Future OCM development undertaken to account for learning behavior should address this issue.

Review of the Data Base

The data base analyzed in this study was provided by two preceding experimental studies. The study referred to as the "stable plant" study was conducted primarily to explore the effects on pilot performance of a delay between roll-axis platform motion and visual cues [7], and employed simulated vehicle dynamics representative of a high-performance fighter in roll. The study referred to as the "unstable plant" study was conducted as part of a research project to develop a tracking task for which performance would be highly sensitive to task- and environment-related stress [8], and employed plant dynamics having a divergence time constant of 0.5 seconds. The latter study was conducted fixed-base.

All subject populations were first given exposure to the tracking dynamics, without external disturbances, to allow subjects to develop an appropriate control strategy for stabilizing the plant dynamics. This initial exposure typically lasted a few minutes. A considerable period of practice followed, spread out over a number of days, in which the subjects attempted to minimize the effects of an external forcing function on mean-squared tracking error. The earliest data available for analysis of the type reported here, then, occurred early in the

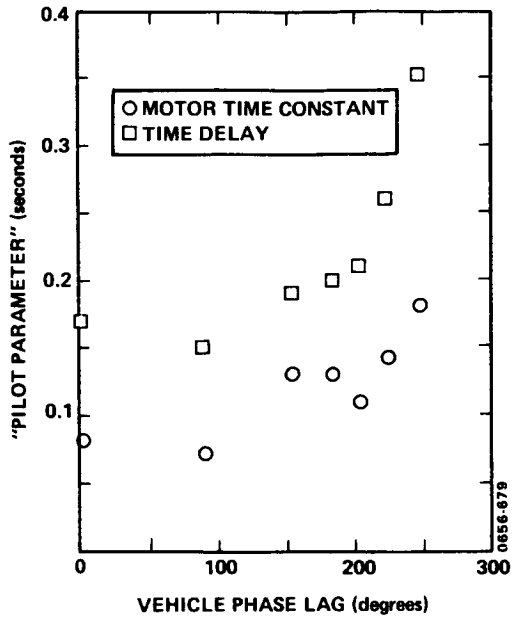


FIG. 3. EFFECTS OF TASK COMPLEXITY ON INDEPENDENT MODEL PARAMETERS

post-initialization phase; they do not represent the subjects' first exposure to the tracking dynamics.

Separate populations of 4-5 subjects were given initial training on the various conditions explored in the stable-plant study. "Learning curves" are shown in Figure 4 for the subject populations trained (a) fixed base and (b) moving-base with synchronous visual and motion cueing. Averaged mean-squared tracking error is plotted as a function of practice session, where each session consisted of four experimental trials of approximately three minutes each.

Figure 4 shows that both the fixed- and moving-base subject populations improved their tracking performance scores with continued practice, with the latter subject group apparently reaching an asymptotic performance level during the course of training. The moving-base population not only achieved lower error scores than the fixed-base population, but exponential fits to the learning curves suggest that the moving-base population reached asymptote with fewer practice sessions [7]. Learning curves for the subjects participating in the unstable-plant study showed similar trends in that tracking error scores were reduced with continued practice in an approximate exponential manner with respect to an asymptotic performance level.

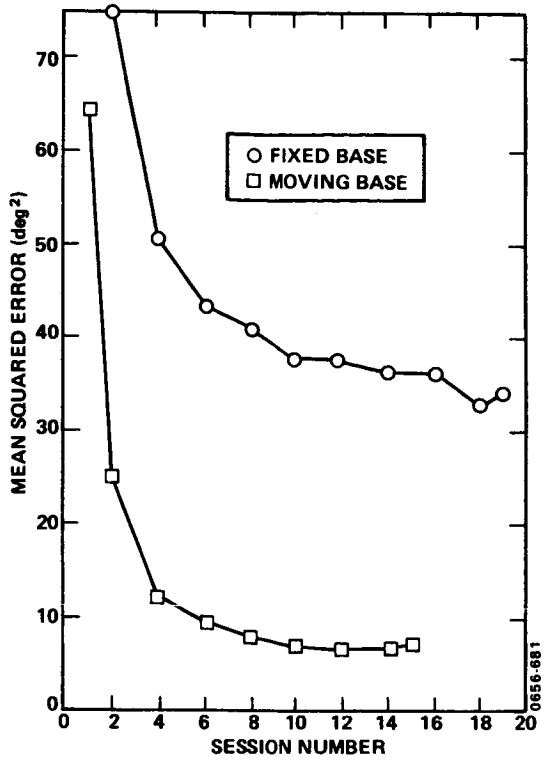


FIG. 4. EFFECTS OF PRACTICE ON MEAN-SQUARED TRACKING ERROR, STABLE PLANT

Average of 4-5 subjects, 4 trials each.

RESULTS

The three data bases described above (stable plant fixed- and moving-base, and unstable plant) were subjected to model analysis to provide a first cut at modelling learning behavior. This analysis was performed in three phases: (1) OCM parameter identification to determine the empirical relations between stages of training and the pilot-related model parameters; (2) tests of hypotheses concerning the underlying effects of practice on control-strategy development; and (3) preliminary analysis to explore relationships between various aspects of the task environment and the pilot's internal model. Principal results are summarized below; additional details are to be found in [2].

Parameter Identification

As in previous model applications of this type, model

analysis was performed not on the time histories directly, but on derived performance metrics consisting of variance scores (error, error-rate, control, and control-rate) and frequency-response measures (pilot gain, phase shift, and remnant). Two stages of practice were explored: very early in the training phase ("early"), and the end of the training phase in which near-asymptotic performance was obtained ("late").

Figure 5 shows that the three subject populations exhibited similar practice trends. Performance early in practice, relative to near-asymptotic performance, was characterized by lower pilot gain, minimal difference in phase shift, and higher remnant.

Pilot-related model parameters were identified via the quasi-Newton gradient search scheme, and practice-related parameter differences were subjected to tests for statistical significance [9,10]. Figure 6 shows the effects of practice on observation noise/signal ratio and motor time constant -- the only model parameters significantly affected by practice. For this analysis, a uniform noise/signal ratio was imposed on all observational quantities assumed utilized by the subjects (error displacement and rate for fixed-base tasks; displacement, rate, and acceleration for the moving-base task). Thus, a single observation noise parameter was identified, rather than two or three.

Practice led to a substantial and statistically significant decrease in observation noise for all three tasks. A consistent decrease was also exhibited by the motor time constant, but only the fixed-base, stable-plant task yielded a significant difference.

Experimental and modeling results reviewed here and in the background section lead to the following conclusions:

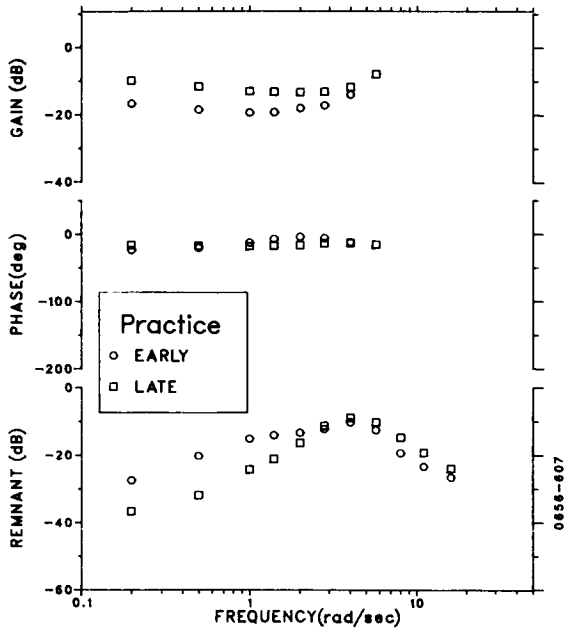
1. Time delay and motor time constant appear to increase with increasing task difficulty for trained subjects.

2. Observation noise/signal ratio and (to a lesser extent) motor time constant appear to decrease with continued practice for incompletely-trained subjects.

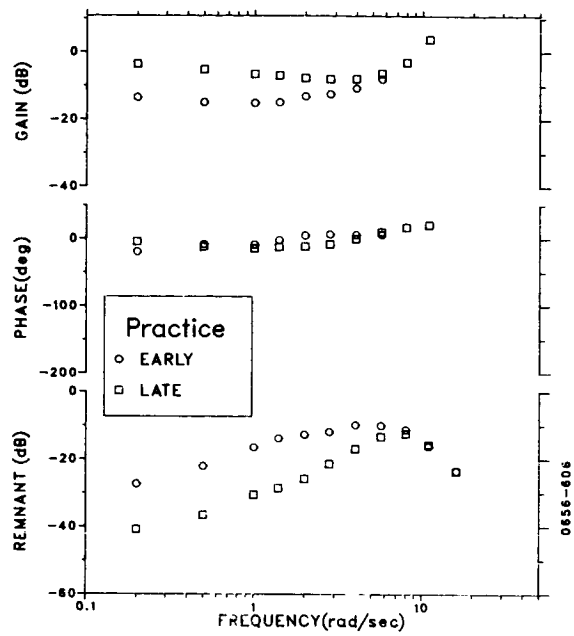
3. The availability of whole-body motion cues allows subjects to train faster and better; "faster" in that apparent asymptotic performance levels are reached with less practice time, and "better" in that lower values of observation noise and motor time constant can be achieved.

These conclusions are intended for the data base considered in this study, not as general comments.

a) Stable Plant, Fixed Base



b) Stable Plant, Moving Base



c) Unstable Plant

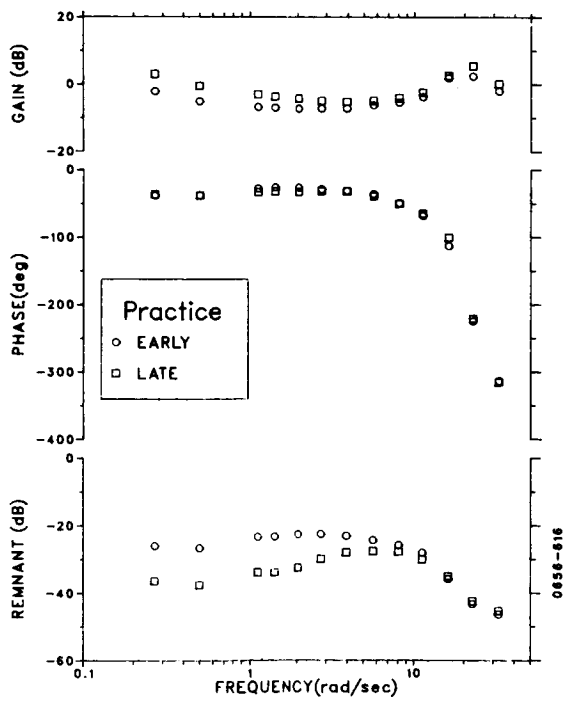


FIG. 5. EFFECTS OF PRACTICE ON AVERAGE PILOT FREQUENCY RESPONSE

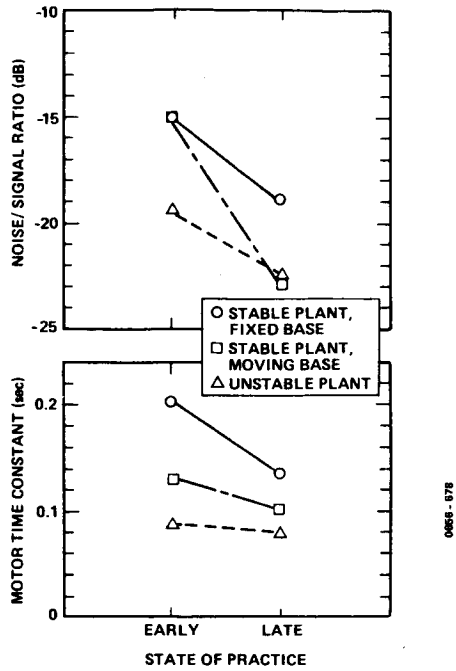


FIG. 6. EFFECTS OF PRACTICE ON INDEPENDENT MODEL PARAMETERS

Average of 4-6 subjects.
Average 2-4 trials/subject early, 4 trials late.

Hypothesis Testing

Various hypotheses were explored concerning potential information-processing deficits early in the training phase, compared to information-processing capabilities after substantial practice. Some of these hypotheses were suggested by the analysis summarized above, others by existing theories of learning. Potential deficits were first explored individually, then in combination. The goal of this analysis phase was to characterize learning effects in a manner that was parsimonious and yet applicable to a predictive model for learning.

The following tests were performed on the three subject-averaged data bases discussed above. Results were similar for all cases. To save space, only the data relevant to the unstable-plant task (the task yielding the greatest measurement bandwidth) are shown below.

Observation Noise/Signal Ratio

The results presented above suggest that observation noise differences should account for much of the practice effects. The following test was therefore performed: (1) independent model parameters were initially set to the values identified from the

late-practice data, and (2) the observation noise/signal ratio was increased until the early-practice average MSE score was matched to within 10%. A uniform noise/signal ratio was maintained for the position and rate components, and all other parameters were kept fixed at the values appropriate for late practice. The objective of this analysis, then, was to determine how well practice effects could be accounted for by a change in a single model parameter; specifically, observation noise/signal ratio.

The effects of changing the observation noise are shown in Figure 7a. In order that one may qualitatively judge the accuracy of the model trends, average frequency-response data are shown for comparison.

As the reader will shortly see, this hypothesis provides the best single-parameter explanation for observed performance trends.* Predicted gain is decreased and predicted remnant is increased, but not all details of the early/late differences are modeled. In particular, the model predicts less of a differential between low-frequency and high-frequency remnant changes, and more of an effect on high-frequency gain peak, than were observed experimentally. There also seems to be a trend for predicting too much phase lag at mid-to-high frequencies.

Lack of Information on Error Derivatives

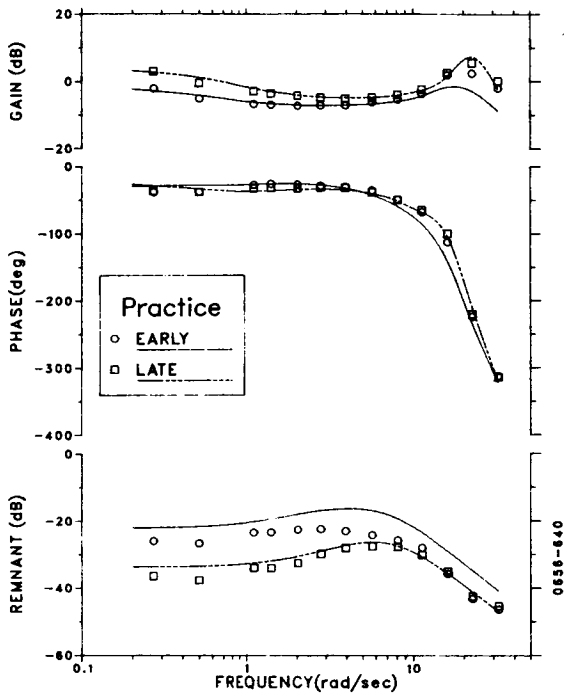
It has been suggested that an untrained operator makes relatively little use of rate information provided by the velocity of the error indicator, and that the reliance on this type of information increases with continued practice [11]. To test the validity of this hypothesis, model parameters were initially selected as in the previous test, and rate cues were eliminated from the set of assumed observations. No other parameter changes were made.

Figure 7b shows that this hypothesis produces some predicted performance trends that are counter to experimental observations. Although the effects of practice on low-frequency gain are matched reasonably well, too much high-frequency phase lag is predicted, and the predicted remnant trends are wrong. Specifically, this hypothesis predicts that remnant will be increased largely at high frequencies, whereas the data show the increase to be primarily at low frequencies. Thus, the hypothesis of a specific deficiency in extracting and utilizing velocity information early in practice is not supported.

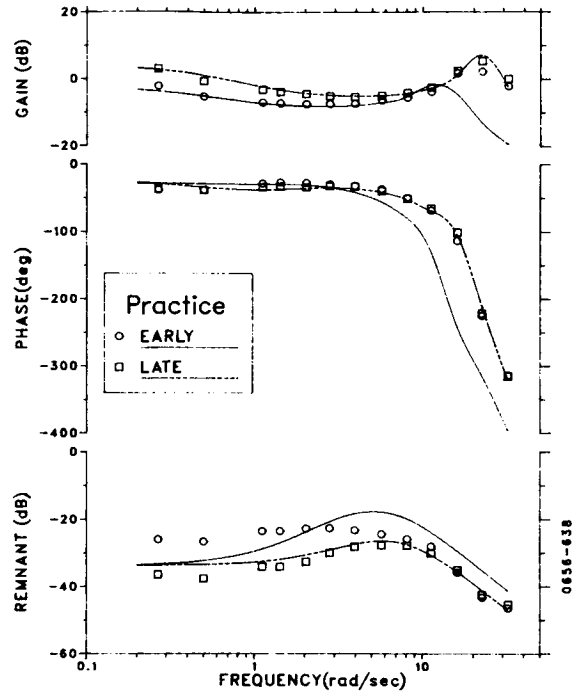
---

*Throughout this discussion, a comparison of early-practice behavior to late-practice behavior is implied.

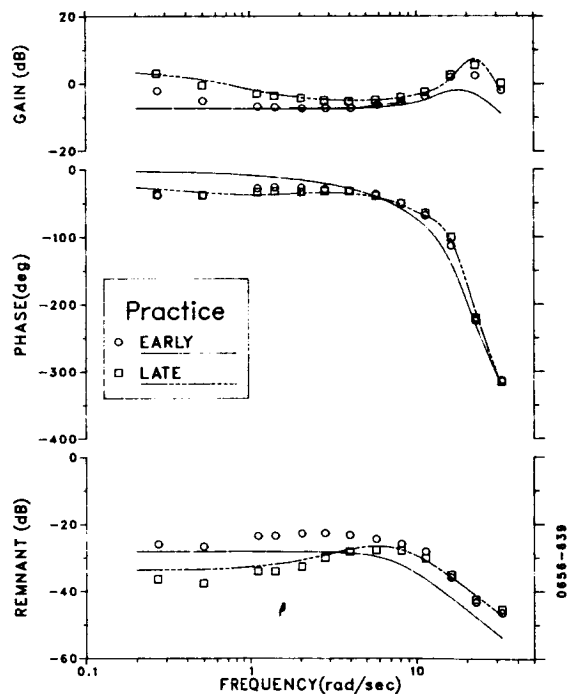
a) Observation Noise



b) Lack of Rate Information



c) Deficient Internal Model of Tracking Input



d) Observation Noise and Deficient Internal Model

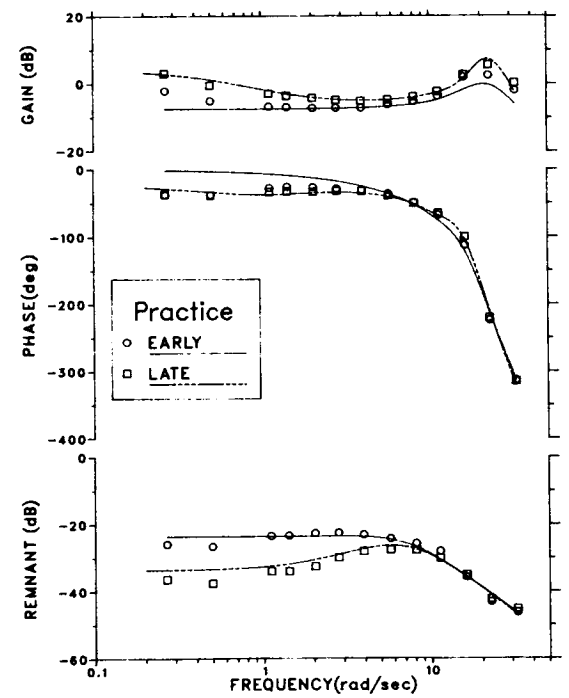


FIG. 7. RESULTS OF HYPOTHESIS TESTING, UNSTABLE-PLANT TASK

Deficient Internal Model of the Tracking Input

As noted in the Background section, the "early" results presented here do not represent the subjects' first exposure to the tracking tasks, but rather the first available opportunity to obtain consistent performance measures once the subjects have learned some rudimentary strategy for stabilizing and controlling the vehicle. Consequently, it is possible that a substantial portion of the learning taking place during the practice intervals relevant to these data bases was devoted to learning the statistics of the external inputs. This hypothesis is consistent with the "Sensory Organization of Perception" theory [12,13], in which control-strategy development is assumed to undergo three stages: (1) development of "compensatory" skills, in which appropriate feedback laws are established for stabilization and control; (2) development of "pursuit" skills, in which sufficient knowledge is gained to allow application of a partial feedforward strategy; and (3) a "precognitive" mode in which full advantage is taken of any predictability of the external inputs.

In keeping with these notions, a third hypothesis was tested; namely, that performance deficiencies early in practice can be accounted for by a deficient internal model of the input statistics. To test this hypothesis, the (mathematical) "pilot" was assumed to have no knowledge of the correlations inherent in the forcing function. A white-noise (internal) model of the input was assumed, and the covariance of this white noise was adjusted to match early-practice MSE scores, with other pilot-related parameters fixed at values appropriate for asymptotic tracking.

Figure 7c shows that practice effects on pilot gain are matched over much of the spectrum. However, low-frequency phase effects (not seen in the data) are predicted, and the predicted remnant spectrum is uniformly too low. Thus, the notion of a deficient internal representation of the tracking input is not sufficient as a stand-alone hypothesis to account for practice effects.

While it is reasonable to expect that the subject's internal model of the plant was also being refined during practice, model analysis was restricted to consideration of input modeling for two reasons. First, because of the mathematics involved, consideration of certain types of input-model deficiencies can be treated with considerably less computational cost than plant-model deficiencies. Second, the plant can be stabilized with a grossly deficient input model, and one has a rationale for choosing such a deficient model (e.g., the untrained subject is unaware of correlations within the input signal). On the other hand, severely deficient plant models cannot be explored because of resulting closed-loop instabilities, and there is no correspondingly obvious candidate for a plant model deficiency.

Combination of Observation Noise and Deficient Input Model

The foregoing results suggested a test of the combined hypothesis of increased observation noise and a deficient input model early in training. A gradient search was performed jointly on the internal white-noise input covariance, and the observation noise/signal ratio, to provide the best overall match to the early-practice data. Remaining parameters were fixed at values appropriate to asymptotic performance. Figure 7d shows that, on balance, a good match to the early data was achieved with this hypothesis.

Summary of the Hypothesis-Testing Exercise

This analysis phase was begun with essentially a two-parameter match to practice effects (observation noise/signal ratio and motor time constant), and it concluded with a different two-parameter match (observation noise/signal ratio and a deficient internal model).* Since either treatment is able to provide an acceptable data match, the decision concerning further model development must be guided by the remaining criteria proposed in the introduction; namely, one should pursue an approach that makes sense in terms of psychomotor capabilities and is most likely to lead to a model of predictive value.

As noted in a previous publication [1], practice-related reduction in observation noise makes intuitive sense. Since observation noise is the mathematical device by which most of the pilot's "remnant" is accounted for, time variations and nonlinearities are reflected in the observation noise. One would expect that continued practice would bring about a more linear and stable response strategy.

One might also expect continued practice to enhance the operator's overall motor capability, which might well be reflected as a practice-related reduction in the motor time constant. However, extensive analysis of the data base [2] suggests that the apparent changes in motor time constant are more directly related to the perceptual environment than to direct "motor training". For example, the subject population trained on the moving-base, stable-plant task had a lower average motor time constant early in practice than the fixed-base population had late in training. Since the subject populations

---

*That is, variations in only two parameters are required to account for early/late performance differences. To match performance in a specific condition, non-zero values are typically associated with at least four "pilot-related" parameters, given the current model structure.

were matched for tracking ability prior to the study, it is not likely that the apparent differences in motor time constants reflected inherent differences in response bandwidth capabilities. On balance, the evidence suggests that practice- and plant-related motor time constant represent differences in "perceptual efficiency" perhaps due to differences in the quality of the operator's internal model.

If we accept the notion that the subject's internal model changes with practice, we should be prepared to consider changes in the internal model of the plant as well as of the input. Furthermore, subsequent model development should address the problem of predicting how the internal model (and the observation noise/signal ratio) changes with practice, and how these practice effects are modified by the details of the task environment. These issues are addressed in a preliminary manner in the following discussion.

Preliminary Treatment of the Influence of the Task Environment on the Operator's Internal Model

One way of treating internal model development is to add a fourth adaptive element -- the "optimal identifier" -- to the OCM model structure as indicated in Figure 8. As the name implies, this model element would mimic the way in which the human operator identifies plant and input dynamics, and it would use this information to properly configure the remaining adaptive model elements.

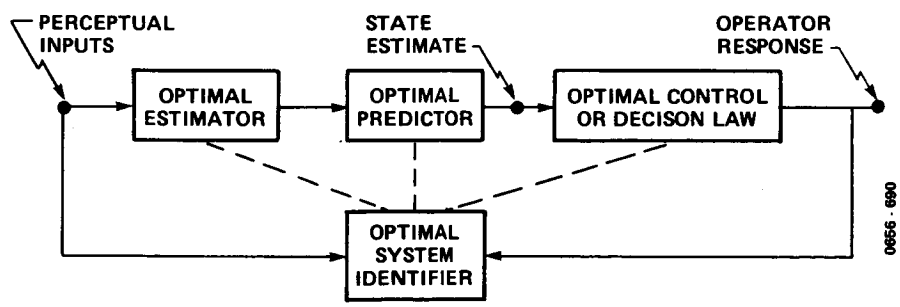


FIG. 8. PROPOSED MODEL STRUCTURE FOR OPERATOR RESPONSE STRATEGY

Because of limited resources, a highly simplified model analysis was performed: one not in keeping with the framework of the OCM, but yet sufficient to allow a feasibility test. A two-

operator task was considered, in which a pilot performed the roll-axis tracking task used in the stable-plant study referenced above, and an independent observer attempted to identify the plant dynamics. The observer performed this identification on the basis of the pilot's control input and the tracking error, and was prevented from making a perfect identification because of the effects of the tracking input (not directly available to the observer). The observer performed the identification by adjusting the parameters of a fixed-form model to reduce the mean-squared difference between "predicted" and observed tracking error.

No experimental data were used for this analysis. Instead, tracking "data" were generated by the OCM using nominal values for pilot-related parameters. Analysis was performed in the frequency domain. Parameters of the analysis included plant dynamics, perceptual cues available to the observer, and the order of the internal model adopted by the observer.

Note that the identification task posed here is the mathematical dual of the problem of identifying pilot response behavior. Additional details of this analysis are given in [2].

Because of the preliminary nature of this modeling effort, the following results are not intended as accurate predictions of a human operator's ability to identify system dynamics, but rather to explore gross trends relating internal models to elements of the task environment.

The results of four tests performed with this approximate model are shown in Figure 9. We first describe the test results and then discuss the implications with regard to pilot model development.

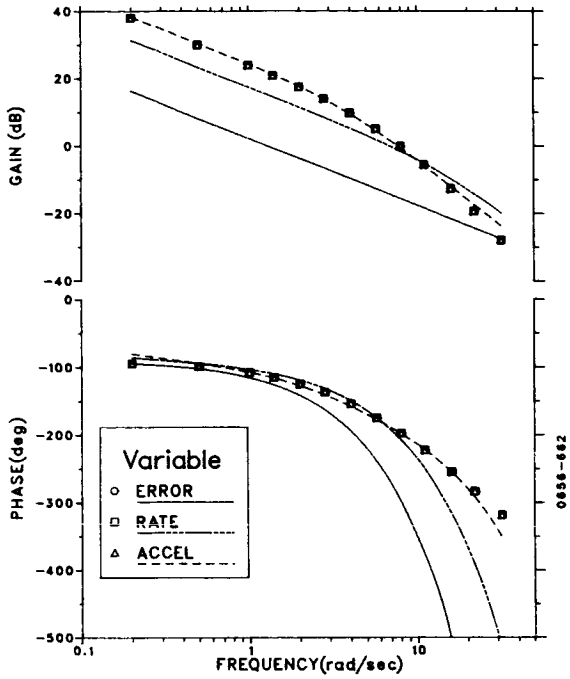
To explore the influence of the cueing environment on the operator's internal model of the plant, analysis was performed with error displacement, rate, or acceleration information assumed available to the operator.* The observer was assumed to have a full- (in this case, second-) order model of the plant. Figure 9a shows that the transfer function of the identified vehicle dynamics (continuous curves) progressively approached the true plant response (discrete points) as higher-order information was made available.

A subsequent analysis explored the interaction between the

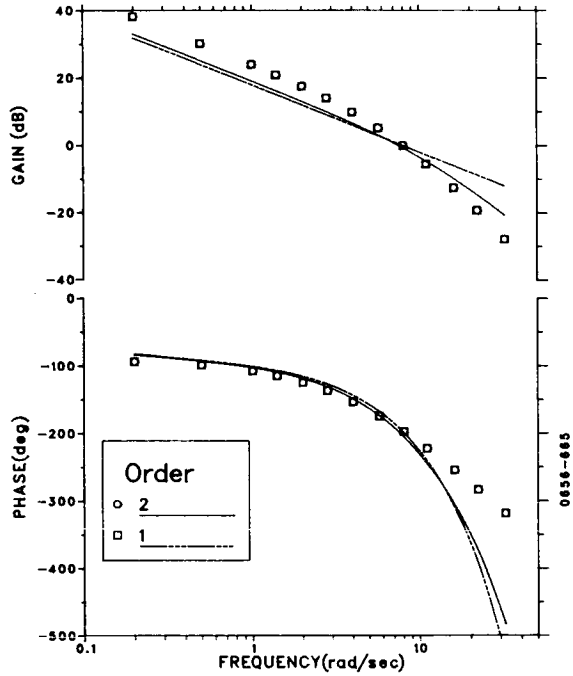
---

*For simplicity, combinations of these cues were not considered, although in practice one would expect a human operator to use all available and relevant information.

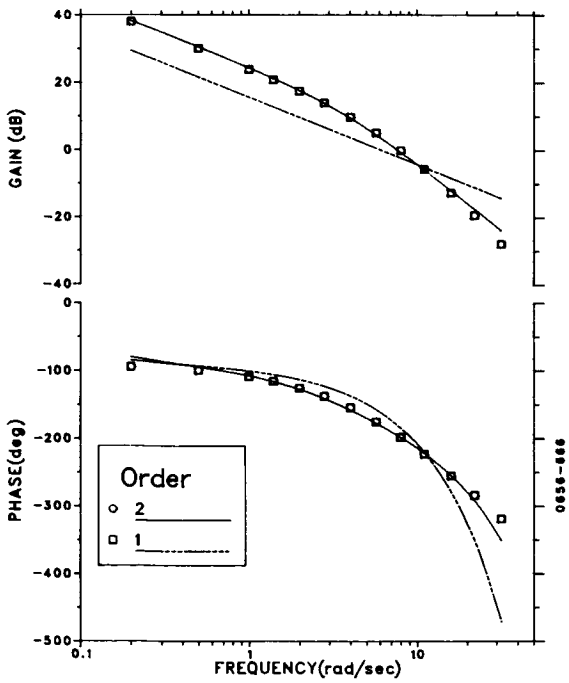
a) Effect of Primary Cue



b) Effect of Model Order, Error-Rate Cue



c) Effect of Model Order, Error Acceleration Cue



d) Effect of Plant Dynamics

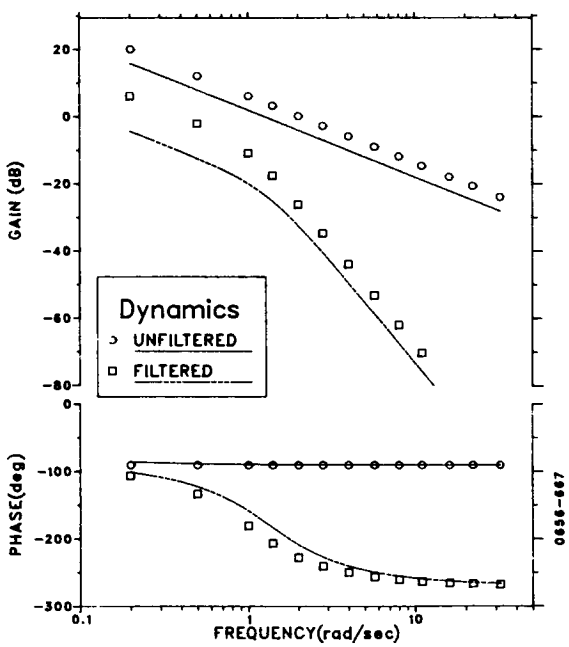


FIG. 9. INFLUENCE OF TASK ENVIRONMENT ON PREDICTED INTERNAL MODEL OF THE PLANT

available perceptual cues, the order of the observer's internal model, and the identified plant dynamics. Error-rate and error-acceleration cues were considered, as were first- and second-order plant models. Figure 9b shows that, with only error-rate available, the order of the internal model had little influence on how accurately the plant transfer function was identified (except at the highest frequencies, where signal power is relatively low). However, when acceleration cues were available (Figure 9c), the second-order plant model gave a noticeably closer match to the true plant dynamics.

A final test was performed to explore the effects of plant dynamics on the fidelity of the predicted internal plant model. Two sets of plant dynamics were considered: simple rate control (designated as "unfiltered" dynamics in Figure 9d), and rate control cascaded with a second-order Butterworth filter having a break frequency of 1 rad/sec ("filtered" dynamics). In each case, the observer was assumed to have an internal model of the appropriate order. Figure 9d shows that a qualitatively better estimate of plant dynamics was obtained for the lower-order plant.

The trends revealed by this analytical feasibility study are consistent with the experimental trends reported above: conditions which led to an improvement in the predicted internal model of the plant correspond to experimental conditions that led to an apparent increase in information-processing efficiency. Furthermore, certain cause-and-effect relationships are suggested:

1. The results of Figure 9a suggest that the reduced values for pilot-related model parameters associated with whole-body motion cueing are due to the improved internal plant model attainable in this cueing environment.

2. Figures 9b and 9c suggest that learning is more rapid in a motion-base environment partly because the operator is more readily able to determine the correct order of the plant dynamics.

3. Figure 9d supports the notion that the apparent degradation in information-processing capabilities associated with high-order plants is, in part, a reflection of a degraded internal model of the task environment.

It must be re-emphasized that the model results summarized in Figure 9 are highly qualitative and are intended only to illustrate trends. The intent of this final analysis phase has been to demonstrate the feasibility of considering deficient

internal models and of pursuing mechanisms to predict such deficiencies. The preliminary results reported here suggest that such an approach is feasible.

CONCLUSIONS

Experimental and model results reported in this paper have revealed the following relationships between the task environment and human operator performance:

1. Although a fixed set of values for the "pilot-related" parameters of the optimal control model (OCM) matches a substantial range of tracking tasks, the time delay and motor time constant appear to increase as the effective delay of the controlled element dynamics is increased.

2. For the motion-base task reviewed in this study, both the rate and degree of learning were apparently enhanced by the availability of whole-body motion cueing.

3. The OCM could account for practice effects by a decrease in observation noise/signal ratio, along with either a decrease in motor time constant or an improvement in the internal model of the tracking input statistics.

Preliminary exploration of techniques for predicting the operator's internal model of the plant were encouraging in that manipulations of the task environment that enhanced the information-processing capabilities of the experimental subjects also yielded superior correspondence between the predicted internal model of the plant and the actual plant dynamics. It is therefore suggested that the OCM be enhanced by the addition of an adaptive element, the function of which is to reflect the human operator's task of identifying the dynamical properties of both the controlled element and the external input(s). If properly formulated, the resulting model should be able to address issues not only of operator skill development, but also of performance difficulties (such as PIO's) that arise when substantial lags and delays are introduced into the system.

ACKNOWLEDGEMENT

This work was supported under Air Force Contract No. F33615-81-C-Ø517.

REFERENCES

1. Levison, W.H., "Effects of Practice on Pilot Response Behavior", Proc. of the Eighteenth Annual Conference on Manual Control, Dayton, OH, June 8-1Ø, 1982.

2. Levison, W.H., "Development of a Model for Human Operator Learning in Continuous Estimation and Control Tasks", Report No. 5331, Bolt Beranek and Newman, Inc., Cambridge, MA (in preparation).

3. 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.

4. Levison, W.H., "The Optimal Control Model for the Human Operator: Theory, Validation, and Application", in Frazier, M.L. and Crombie, R.B., (eds): Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics, AFFTC-TR-82-5, Air Force Flight Test Center, Edwards Air Force Base, CA, May 1982.

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

6. Levison, W.H., "The Effects of Display Gain and Signal Bandwidth on Human Controller Remnant", AMRL-TR-70-93, Aerospace Medical Research laboratory, Wright Patterson Air Force Base, OH, March 1971.

7. Levison, W.H., Lancraft, R.E. and Junker, A.M., "Effects of Simulator Delays on Performance and Learning in a Roll-Axis Tracking Task", Proceedings of the Fifteenth Annual Conf. on Manual Control, AFFDL-TR-79-3134, Air Force Flight Dynamics Laboratory, Wright Patterson Air Force Base, November 1979.

8. Zacharias, G.L. and W.H. Levison, "A Performance Analyzer for Identifying Changes in Human Operator Tracking Strategies", AMRL-TR-79-17, Aerospace Medical Research Laboratory, Wright Patterson Air Force Base, March 1979.

9. Levison, W.H., "A Quasi-Newton Procedure for Identifying Pilot-Related Parameters of the Optimal Control Model", Proceedings of the Seventeenth Annual Conf. on Manual Control, Los Angeles, CA, June 1981.

10. Levison, W.H., "Effects of Whole-Body Motion Simulation on Flight Skill Development", Report No. 4645, Bolt Beranek and Newman, Inc., Cambridge, MA, October 1981, ADA 111-115.

11. Fuchs, A.H., "The Progression-Regression Hypotheses in

Perceptual-Motor Skill Learning", Journal of Experimental Psychology, 63, 1962, pp. 177-182.

12. Krendel, E.S. and McRuer, D.T., "A Servomechanisms Approach to Skill Development", J. Franklin Inst., pp. 24-42, January 1960.

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

CONTRIBUTIONS OF MYOTATIC MECHANISMS TO LOAD COMPENSATION

Gyan C. Agarwal, Gerald L. Gottlieb and Robert J. Jaeger

College of Engineering
University of Illinois at Chicago
P. O. Box 4348
Chicago, Illinois 60680

and

Department of Physiology
Rush Medical College
1753 West Congress Parkway
Chicago, Illinois 60612

Abstract

Sudden torque perturbations of the foot were imposed by a torque motor while blood flow to the lower leg was occluded by a sphygmomanometer cuff applied to the thigh of seated subjects. Subjects were instructed to resist the torques and to make voluntary ankle flexions to restore original position or to go beyond the starting position to a given target. The footplate system was controlled by an electronic servosystem to act like a moderately damped, stiff spring. By making one out of every three perturbations a zero amplitude pulse, this paradigm allowed us to record intermixed stretch evoked and visually evoked contractions before ischemia and after myotatic reflex had vanished.

Visually triggered EMG responses recorded from soleus and tibialis anterior muscles were less affected by ischemia than torque evoked responses. When the earliest (myotatic) response has been totally abolished, the later (postmyotatic) response had usually reduced. In contrast, the visually triggered response was usually unchanged or had increased. By comparing these responses, possible contributions made by myotatic mechanisms to the tasks of load compensation are evaluated.

INTRODUCTION

Since the early quantitative descriptions of the behavior of the myotatic reflex (30), this physiological system has been extensively and systematically studied (8,9,19,23). The reasons for such interest lie, in part, in the unique and fascinating anatomy of the structures underlying it. There is also a presumption that the reflex is physiologically, that is behaviorally, important to the normal functioning of the motor system.

The classical notion of a reflex is derived from processes that are both automatic and understandably useful or homeostatic in the sense addressed by Cannon (6). Many motor reflexes fall into this category such as those mediating nociceptive responses and crossed extension. The myotatic reflex is not like these, however. Although in the spastic quadruped it can provide a simulation of normal posture, it does not do this in humans.

The teleological problem posed by the myotatic reflex is its failure to provide demonstrable length-regulating or load-compensating properties to the normal limb. This failure was first noted by Hammond (21) and later by Melvill Jones and Watt (29). If that is not its role, then what role has

it? Several alternatives have been offered (2,14,24) and recently summarized by Stein (33).

In earlier papers (13-16) we have described responses to ankle perturbations observed in plantar-dorsal muscles during various voluntary actions involving those muscles. Our observations have all been based upon the assumption that the physiologically meaningful stimulus was mechanical and could be quantified by evaluating the rate of ankle rotation. On this assumption, we integrated electromyographic responses seen at various latencies after the onset of perturbation and took the ratio of electrical response to mechanical stimulus as a measure of the gain of segments of the motor system.

While the use of electromyographic information from contracting muscles is of enormous utility in many ways, it remains ultimately unsatisfying. That is because the physiologically significant output variable of muscle is mechanical tension but the dynamic relationship between its electrical and contractile responses has not been established for most physiological conditions (1).

In order to determine what, if any, mechanical advantages derive from the myotatic reflex, we wish to evaluate the peformance of the limb both with normal reflex function and in its absence. Although it is easy, by a number of means, to abolish the reflex, it is probably impossible to do only this and alter nothing else in the system. Nevertheless, we have attempted to deafferent the limb by reversible ischemic hypoxia induced by a sphygmomanometer cuff.

Interruption of the flow of blood to a limb leads to hypoxia of the peripheral tissues and the accumulation of metabolic byproducts. The effects of hypoxia begin distally and progress proximally in both sensory and motor fibers (26). Large fibers are affected before small ones and, of particular importance, sensory fibers are affected before motor fibers (28,32).

Tendon jerks and Hoffmann reflexes can both be abolished by ischemia (5,26,28,30). Magladery et al (28) showed that with the stimulus distal to the cuff, the M-wave could be retained and virtually unaffected at a time when the H-wave had vanished. Moddell (30) placed the cuff distal to the politeal fossa and estimated from the M-wave that when the tendon jerk was abolished, an average of 75% of the alpha fibers were still conducting.

Kots (26) studied the effects of cuff-produced occlusion on the responses to electrical stimulation both above and below the cuff. For either electrode position, he found a progressive decrease in the M-wave. The conclusions drawn were that nerve conduction decreased and eventually failed due to ischemia, central motor neuron excitability initially increased, and primary afferents were more sensitive to the ischemia than their companion efferents.

It has also been shown that the effects of a local anesthetic block at the ankle does not alter postmyotatic reactions to SOL stretch (7). Bawa and McKenzie made similar observations at the wrist (4). Hence we assume that it is the blockage of muscle afferents rather than joint or distal cutaneous receptors that is critical in these experiments.

The experiments to be described here confirm the earlier findings that myotatic reflexes are abolished by ischemic hypoxia while a considerable degree of efferent conduction remains. The Hoffmann reflex is also abolished but, when evoked proximal to the cuff, outlasts the myotatic reflex. Our new

findings concern the attenuation of the postmyotatic responses to joint perturbation and the accompanying functional, mechanical consequences of these myoelectric deficits.

METHODS

Our basic procedures remain as described in earlier papers (13, 18). One second long pulses of torque were applied to dorsiflex or plantarflex the ankle of a seated human subject. Torque, foot angle and electromyograms (EMGs) from the soleus (SOL) and anterior tibial (TA) muscles recorded with disk surface electrodes were digitized and recorded by computer. The EMGs were amplified 1000X, full wave rectified and demodulated by a third-order averaging filter (11). The subjects were instructed to respond to the perturbation by restoring the foot to its initial angle as rapidly as possible.

In most of the experiments reported here, an 8 inch wide sphygmomanometer cuff was wrapped around the thigh, about 8 inches above the knee. A series of torque pulses was delivered to obtain a control response and the cuff was then inflated to 150 mm Hg. Higher pressures can be used but we found this to be unnecessary and considerably more uncomfortable to the subject. Torque pulses were then applied at a rate of about 10 per minute, with a one minute rest period out of every five, until the myotatic reflex had disappeared in the EMG response, as monitored on an oscilloscope. This took from 18 to 25 minuntes.

We would continue to collect data for about five more minutes. By the end of this period, light touch sensation was lost from the foot and reduced above the ankle. Subjects could make quick, brief voluntary plantarflexions and dorsiflexions of the foot throughout but could maintain a contraction for more than a few seconds only with great effort and muscular discomfort. (The greatest discomfort was always in the contracting muscles, not at the cuff itself.)

After collecting sufficient data in the areflexic state, the cuff was abruptly deflated. Reflexes and sensation were promptly restored in less than 30 seconds and testing continued. About one minute after deflation, a period of hyperesthesia ensued in which the torque pulse evoked a powerful, ringing "pins and needles" sensation. This frequently resulted in terminating the experiment. The hyperesthesia lasted for about one minute, after which the subject felt and functioned as before the experiment.

The effects of ischemia on the responses in this experiment were quantified electrically and mechanically. We have previously shown that integrating the rectified EMG from a responding muscle could provide a measure which is highly and linearly correlated with the rate of muscle stretching (13, 14). The slope of the EMG-rate regression line provides a measure of the gain of part of the control loop. This slope can be measured by applying perturbations of different amplitudes. Because this line goes through the origin, it can also be estimated from the simple ratio of integrated EMG to rate of joint rotation. Measuring this ratio as it changed over time provided a myoelectrical measure of ischemic effects upon reflex gain.

To evaluate the contractile deficits that accompany ischemia we chose three simple measures of performance. The subjects were instructed to restore the foot to the baseline position as rapidly as possible. We integrated the area between the foot angle and the baseline over the 200 ms after the perturbation. The more rapidly the perturbation could be arrested and

corrected, the smaller this number.

As a second evaluator we measured the peak angular rate of the foot as it was returned to the baseline by the subject. As a third measure, we computed the length of time it took the foot, after reaching its peak displacement by the torque, to return half way back to the baseline. When all load compensating mechanisms are intact, we would presumably measure the largest velocities and shortest halfway-back times.

As an alternate way of visualizing the effects of ischemia, we also plotted responses to torque perturbations in the angle-velocity plane or "phase-phase". The angular position of the foot, plotted against its instantaneous angular velocity yields characteristic trajectories for both SOL and TA stretching perturbations. Changes in these trajectories during ischemia can be clearly seen and used to estimate when various components of the load compensating reaction have their effect.

In three of our subjects, an additional series of experiments were performed in which 1.5 ms electrical pulses were applied to the posterior tibial nerve at the popliteal fossa, distal to the cuff. This evoked an Hoffmann reflex (H-reflex) or, at high stimulus intensities, a direct motor response (M-wave). Isometric foot torque was measured with the foot plate rigidly clamped. The H-reflex and M-wave were measured until the reflex vanished due to ischemia. The cuff was then deflated and measurement continued.

The above H-reflex and stretch experiments were separate procedures. They were not done simultaneously because we were not confident of maintaining a stable position of the stimulating electrode during the vigorous perturbations of the foot by the torque motor.

Two additional experiments were performed with the cuff just below the knee so that electrical stimulation could be given proximal to the cuff. One supramaximal electrical stimulus (120 v) and one stimulus to evoke a maximum H-reflex (50-60 v) were delivered between sequences of ten torque pulses. Maximal responses were evoked and measured to reduce the effects which might be caused by slight movement of the stimulating electrodes. These stimuli were alternated until the myotatic reflex had vanished.

Lastly, three subjects were tested with torque pulses using a slightly different procedure than previously described. Our footplate system was controlled by an electronic servosystem to act like a moderately damped, stiff spring. The subjects were instructed to make voluntary ankle flexions of about 10° on receiving a visual stimulus. This stimulus was given simultaneously with the torque pulse. Because the latency of the postmyotatic reaction is significantly shorter than visual reactions by 70-150 ms (14,25) the visual target task was assumed to have only small affect on the postmyotatic response. In one experiment, the difference between those two situations was studied. By making one out of every three perturbations a zero amplitude pulse, this paradigm allowed us to record intermixed stretch evoked and visually evoked contractions to see if they evolved differently during ischemia. To further minimize the effects of fatigue, testing was done before ischemia began, after 15 minutes of ischemia with the subject sitting quietly and after the myotatic reflex had vanished. Loss of the myotatic reflex was determined by monitoring the achilles tendon jerk.

These experiments were performed on five normal, stoic adults (four male,

one female). They were approved by the Medical Center committee on Human Investigation.

RESULTS

Inflation of an occluding thigh cuff has little immediate effect upon either the reflex or voluntary responses to a perturbing torque. Figure 1 shows one experiment in which torque pulses at eight different amplitudes, four plantarflexing and four dorsiflexing, were delivered. Average angular displacement and EMG from the stretched muscle are shown at successive intervals following cuff inflation. The bottom set shows responses two minutes after deflation of the cuff.

The most striking feature of the EMG records is the disappearance of all myotatic reflex activity sometime between 18 and 25 minutes after inflation. This is equally true of both muscles although it is less apparent in TA than in SOL. At 25 minuntes, the postyotatic response remains with unchanged latency but attenuated. Rapid recovery of both myotatic and postmyotatic responses takes place after deflation.

Abolition of the myotatic reflex was a uniform finding in every subject studied. The effects of ischemia upon the later responses were more varied. This is shown in Figure 2 where we show changes in the ratio of the EMG (integrated over selected intervals) to the rate of ankle rotation (measured 20 ms after torque onset) as a function of the duration of ischemia. Three integration intervals were examined; the myotatic reflex intervals (40-100 ms), the postmyotatic response interval (120-250 ms) and a late interval (400-600 ms) during which the foot is stabilized against the torque after returning to the original position. This late EMG is correlated with initial velocity as a consequence of initial velocity and perturbing torque being correlated by laws of mechanics. Hence the late EMG is also correlated with the postmyotatic response but this represents their dependence on a common stimulus (torque) rather than any direct relationship between them.

The results for four subjects in Figure 2 show that by the time the myotatic reflex had vanished, the postmyotatic response was greatly dimished in amplitude in 2 of the 4 subjects. At this time, the late stabilizing EMG was also reduced.

To quantify the mechanical response, we measured 1) the area between the foot angle record and the baseline over a 200 ms interval after perturbation, 2) the peak angular velocity of the foot as it returned to the baseline after the perturbation, and 3) the time it took for the foot to return halfway back to the baseline from its peak displacement. In all cases, when the reflex was lost, area and time to return increased and the peak rate of return decreased. This is shown in Table 1.

Another method of demonstrating the mechanical consequences of areflexia is to look at phase-plane plots of the foot's trajectory immediately after cuff inflation and after loss of the myotatic reflex. Figure 3 shows three examples. The inserts show the angle and EMG records versus time.

In Figure 3a, peak angular velocity is produced by the perturbing torque at about 55 ms and peak angular displacement at about 110 ms with the reflex intact. A myotatic reflex which occurs in the EMG at about 40 ms, quickly returns the foot about 40% of the way back to the initial position where it briefly pauses. For this subject, the postmyotatic response appears in the EMG at about 150 ms and at 190 ms moves the foot on towards the center

position.

Without the myotatic reflex, the foot's trajectory is identical for the first 70 ms. It is displaced slightly further (0.5 degrees or 14%) and remains very near its peak displacement until 190 ms when the return begins. This latency is appropriate to the development of muscle tension by the post-myotatic response. Cusps in the trajectory, denoting pauses in the return to the origin, occur with the same latency in both cases.

In Figure 3b analagous results from TA stretch are shown. The two trajectories start to diverge most clearly after about 70 ms. In both cases the foot reaches its maximum excursion at 140 ms but has been displaced further in the areflexic case. The TA muscle and its reflexes are much weaker than SOL so the actual return of the foot is accomplished after 140 ms by the postmyotatic response. In the ischemic case this response is weaker.

In Figure 3c slightly different results are seen. Subject GCA differs from RJJ by having his foot displaced significantly less when the reflex was functioning than when it was not. This probably results from small changes in the degree of resting muscle tone. Tone increases the muscle's viscoelastic resistance to stretching. The muscle's tone diminishes when the myotatic reflex vanishes. Starting from a point of maximum displacement, however, a significant portion of the return is generated by the myotatic reflex when it is present. The postmyotatic response accomplishes the rest, with its movement beginning at about 150 ms for this subject. In both reflexic and areflexic cases the cusps occur with the same 150 ms latency.

One further question is whether the latency of the postmyotatic reflex is changed by the ischemic block of the spindle afferents. If these afferents provide an essential sensory input for the postmyotatic response, as they do for the myotatic reflex, latency changes seem inevitable. A delayed "postmyotatic response" would still be expected through other sensory modalities as long as the alpha motoneurons functioned.

The averaged EMG waveforms of Figures 1 and 3 do not suggest any latency changes but this can be verified only by examining latencies of single records. Figure 4 shows four histograms of postmyotatic latency for one experiment. The first histogram was collected before cuff inflation, the second during the first 10 minutes after inflation, the third during the second 10 minutes period after inflation and the last during the final 8 minute period of data collection after the myotatic reflex had vanished from the oscilloscope monitor. These histograms give no evidence of increased postmyotatic latency, an observation also made by Allum et al (3) in a similar experiment.

HOFFMANN REFLEX EXPERIMENTS: In order to correlate our experiments with those of earlier investigations, we also measured H-reflexes by electrically stimulating the posterior tibial nerve at a site below the cuff. Stimulus intensity was adjusted manually and stepped sequentially from just below threshold to supramaximal levels in a repetitive manner. Stimuli were given at a rate of about 10 per minute. Figure 5 shows the peak-to-peak amplitudes of individual H-waves, M-waves and isometric twitches from the start of cuff inflation until after its release. All measurements were normalized relative to the largest single value (H-wave, M-wave or twitch) recorded during the experiment.

In Figure 5a we show the evolution of the H-reflex response to a 46v

stimulus and a 60v stimulus. The 46v stimulus was initially subthreshold for the H-wave. The 60v stimulus elicited a maximal H-wave and a small M-wave as well (Figure 5b). Figure 5b also tracks a maximal M-wave evoked by a supramaximal stimulus of 120v, a level at which no H-wave was present because of antidromic inhibition of the alpha motoneuron pool (12). Figure 5c shows the reflex twitches evoked at 60v and 120v.After about 6 minutes of occlusion, in confirmation of Kots (26), the excitability of the reflex arc increased so that an H-wave was elicited at 46v. This increase in spinal excitability is consistent with the increases in the myotatic reflex sometimes seen early in the ischemic period (Figure 2). By 14 minutes, this 46v response grew to about 50% of the maximal H-wave amplitude and then declined. By the 21 minute point, all H-reflexes had vanished.

The cuff was deflated after 23 minutes. This was followed by an almost immediate 80% recovery of the maximal H-wave and a rise in the 46v H-wave as well. By about 26 minutes, the maximal H-wave had returned to pre-occlusion levels and the 46v stimulus was again subthreshold.

While the excitability of the afferent limb of the reflex arc rose and then fell, little change could be observed in the efferent limb. Neither the small (60v stimulus) nor the supramaximal (120v stimulus) M-waves were affected by the occluding cuff. The 60v twitch approximated the course of the 60v H-wave, but did not entirely vanish because of the persistence of the small M-wave.

The two experiments performed with a below-the-knee cuff distal to the site of electrical stimulation found the H-reflex to outlast the myotatic reflex. At the end points of the experiments, in one subject the relative responses (compared to immediately after cuff inflation) were: myotatic reflex 19%; maximum H-reflex 108%; maximum M-wave 80%; postmyotatic responses 65%; integrated EMG at 400-600 ms was 104%. In the other subject, these values were: myotatic reflex 8%; maximum H-reflex 104%; maximum M-wave 74%; postmyotatic response 69% and integrated EMG at 400-600 ms was 200%. Our failure to completely abolish the myotatic reflex was due to the considerably greater discomfort experienced by our two subjects with this cuff location.

VISUAL REACTION EXPERIMENTS:

In order to better differentiate between efferent or afferent failures during ischemia, we also looked at the changes in voluntary contractions in response to visual triggering signals. Our hypothesis was that postmyotatic reactions in a stretched muscle are enhanced by spindle afferent input and the loss of that input during ischemia would diminish this response. In contrast, although ischemia might reduce tonic spindle activity, visually evoked responses should be less affected because they do not incorporate a stretch evoked component of which they can be deprived by deafferentation.

Figure 6 shows results from one of the three experiments in which visual and torque stimuli were intermixed. The visual reactions are plotted at the top before cuff inflation and after loss of the tendon-jerk reflex. The rightmost traces are velocity (inverted relative to angle) and torque, respectively.

For this experiment, the footplate was controlled via a feedback servo system so that the movement was resisted by a spring force. The visual trigger data have been shifted to the left in time by 140 ms for precuff and 100 ms for the second trace in the plantar contraction. For the dorsal

construction, both of these traces have been shifted by 110 ms. This shift was done so that the start of the voluntary EMG burst occurs at about the same time as the postmyotatic burst would in a stretched muscle.

In SOL, the ischemic integrated EMG burst was diminished by 24% while in TA it increased by 55%. Consistent with this, the rate of plantarflexion (and rise in plantar torque) diminished slightly while the rate of dorsiflexion increased.

The torque evoked reactions in this experiment are shown in the bottom parts of Figure 6. Only SOL had a myotatic reflex, the loss of which could clearly be identified. Both muscles show reductions in the initial postmyotatic bursts however in the areflexic state. The data also show the electrical and mechanical ability of the ischemic muscle to support contraction for durations up to several tenths of seconds.

The figures show that as predicted, visually triggered EMG responses were less affected by ischemia than torque evoked responses. This dependence varied with the latency of the EMG interval examined however. When the earliest (myotatic) response had been totally abolished by ischemia, the later (postmyotatic) response had usually been reduced. In contrast, the visually triggered response was usually unchanged or had increased. Beyond 350 ms, in an interval when the foot had reached its final position, the mean EMG was less affected. The percentage changes for our three subjects are shown in Table 2.

In a third subject, we also looked explicitly at the differences in the postmyotatic response introduced by the change in task from "restore your foot to its original position" to "restore your foot ten degrees beyond its original position." This resulted in an approximate doubling of the integrated EMG over the whole postmyotatic interval (ca 100-300 ms) when the foot was flexed the additional ten degrees. However, almost all of this increase occurred in the second half of the interval, appearing as a prolongation of the postmyotatic burst rather than an increase in its peak value.

DISCUSSION

The first conclusion to be drawn from the foregoing experiments is that both myotatic and Hoffmann reflexes (with a proximal cuff) are blocked by peripheral ischemia. Our distal cuff M-wave measurements confirm that this block occurs while most of the alpha efferents remain functional as others have found (26,30). The electrical stimulation experiments show, by the perpetuation of the M-wave, that propagation of the action potentials down the distal motor axon, across the myoneural junction and along the muscle fibers is not seriously altered. The continued ability of the muscle to produce a maximal twitch supports the relative normality of this part of the efferent pathway.

Resting muscle tone is also affected by the myotatic reflex arc. This has been shown in animals (22) using EMG measures. Our data (Figure 3) show that the deafferented human limb is more compliant to torque perturbations than a "relaxed" limb with intact reflexes. These particular figures show subjects who differ widely from each other in their normal level of tone, but all show reductions with ischemia.

The effect of ischemia on the postmyotatic response is only to reduce its amplitude, not alter its latency. Retention of an undelayed response in the

absence of spindle feedback from a stretched muscle is consistent with our previous finding that responses with this 120 ms latency could be elicited from unstretched muscles by instruction to the subject (14). This finding is in contradiction to that of Chan et al (7) who concluded that muscle stretch receptors were necessary contributors for responses of this latency. However, these latter investigators used a different form of mechanical perturbation which may account for their differing observations.

The next conclusion drawn from our data is that alterations in the limb trajectory following a torque perturbation can be seen at latencies appropriate to segmental pathway mediation. In the experiments we have performed, this change in trajectory is caused by an early muscle twitch which tends to restore the limb to its original position. The twitch does not fully accomplish this however. Postmyotatic and later responses are required. This is similar to the findings by Allum et al (3) of reductions in ischemic soleus muscle force just prior to the postmyotatic response.

Such a failure should not be taken as a sign of reflex ineffectuality. In normal humans, stretch reflexes are not likely to be called upon to control limb position except under the guiding intelligence of higher centers. We would suggest that it is more appropriate to ask if the reflex makes any contribution at all to the mechanical response and our finding is affirmative. This contribution is proportional to the magnitude of the perturbation over the range of perturbations we have examined. Since regression analysis of the EMG response to stretch (13) shows a very low threshold, at least in soleus, we would expect this effect to occur over a wide range of situations. Given the wide range in power and precision of the human motor system, even small increments in performance are potentially of great significance but may not be quantifiable by present laboratory tests. The failure of reflex mechanisms to provide gross load compensation does not exclude the possibility that they make an important but more delicate or precise contribution.

Early muscle contraction in response to imposed loads is not the only role we would like to ascribe to myotatic mechanisms. The data are also compatible with the hypothesis we proposed previously that the myotatic reflex modulates descending signals (14). In this manner, central commands are integrated with the most recent sensory information to optimize the efferent signal.

The fact that primary spindle afferents are not obligatory for a postmyotatic response does not imply that they make no contribution to it. Joint perturbations to the normally innervated limb result in a succession of bursts from the primary muscle spindle afferents (20) which will converge in the anterior horn with descending signals. Consequently, we would expect postmyotatic responses to be stronger if spindle feedback is present and that is our finding. It is also reasonable that spinal reflex mechanisms can, on their own, support some level of activity at greater than monosynaptic latency as Ghez and Shinoda (10) and Tracey et al (34) have found in spinal cats and spinal monkeys, respectively.

Our conclusion that the reduction in the postmyotatic component of the stretch-evoked response is due to a loss of afferent input is consistent with the comparisons drawn between visually and stretch evoked responses during ischemia. In experiments such as shown in Figure 6, the two kinds of responses were intermingled during the experiment. Any changes in the ability

of the efferent pathway should therefore appear in both kinds of responses in equal measure. Responses which incorporate afferent input from the spindles will however, be attenuated by loss of that component. This attenuation is clear in most postmyotatic responses after failure of the myotatic reflex. At this time, visually triggered EMG activity is even enhanced, a finding that is, in fact, to be expected during muscle fatigue.

Our conclusion can be summarized by the following simple equation.

$$\text{voluntary EMG} + C1 + C2 * f(\theta) \qquad \text{Equation 1}$$

This equation says that the EMG activity of a "voluntary" response such as the postmyotatic response, is proportional to two separate signals. The most obvious is a descending voluntary component (C1). It also has a component due to spindle afferent input $f(\theta)$, where we have assumed $f(\theta) = d\Theta/dt$ as a reasonable first approximation, but that component is modulated by a portion of the centrally originating command signal (C2).

This is a clear example of gain modulation of postmyotatic activity, quite analagous to the gain modulation that occurs with myotatic (13) and Hoffmann reflexes (18). A separate question that arises is whether C1 and C2 are independently controllable command signals. It might be argued that they are locked and must always covary. We have addressed this question previously (15) and do not think the data here are relevant to that issue. What we wish to stress however, is that the modulation of segmental gain is a real phenomenon that affects both reflex and voluntary responses to external perturbations.

Support for our hypothesis of reflex gain modulation is also provided by the two experiments in which maximal M-waves were elicited by a stimulating electrode proximal to the cuff. In both cases the postmyotatic response was reduced by a greater degree than was the maximal M-wave. We take the relative persistence of the maximal M-wave to be a measure of the continued function of the entire efferent path. The more substantial fall of the postmyotatic response may therefore also reflect reduced recruitment of motor units by all converging inputs to the anterior horn. The only excitatory input we know with confidence to be seriously attenuated is that from the muscle spindle primaries.

The two experiments performed with a distal (below-the-knee) cuff can be used to estimate the direct contribution made by the primary afferents to the postmyotatic response. Let the integrated postmyotatic EMG response to a perturbation be described by equation 2.

$$e_n = C * (1 + \delta_n A) * \alpha_n \qquad \text{Equation 2}$$

In this equation C is the central command which is determined by the perceived stimulus and the subject's intentions (e.g., if the subject chose not to react, C would be zero). The variable A is the primary afferent signal which we wish to determine. The effects of this afferent signal are altered by δ, the fraction of the Ia pathway which is functioning ($\delta = 1$ normally and $0 < \delta < 1$ under ischemic conditions). Similarly α is the fraction of the motoneuron pool which is functioning ($\alpha = 1$ normally and $0 < \alpha < 1$ response. The subscript n indicates normal conditions without ischemia.

Under ischemic conditions, the postmyotatic response is described by Equation 3, where we have merely changed the subscript.

$$e_i = C * (1 + \delta_i A) * \alpha_i \qquad \text{Equation 3}$$

The EMG response is now e_i and the ischemia has altered the functioning fractions of both the afferent (δ_i) and efferent (α_i) pathways. We assume that C and A are unchanged.

Since we are only interested in relative changes, it is convenient to define

$$R_e = e_i/e_n \quad , \quad R_\alpha = \alpha_i/\alpha_n$$

so that

$$R_e = \frac{1-\delta_i A}{1-\delta_n A} R_\alpha . \qquad \text{Equation 4}$$

Equation 4 can now be solved for A which gives

$$A = \frac{R_\alpha - R_e}{\delta_n R_e - \delta_i R_\alpha} \qquad \text{Equation 5}$$

In terms of our experimental data, R_α is measured as the ratio of the M-waves at the beginning and end of the ischemic interval; R_e is the ratio of the postmyotatic EMGs; δ_n is assumed to be unity and δ_i is the ratio of the myotatic reflexes (which were not totally abolished in these two experiments). For the two cases we have:

	Case 1	Case 2
R_a	0.80	0.74
R_e	0.65	0.69
δ_i	0.19	0.08
A	0.30	0.08

By this computation, the afferent contributions to the postmyotatic response were about 30% and 8% for the two subjects tested with the distal cuffs. Such estimates must be rather tentative given the assumptions which underlie them. Descending signals are both necessary and sufficient to produce a postmyotatic component in response to joint perturbation. Although of diminished amplitude, they appear at constant latency even in the absence of primary spindle afferents.

Our hypothesis is that spindle afferents interact with and modulate descending signals. Under this hypothesis, myotatic mechanisms have a functionally significant role at the segmental level in normal human subjects. In addition, muscle spindles provide information about muscle length which may be used at many levels of the motor hierarchy. This aspect

of spindle function requires a great deal of further study.

ACKNOWLEDGEMENTS

We would like to thank Ms. Bernice Kaufman for technical assistance with these experiments. Dr. Jaeger's current address is: Pritzker Institute of Medical Engineering, Illinois Institute of Technology, E1 Room 125, Chicago, Illinois 60616.

This work was supported, in part, by NIH Grant NS-12877 and NSF Grant ENG-7608754.

REFERENCES

1. Agarwal, G.C. and Gottlieb, G.L. Mathematical modeling and simulation of the postural control loop. Part I. CRC Critical Reviews in Biomedical Engineering 8:93-134, 1982.

2. Allum, J.H.J. Responses to load disturbances in human shoulder muscles: The hypothesis that one component is a pulse test information signal. Exp. Brain Res. 22:307-326, 1975.

3. Allum, J.H.J., Mauritz, K.-H., and Vogele, H. The mechanical effectiveness of short latency reflexes in human triceps surae muscles revealed by ischemia and vibration. Exp. Brain Res. 48: 153-156, 1982.

4. Bawa, P. and McKenzie, D.C. Contribution of joint and cutaneous afferents to longer latency responses in man. Brain. Res. 211:185-189, 1981.

5. Burke, D., Andrews C., and Ashby, P. Autogenic effects of static muscle stretch in spastic man. Arch. Neurol. 25:367-372, 1971.

6. Cannon, W.B. The Wisdom of the Body, New York: Norton, 1932.

7. Chan, C.W.Y., Melvill Jones, G. and Catchlove, R.F.H. The 'late' electromyographic response to limb displacement in man. II. Sensory origin Electroencephalog. Clin. Neurophysiol. 46:182-188, 1979.

8. Denny-Brown, D. On the nature of postural reflexes. Proc. Rpy. Soc., B. 104:252-301, 1929.

9. Desmedt, J.F. (editor), Cerebral Motor Control in Man: Long Loop Mechanisms, Progress in Clinical Neurophysiology; Volume 4; Karger, Basel, 1978.

10. Ghez, C. and Shinoda, Y. Spinal mechanisms of the functional stretch reflex. Exp. Brain Res. 32:55-68, 1978.

11. Gottlieb, G.L. and Agarwal, G.C. Filtering of electromyographic signals. Am. J. Phys. Med. 49:142-146, 1970.

12. Gottlieb, G.L. and Agarwal G.C. Extinction of the Hoffmann reflex by antidromic conduction. Electroenceph. Clin. Neurophysiol. 41:17-24, 1976.

13. Gottlieb, G.L. and Agarwal, G.C. Response to sudden torques about ankle in man: Myotatic reflex. J. Neurophysiol. 42:91-106, 1979.

14. Gottlieb, G.L. and Agarwal, G.C. Response to sudden torques about ankle in man: II Postmyotatic reaction. J. Neurophysiol. 43:86-101, 1980.

15. Gottlieb, G.L. and Agarwal, G.C. Response to sudden torques about ankle in man: III Suppression of stretch-evoked responses during phasic contraction. J. Neurophysiol. 44:233-246, 1980.

16. Gottlieb, G.L., Agarwal, G.C. and Jaeger, R.J. Response to sudden torques about ankle in man: IV A functional role of α-γ linkage. J. Neurophysiol. 46:179-190, 1981.

17. Gottlieb, G.L., Agarwal, G.C., and Jaeger, R.J. Response to sudden torques about ankle in man: V Effects of peripheral ischemia. J. Neurophysiol., 50:297-312, 1983.

18. Gottlieb, G.L., Agarwal, G.C. and Stark, L. Interactions between voluntary and postural mechanisms of the human motor system. J. Neurophysiol. 33:365-381, 1970.

19. Granit, R. The Basis of Motor Control. London, Academic Press, 1970.

20. Hagbarth, K.E., Hagglund, J.V., Wallin, E.V. and Young, R.R. Grouped spindle and electromyographic responses to abrupt wrist extension movements in man. J. Physiol. (London) 312:81-96, 1981.

21. Hammond, P.H. Involuntary activity in biceps following the sudden application of velocity to the abducted forearm. J. Physiol. (London), 127:23-25P, 1954.

22. Hnik, P., Vejsada, R., and Kasicki, J. Reflex and locomotor changes following unilateral deafferentation of rat hind limb assessed by chronic electromyography. Neuroscience 6:195-203, 1981.

23. Homma, S. (editor) Understanding the Stretch Reflex, Progress in Brain Research, Volume 44, Elsevier, 1976.

24. Houk, J.C. Regulation of stiffness by skeletal motor reflexes. Ann. Rev. Physiol. 41:99-114, 1979.

25. Jaeger, R.J., Gottlieb, G.L. and Agarwal, G.C. Myoelectric responses at flexors and estensors of human wrists to step torque perturbations. J. Neurophysiol. 48:388-402, 1982.

26. Kots, Y.A. The organization of voluntary movement, New York: Plenum Press, 1977.

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

28. Magladery, J.W., McDougal, D.B. and Stell, J. Electrophysiological studies of nerve and reflex activity in normal man II: The effects of peripheral ischemia. Bull. Johns Hopkins Hospital 86:291-312, 1950.

29. Melvill Jones, G. and Watt, D.G.D. Observations on the control of stopping and hopping movements in man. J. Physiol. (London) 291:709-727, 1971.

30. Moddell, G., Best, B. and Ashby, P. Effect of differential nerve block on inhibition of the monosynaptic reflex by vibration in man. J. Neurol. Neurosurg. Psychiat. 40:1066-1071, 1977.

31. Sherrington, C.S. The integrative action of the nervous system, New Haven: Yale University Press, 1906.

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

33. Stein, R.B. What muscle variable(s) does the nervous system control in limb movement? Behav. and Brain Sci., 5:535-577.

34. Tracey, D.J., Walmsley, B. and Brinkman, J. Long loop reflexes can be obtained in spinal monkeys. Neuroscience Letters 18:59-65, 1980.

Table 1. Relative changes in the mechanical reaction of the limb to a torque step, induced by ischemic differentiation. Each value is the ratio of the final areflexic response to precuff response (average of 10 for each). Area is the area between the deflected foot angle and the baseline computed for 200 ms after the perturbation. Peak rate is the highest angular velocity achieved by the subject on return to the original position. Time-2 is how long it took to restore foot angle one half way back from its peak deflection.

Subject	Sol Muscle			TA Muscle		
	area	peak rate	time2	area	peak rate	time2
RJ	1.63	0.57	1.83	1.05	0.47	1.92
GA	1.99	0.61	1.72	1.61	0.39	2.24
GG	1.40	0.81	1.42	1.12	0.34	2.37
BK	1.09	0.51	1.25	1.11	0.24	1.47

Table 2. Relative changes in integrated EMG induced by ischemic deafferentiation. MR - myotatic reflex (ca 40-100 ms, very small and not measured in TA). PMR - postmyotatic response (ca 120-125 ms). VTR - visually triggered response (ca 220-350 ms). TR - tonic response (ca 400-600 ms). Each value is the ratio of the final areflexic response to the precuff response.

Subject	Soleus Muscle				Anterior Tibial Muscle		
	MR	PMR	VTR	TR	PMR	VTR	TR
RJ	0.02	1.70	2.45	1.13	0.50	0.79	1.00
GA	0.01	0.94	1.64	1.45	0.47	1.19	0.70
GG	0.00	0.80	0.76	1.09	0.81	1.55	2.07

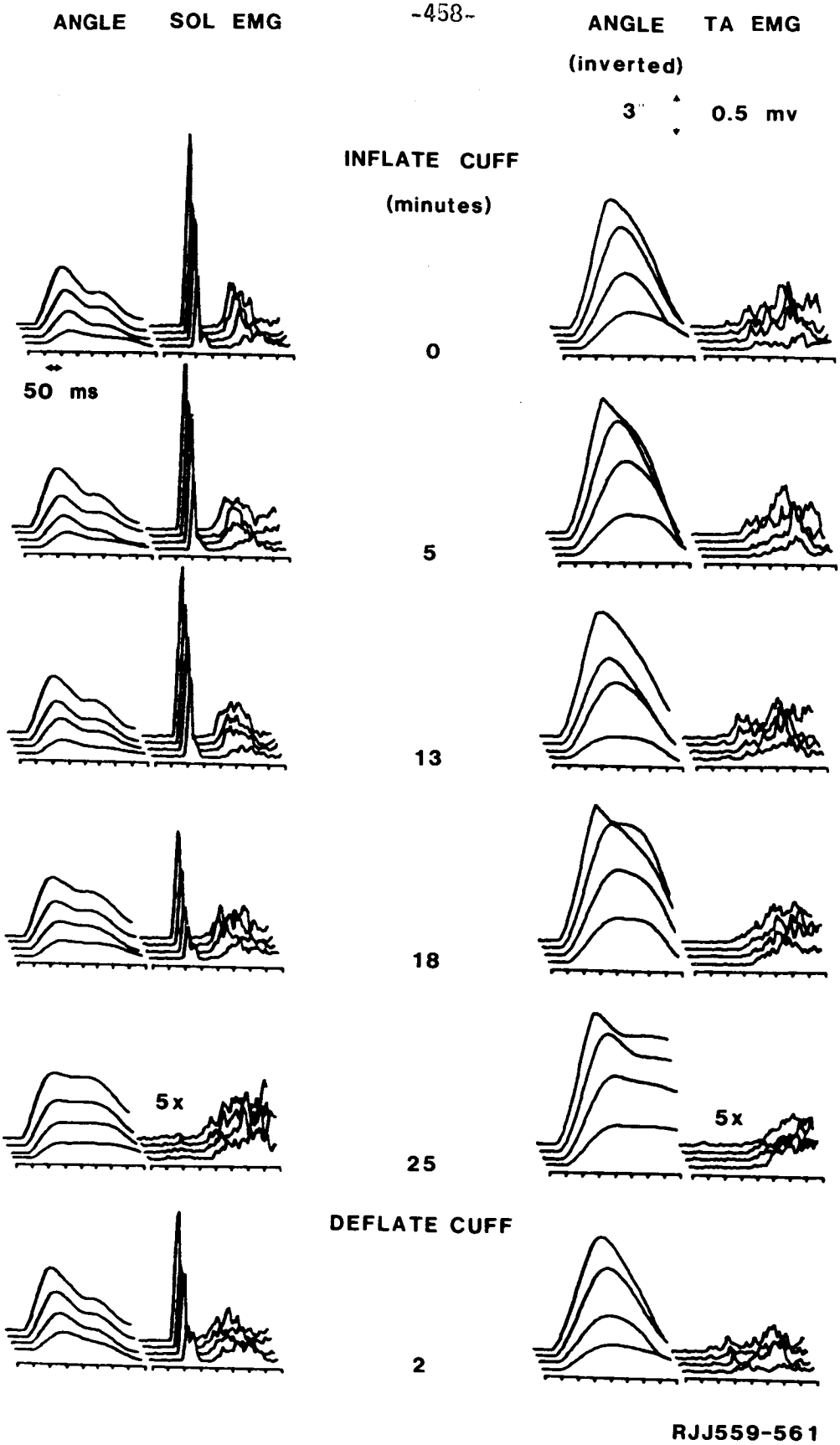


Fig. 1. Effects of progressive ischemia on responses to joint perturbation. Joint angle and EMG from the stretched muscle are plotted with responses to stronger perturbing torques plotted towards the background. Each record is an average of 10 responses. Minutes of ischemia are in the middle column. Note the EMG scale is amplified 5X for the set at 25 minutes.

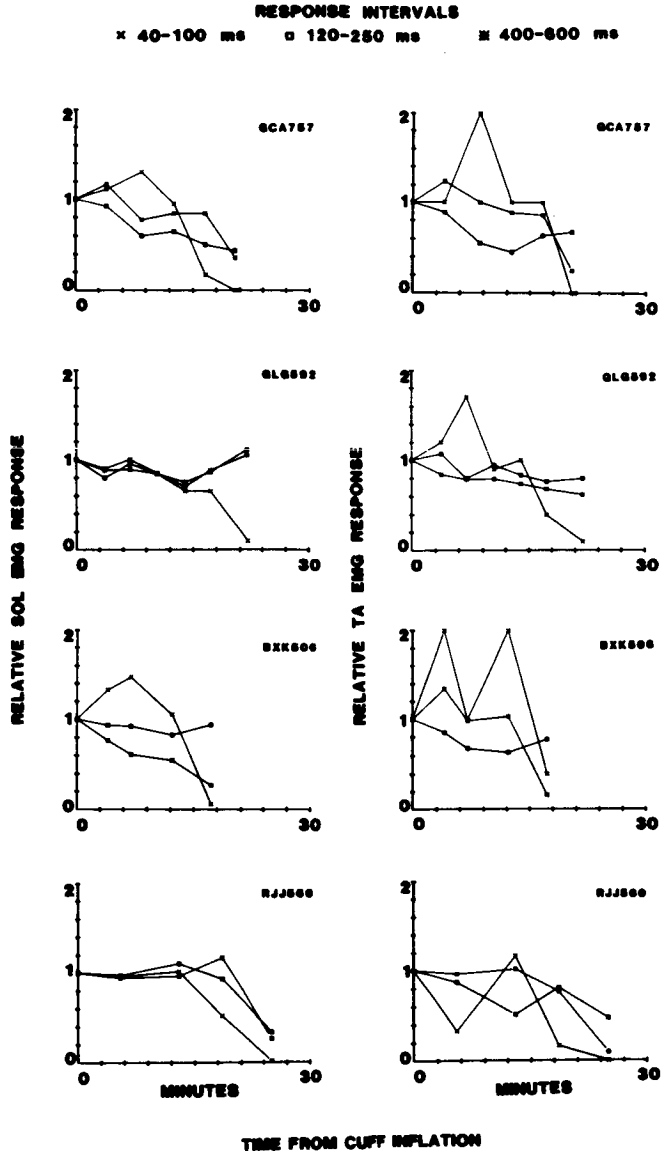


Fig. 2. EMG activity integrated over three selected intervals versus duration of ischemia. Interval 1 (X, 40-100 ms from perturbation) is the myotatic reflex. Interval 2 (□, 120-250 ms) is the postmyotatic response. Interval 3 (*, 400-600 ms) represents the stabilizing or tonic response to a sustained torque. Left side graphs are SOL responses to dorsiflexing torques. Right side graphs are TA responses to plantarflexing torques.

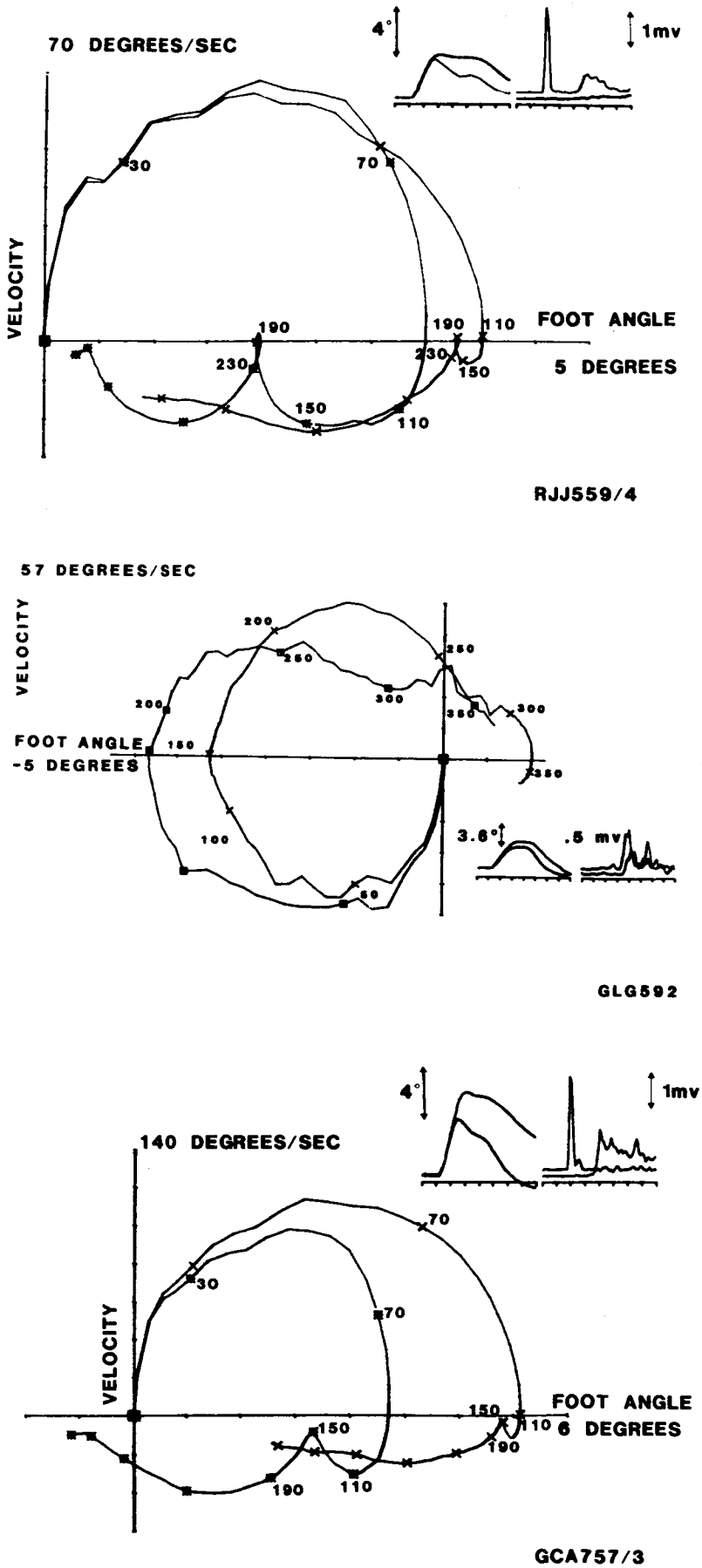


Fig. 3. Phase plane plots of foot trajectory following a torque perturbation (average of 10 responses) (* with reflex intact; X with ischemic, areflexic limb). Numbers along trajectory are time (in ms) from torque onset. Inset shows angle and EMG vs time for both trajectories. EMG traces have been vertically offset for clarity. Parts a and c show the SOL response to dorsiflexing torque. Part b shows the TA response to plantarflexing torque. Time markers for the insert records are 50 msec as in Fig. 1.

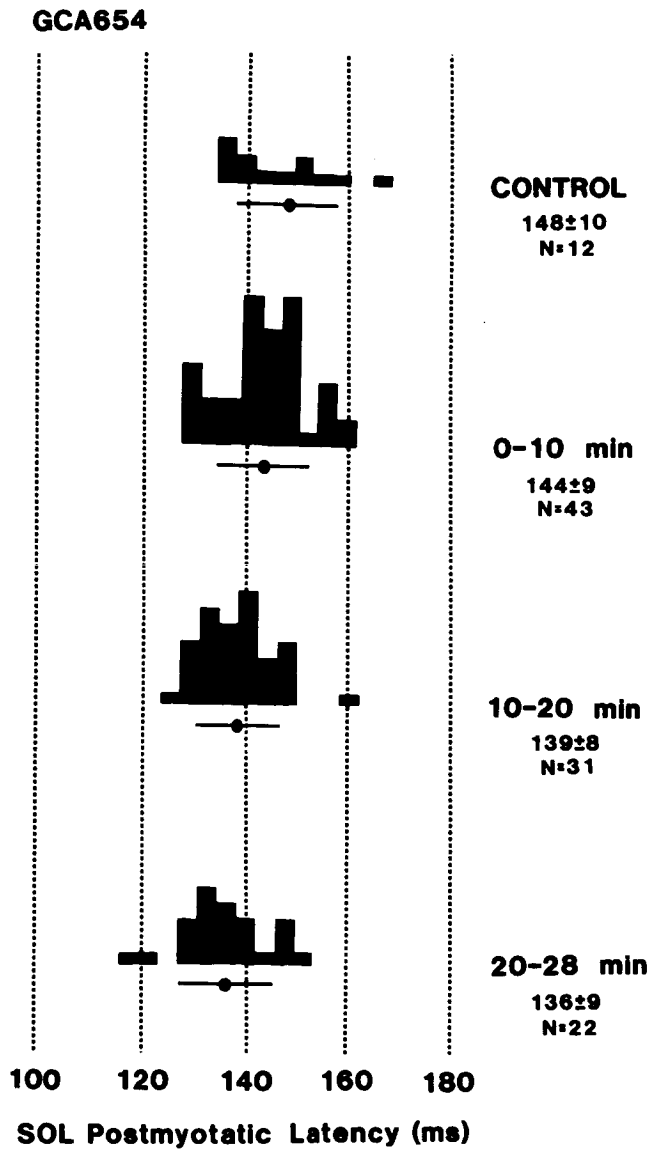


Fig. 4. Latency histograms for the SOL postmyotatic response. Each histogram is collected from successive responses in intervals i) preischemic, ii) first 10 minutes, iii) second 10 minutes, and iv) last 8 minutes (areflexic). Measurements made by visual inspection of individual EMG records. Bars under histograms show mean and one standard deviation.

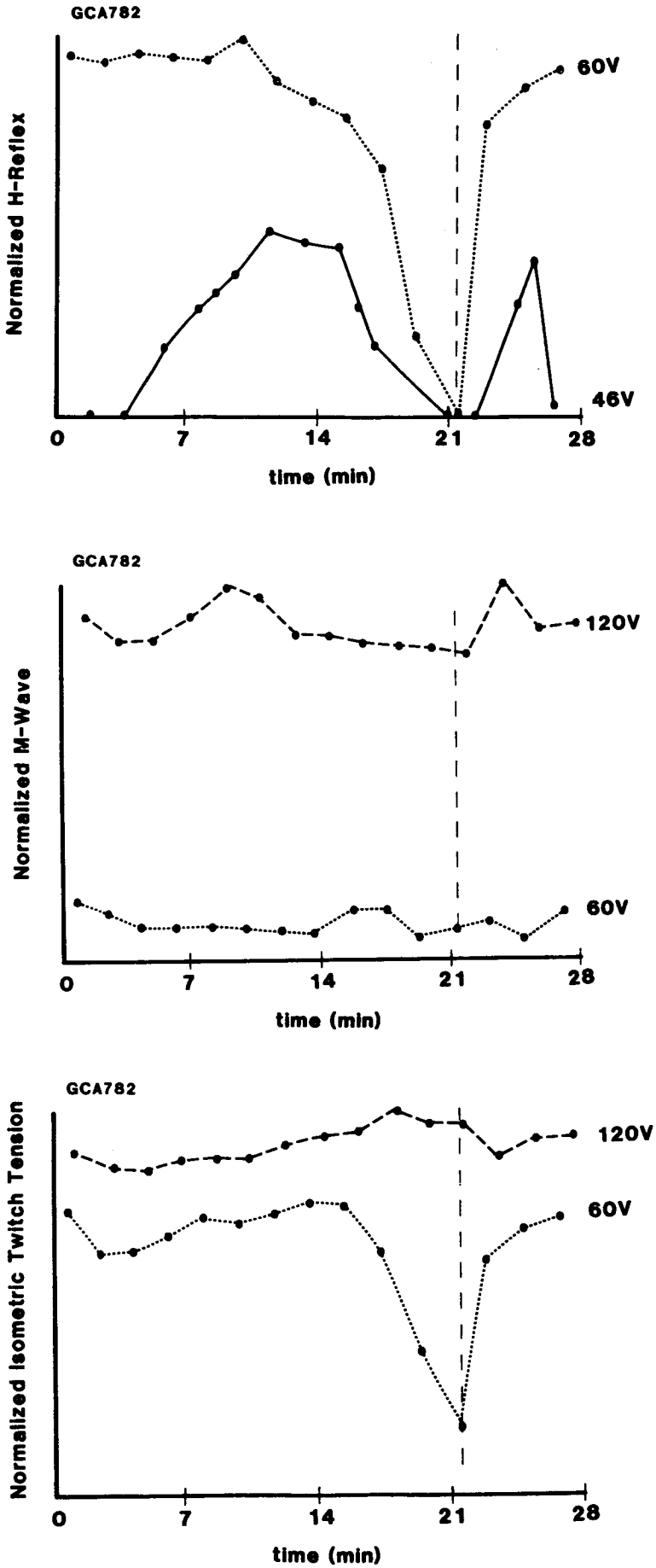
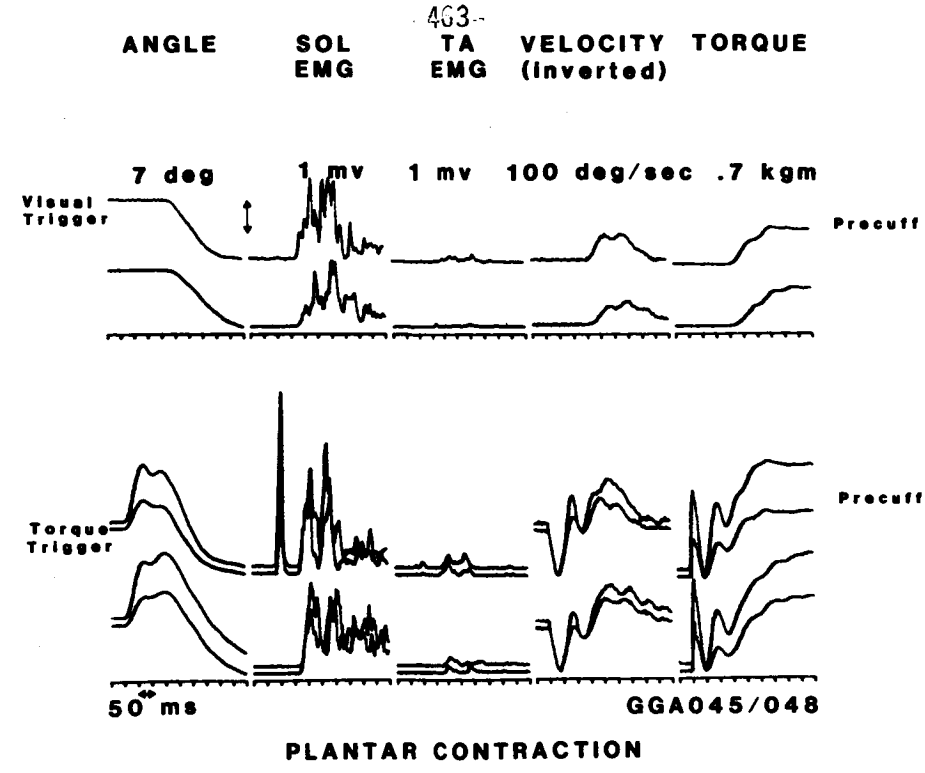


Fig. 5. Alterations in a) H-reflex, b) M-wave and C) muscle twitch during ischemia for three levels of electrical stimulation of the posterior tibial nerve. EMG measures are peak-to-peak values from unrectified surface recordings. Twitch measure is peak isometric tension. Vertical dashed line shows end of ischemia. 46v stimulus was initially, subthreshold for all responses. 60v stimulus evoked a maximal H-wave with a small M-wave. 120v stimulus evoked a maximal M-wave with no H-wave.



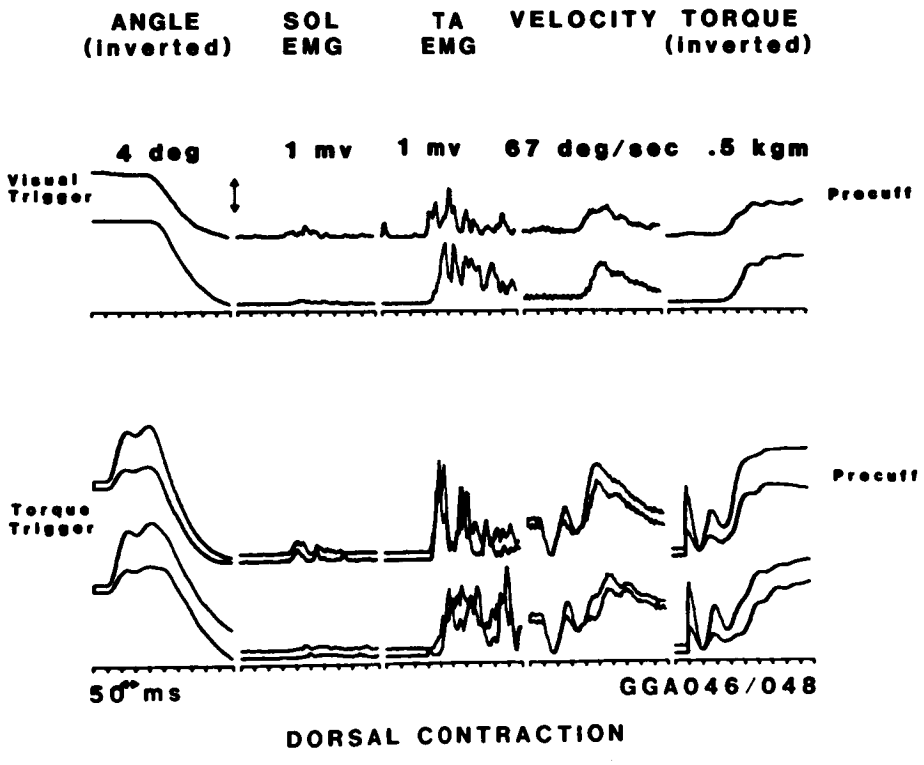


Fig. 6. Comparison of visually and stretch evoked responses in SOL and TA before and after ischemic deafferentiation: a) SOL responses and b) TA responses. The visual reactions are plotted at the top and the data has been shifted to the left so that the start of the EMG burst occurs at about the same time as the post-myotatic burst in the stretched muscle shown in the bottom part. The separate parts are joint angle, SOL EMG, TA EMG, joint velocity (inverted relative to the angle record) and joint torque. Each record is the average of 6 responses.

Neurological Control of Head Movement with Added Viscous Load

by

V.Lakshminarayanan, B.Hannaford, M.H.Nam and L.Stark

Depts. of Physiological Optics, Engineering Science and Neurology, University of California, Berkeley, CA.94720

ABSTRACT

The effect of added viscous load on time optimal horizontal head rotation was investigated. Quantitative measurements of the dynamic parameters were made. It was found that the adaptation to the added load was rapid and trajectories were produced which depended on the different detailed instructions given to the subjects, indicating preprogrammed compensatory mechanisms.

Simultaneously with head rotation, surface electromyograms were recorded. Ensemble averaging after preediting the EMG potentials enabled construction of the underlying neurological controller signals. A causal relationship between movement dynamics and the EMG signals was observed.

INTRODUCTION

Head movements play an important role in gaze (Zangemeister and Stark 1981, Dichgans, Bizzi, Morasso and Tagliasco 1974) and studies of the vestibular ocular reflex are widely used clinically and experimentally (Baloh & Honrubia 1979, Wilson & Melvill - Jones 1979, Henn, Cohen & Young 1980). The dynamic characteristics of head rotations are of interest in themselves and can be displayed for example, in main sequence deiagrams (Zangemeister, Jones and Stark 1981, Zangemeister, Lehman and Stark 1981a) to analyze basic neurological control patterns for

movement.

A study of neck muscle electromyograms (EMG), especially the envelope shape provides basic information about the controller signal driving the head movement (Zangemeister,Lehman & Stark 1981a,b, Stark,Lehman & Waite 1981). Review of the correlation between EMG signals and dynamical parameters of general limb movements are available in the literature (Desmedt 1973, Perry & Bekey 1981, Agarwal & Gottlieb 1982).

The load the head presents to the neck muscles includes the inertia due to the mass of bone and tissue, damping arising from the viscous nature of the tissues and muscles and the elastic stiffness. In different experiments researchers have altered the inertial-viscous-elastic load and have studied the direct effects on head movement and the adaptive changes in neurological control to compensate for the added load (Bizzi, Dev, Morasso and Polit 1978, Schirachi,Monk and Black 1978 , Gauthier,Martin and Stark 1981).

In this article, we present experimental results of studies of optimized head movements with different amounts of added viscous damping. By adding large, light weight vanes we used air itself as the viscous medium in which the head rotates. Our aim was to gain insight into neurological control of head movements adapting to changes in the viscosity parameter. We also show that even though interpreting surface EMG has traditionally presented many problems, by using the technique of ensemble averaging, it is possible to construct the underlying controller signals from horizontal head movements selected to have closely similar

dynamics.

METHODS

Horizontal head position was measured by a precision , low torque potentiometer attached with a universal joint between a bicycle helmet frame worn by the subject, and the laboratory frame. On top of the helmet, provision had been made for attaching (and removing easily) light weight rods to which cardboard vanes of varying sizes could be fixed as the source of additional viscous damping. In each experiment, 3 channels of data, ie:- head position and two channels of surface EMG from the left and right splenius muscles were recorded. The data was converted at 100 samples per second from analog to digital form with a Data Translation DT2762, 12 bit analog to digital converter installed in a LSI 11/23 microcomputer. Head velcity and acceleration were calculated digitally by the computer. Position data was calibrated by recording with the subject's head pointing successively in known directions. The viscous drag added by the cardboard vanes is a non-linear function of the velocity (White 1975) and calibration procedures were used to calculate the exact added viscous resistance and are reported in detail elsewhere (Nam,Lakshminarayanan,Lipiansky & Stark 1982).

EMG potentials were measured differentially with two S&W number 737 self adhesive Ag-AgCl disposable surface electrodes approximately 5 cm apart along the major axis of each muscle. A ground electrode of the same type was placed on the left forearm. The voltage difference between pairs of electodes was amplified by Tektronix AM502 differential amplifiers. The splenius muscle

was used because it's location is readily accessible and removed from other muscle which might cause cross talk. The electrodes were placed in a symmetrical pattern so as to get as nearly similar responses from the two muscles as possible. The use of the 1ØØ Hertz sampling rate (which is unusually low for EMG work) is justifiable since the object of the experiment is to observe the envelopes of the EMG signals and not to analyze the frequency content or short time responses of the signals themselves. The EMG signals were calibrated by connecting a 1 mv square wave from an oscilloscope's calibration output to the electrode terminals and the resulting digitized data was used to calculate the parameters of an affine transformation which was used to convert the recorded EMG data into microvolts. The EMG signals were rectified according to:

$$|V(t)| = \text{ABS}(V(t) - v)$$

where

$$v = 1/n \sum_{i=1}^{n} v(i)$$

Ensemble averaging is a valid procedure if each member of the ensemble consists of an underlying deterministic signal (the same for all members of the ensemble) to which is added a noise signal having zero mean at each sampling instant. To apply this the deterministic components of our data must be preselected so that averaging takes place among records of movements whose dynamics are as nearly identical as possible. This insures that any controller signal that is deterministically related to the

movement dynamics will be reliably represented by the ensemble average but does not presume the existence of such a signal. In figure 1a we show fifteen individual time records of position , velocity, acceleration and the two rectified EMG signals that were selected from an experiment (40 degree head movement) superimposed. These records were selected to have similar dynamics in position,velocity and acceleration and to fall within a range that is small compared to the observed phenomena when superimposed (Figure 1b).

Another important factor is time alignment i.e:- at what point is the movement aligned in time with the other members of the ensemble? Two criteria were used: (i) first change in head position and (ii) zero crossing of the acceleration trace. For criterion (i) , an operator visually judges the point in time in each movement where the position of the head just starts to change; the computer then aligns the records at the designated point and computes a new record whose value at each sampling instant is the arithmetic average of each of the selected movements at that particular time. In criterion (ii) the procedure is the same, but the records are aligned at the point where acceleration crosses zero in the middle of the movement. Comparing the two different ensemble averages of the same 16 movements , computed using the two criteria demonstrates excellent agreement between the two methods (figure 1c). Criterion 1 was used for subsequent analysis except where indicated.

A person can move his head in many different ways in the horizontal plane and hence a large number of controller signals

are possible if the movement is unconstrained. Using a repetitive step function generator, fixation lights at various angular seperations (20,40 and 60 degrees) were made to go on and off. After calibration and practice, the subject was instructed to move from one position to the next of the lights in as short a time as possible, that is in a time optimal manner. Ten movements were recorded as normal (N) head movement data. Then the cardboard vanes were added and similar data were taken under two different conditions, both with the vanes on (added B):

(i) The B+V condition : The subject was asked to make time optimal movements; secondary instruction was to maintain a high velocity inspite of positional inaccuracy and

(ii) The B+P condition : The subject's task was as before, to make time optimal movements, and also maintain maintain accurate final position (avoid overshoots and undershoots in the trajectory) in spite of decreased velocity.

A light source and pointer were attached to the head so that a shadow falling on the curved target screen gave the subject direct supplementary visual feedback of the angular position of his head.

We observed that when the external load was added, the subject went through a complex adapting process and the head movement dynamic parameters reached steady state conditions only after a transient adapting period. This was reached after 10-15 secs (4-5 movements). After adaptation, if the load were suddenly removed, a recovery or post-adaptation time of about 5-10 secs was required for the movement dynamics to return to

normal. Data was taken in the post-adaptation condition also.

RESULTS

Trajectories For normal head movements without supplementary position feedback, position traces were seen with at most one overshoot with smaller overshoots for smaller movements (fig 2a). The velocity traces showed peak velocities of 200,300 and 400 deg/sec for the 20,40 and 60 degree movements respectively. With supplementary feedback, the subjects were able to achieve a more accurate or steady state head position , but with higher overshoots (fig 2b). Smaller absolute but higher relative overshoots occured for smaller head rotations; in all cases steady state was very accurate. In contrast to the no feedback condition, positive and negative accelerations were about equal in most cases. Prominent differences in durations and dynamic parameters for the same amplitudes were noted between subjects. This variability was in contrast to lack of significant asymmetry between leftward and rightward movements.

Added Viscosity When the viscosity parameter wasincreased, a major dynamical change was found in the head movement characteristics (fig 2a). The position trace showed a decreased slopeand a longer duration to reach final headposition. The timeduration from peak position to steady state increased from 50-150 ms in the normal case to 150-200 ms with added viscous load. In the B+V paradigm (added viscosity, velocity dominant instruction) overshoots of up to 20% were seen with slightly decreased peak velocity and positive and negative peak accelerations. In the B+P paradigm (added viscosity, position

dominant instruction) overshoots were rarely observed. There was a decrease in peak velocity as well as in positive and negative accelerations (fig 2b).

We found that with normal loads, accuracy of movement was acheived in about 5-8 secs (2-3 movements) implying that the neck muscles generated the correct forces to drive the head to an accurate steady state position within that time. In the post adaptation condition the normal movements showed more fluctuations in positional accuracy than in the pre-adapted state. With added viscosity also adaptation to the new load took place rapidly. However, more variations within a particular subject's performance were seen in consecutive trials and the subject took a longer time to achieve accurate movements. In general we can say that in both the B+P and B+V conditions, the initial adapting and later adapted states were similar and both showed the effect of subject's intent on parameters of the trajectory.

Main Sequence Main sequence plots were obtained fora single trained subject performing normal head movements ranging from 2 to 120 degrees (fig. 4 left panel). The time duration of the movement ranged from 120 ms (5-15 degrees) to 280 ms (100-120 degrees). For the same range, the peak velocities ranged from 62 deg/sec to 1000 deg/sec. The accelerations varied from 900 deg/sec^2 to 10000 deg/sec^2 for peak positive accelerations and from 950-12000 deg/sec^2 for peak negative accelerations.

When the additional viscous load was added it is found that the main sequence plots shifted downwards with the exception of

the duration-magnitude log plot. In all these cases it should be noted that the unloaded movement parameters had still higher values and shorter durations. Durations for the B+V experiments were smaller (faster movements) than those for the B+P movements (fig. 4 left panel). Higher velocities as well as higher positive and negative (in magnitude) accelerations were found for the B+V movements as compared with the B+P movements. The B+V data points were right shifted as expected since the subject was instructed to achieve fast movements rather than precise movements, and larger overshoots and steady state values resulted. Experiments were done with four different vane areas and the resulting main sequence is shown in figure 4 right panel. The head movement trajectories became sluggish when larger vane areas (greater drag forces) were added. The largest vane area was used in other illustrated experiments. The experimental results shown were obtained from a different subject who performed slower movements that the subject who provided fig.4 right panel data.

Analyzing EMG data, we found that the EMG envelopes had three distinct pulses designated by P1,P2 and P3 to represent the first agonist, the antagonist and the second agonist pulse respectively (Nam,Hannaford,Lakshminarayanan and Stark 1983, Hannaford,Nam,Lakshminarayanan etal, these proceedings). Other pulses which may be present are hard to distinguish with our resolution and will be considered to be of secondary importance in the following.

In general, it was found in the case of normal head movements, the width of P1 appears to be positively correlated with the

magnitude of the movement. A time and magnitude locking between the negative acceleration pulse and P3 was also observed. Addition of the viscous load produced changes in the movement dynamics and EMG signals. For a 20 degree movement, the addition of the viscous load resulted in almost total elimination of P3 when compared to 60 degrees. For 40 degrees, P3 is somewhat reduced and is similar in shape to that of the 60 degrees average. P2 appears to be about the same in the 20 and 60 degree cases, and somewhat increased with viscosity in the 40 degree case (fig. 4) (Lestienne 1979). We consistently observed in the 20 and 40 degree data, an earlier onset of P1 when the two traces are synchronized at the first change of position. Since the EMG onset is causally first in the chain of events, synchronizing the records at the onset of P1 would reveal an extra 20 ms dynamical lag between force onset and movement (Zangemeister,Lehman and Stark 1981a,b) with viscous load.

When the loaded movement records are superimposed and aligned at first change of position criterion (i), no consistent pattern of EMG envelope variation with movement amplitude reveals itself. But when aligned using criterion (ii), acceleration zero crossing, the EMG traces fall into place (Fig 4 right panel). The presence of a third agonist pulse suggests a higher order control strategy.

DISCUSSION

By choosing the vane area to provide about 10-20 times the viscosity of the normal head (as estimated from modelling studies)(Zangemeister,Lehman and stark 1981b), we were able to

produce a marked slowing of head velocity and acceleration and a long duration of movement to reach final head position. The cooperative subjects rapidly adapted to the added load. Since all movements were intended time optimal, acceleratory forces were already maximal (for each amplitude movement) before the load was added and increase in muscle force cannnot account for the adaptation to the load.

In the B+V paradigm, with its larger overshoots and steady state values, the higher accelerations were possibly due either to pulse width control (muscles were activated longer) or to pulse height control (additional recruitment and higher firing frequencies). This allowed the subject to keep the reduction of acceleration due to the added load and velocity to a minimum. Thus time optimal movements are only time optimal within the bounds of an additional scaling constraint. In our experiments we could obtain only some indirect evidence regarding strategies the subjects used to alter the pulse height and pulse width of their controller signals. Figure 3 left panel shows evidence of this alteration of control strategy by long trained subjects who can change their main sequence (symbols marked pw+).

In the B+P paradigm, where amplitude accuracy is made an important criterion, even more decrease in velocity and acceleration is seen. However, the longer duration of movement permits a trajectory where additional forces are allowed to become activated and contribute to the adaptation. Recall that related changes in acceleratory forces were observed (equalization of positive and negative acceleration with added

load).

With the added viscous load, changes in EMG reveal a corresponding change in control strategy. There is a complete suppression or reduction in height of the P3 agonsit pulse. It has been suggested (Lestienne 1979,Bizzi,Dev,Morasso & Polit 1978) that the third EMG pulse in conjunction with the muscle's non-linear viscosity and the dynamics of the end of the movement provides extra viscosity to dampen the final stages of movement. From our study it is easy to conclude that the extra viscous damping provided by the P3 is not needed when extra viscous load is added. Gauthier,Martin and Stark (1981) studied the effects of added inertial loads on head rotations; they found much of the adaptation that took place was related to active co-contraction and to alternating contraction of agonist and antagonist to provide for active damping of the destabilising added inertial load. By increased damping the added-B loads we used provided greater stability.

We also note a time difference of approximately 50ms between agonist onsets in the loaded and unloaded cases. If the traces were realigned according to the well defined agonsit pulse onsets, an increased delay would be seen in the viscous loaded case between the onset of EMG activity and the first signs of head movement (Bizzi,Kalil,Morasso and Tagliasco 1972, Zangemeister,Lehman and Stark 1981a,b).

The above observations raise interesting questions beyond the scope of this article and calls for extensions of our research involving both electromyography and modelling studies.

Acknowledgements We are pleased to acknowledge partial support from the NCC2-86 Cooperative agreement, NASA-Ames Research Center, from the National Institutes of health training grant in Systems and Integrative Biology and a grant from the Ministry of Education, Government of Korea (MN). M.H.Nam is at present with the Dept. of Electrical Engineering, Kon-Kuk University, Seoul 133,Korea.

REFERENCES

1. G.C.Agarwal and G.L.Gottlieb (1982): "Mathematical modelling and simulation of the postural control loop: Part I", CRC Crit.Rev.Biomed.Eng.,vol.8,pp.93-134.

2. R.Baloh and V.Honrubia (1979): "Clinical neurophysiology of the vestibular system" Philadelphia,PA; F.A.davis Co.

3. E.Bizzi,P.Dev,P.Morasso and L.Polit (1978): "Effecxt of load disturbances during centrally initiated movements", J.Neurophysiol.,vol.41,pp.542-556.

4. E.Bizzi, R.E.Kalil,P.Morasso and V.Tagliasco (1972): "Central programming and peripheral feedback during eye-head coordination in monkeys", Bibl.Ophthal., vol.82,pp.220-232.

5. J.E.Desmedt,Ed.,(1973): "New developments in electromyography and clinical and clinical neurophysiology", Basel, Switzerland: Karger. .

6. J.Dichgans,E.Bizzi,P.Morasso and V.Tagliasco (1974): "The role of vestibular and neck afferents during eye-head coordination in the monkey",Brain Res.,vol.71,pp.225-232.

7. G.M.Gauthier,B.Martin and L.Stark (1981): "Effects of

inertial load on head-eye movements",OMS'81, Calif. Inst. Tech., Pasadena, CA., January 1981.

8. V.Henn,B.Cohen and L.R.Young ,Eds., (1980): "Visual-vestibular interaction in motion perception and the generation of nystagmu", Neurosciences Research Program Bulletin, Vol.18,No.4, Boston,MA: MIT Press.

9. F.Lestienne (1979): "Effects of inertial load and velocity on the braking process of voluntary limb movements", Exp.Brain Res.,vol.35, pp.407-418.

10. M.H.Nam,V.Lakshminarayanan,E.Lipiansky and L.Stark (1982): "Effect of external viscous load on head movement", (submitted to IEEE Trans.Biomed.Engineering).

11. M.H.Nam,B.Hannaford,V.Lakshminarayanan and L.Stark (1983): "EMG as controller signal with viscous load", (Submitted to J.Motor Behavior).

12. J.Perry and G.A.Bekey (1981): "EMG-Force relationships in skeletal muscle",CRC Crit.rev.Biomed.Eng.,Vol.8,pp.1-22.

13. D.Schirachi, D.Monk and J.Black (1978): "head rotational spectral characteristics during two dimensional smooth pursuit tasks", IEEE Trans.Sys.man & Cybern.,vol.SMC-8,pp.715-722.

14. L.Stark,S.Lehman and T.Waite (1981): "Neural signal envelope shapes and optimization of motor performance", 11th annual meeting, Los Angeles,CA.,CA.,Abstracts,vol.7,pp299,Bethesda,MD: Soc. for Neuroscience.

15. F.White (1975): "Fluid Mechanics",New York,NY: McGraw Hill.

16. V.J.Wilson and G.Melvill-Jones (1979): "Mammalian vestibular Physiology", New York,NY:Plenum Press.

17. W.H.zangemeister,A.Jones and L.Stark (1981): "Dynamics of

head movement: main sequence relationship", Exp. Neurol., vol.71,pp.76-91.

18. W.H.Zangemeister,S.Lehman and L.Stark (1981 a): "Simulation of head movement trajectories : model and fit to main sequence", Biol.Cybern.,vol.41,pp.19-32.

19. W.H.zangemeister,S.Lehman and L.Stark (1981 b): "Sensitivity analysis and optimization for a head movement model",Biol.Cybern.,vol.41,pp.33-45.

2Ø. W.H.zangemeister and L.Stark (1981) : "Active head rotations and eye-head coordination", Ann.NY.Acad.Sci.,vol.374,pp.54Ø-559.

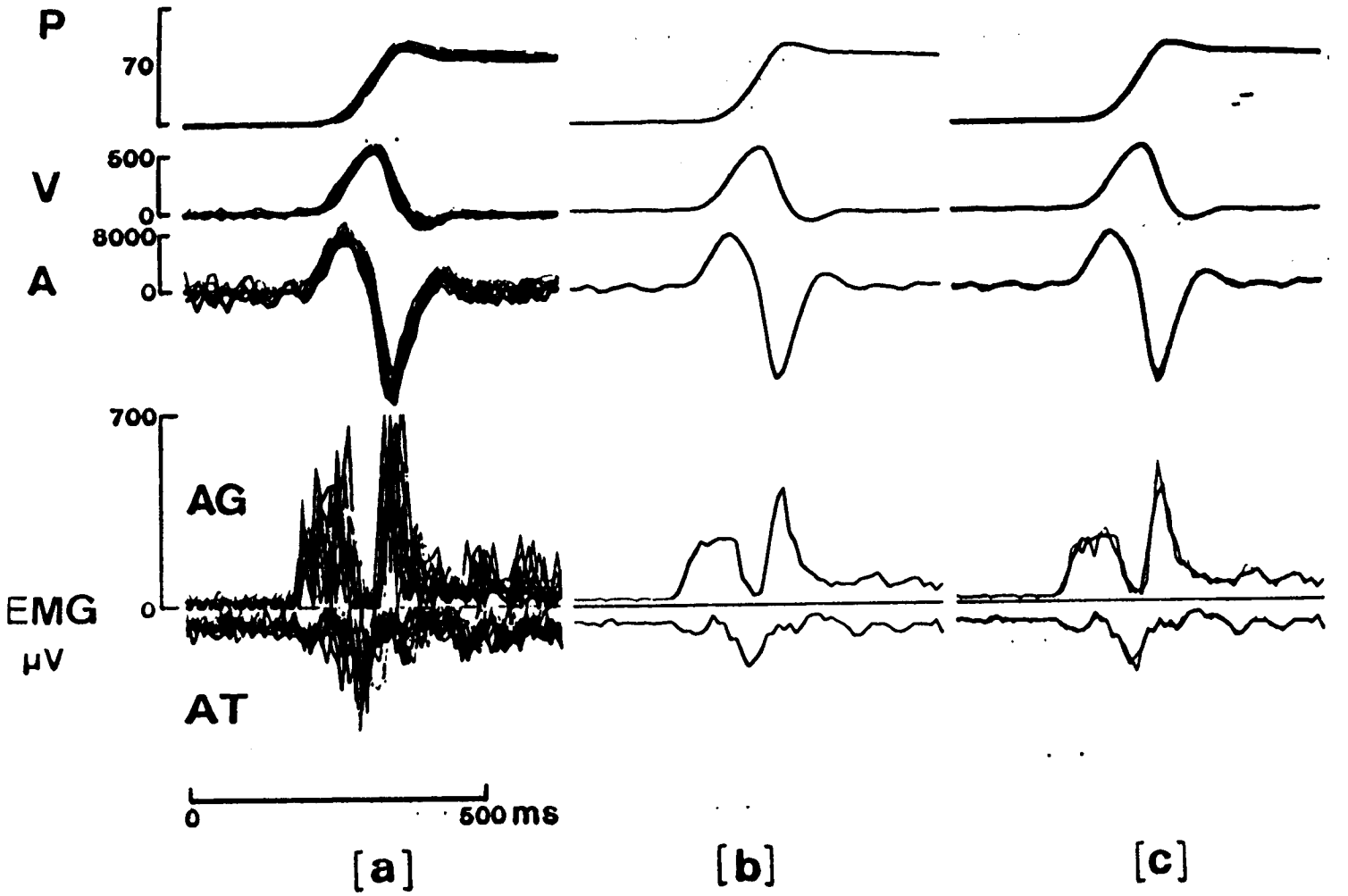
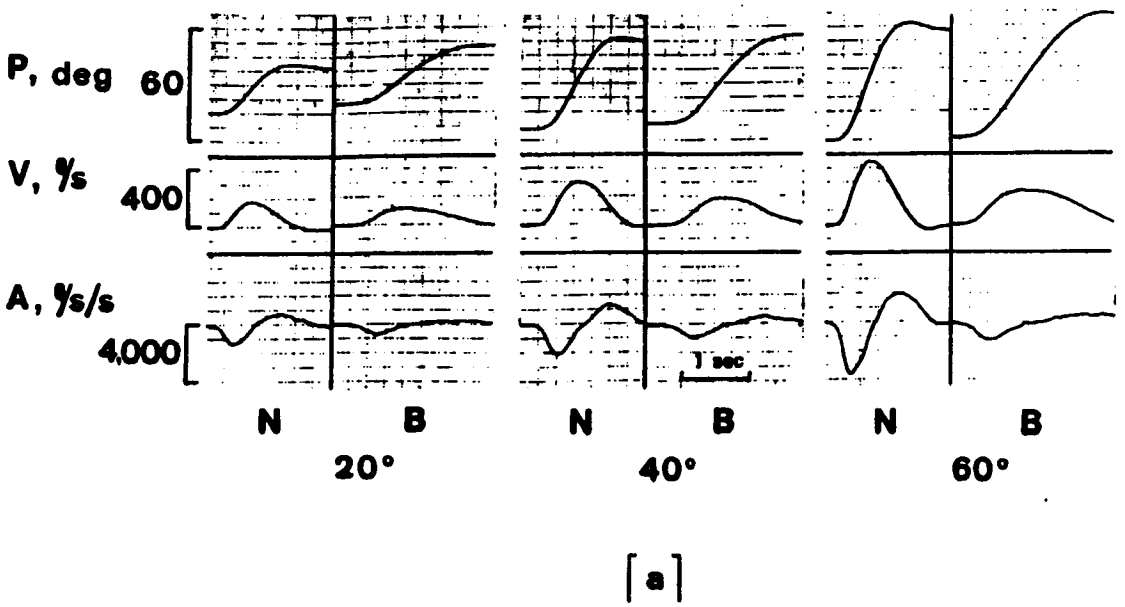
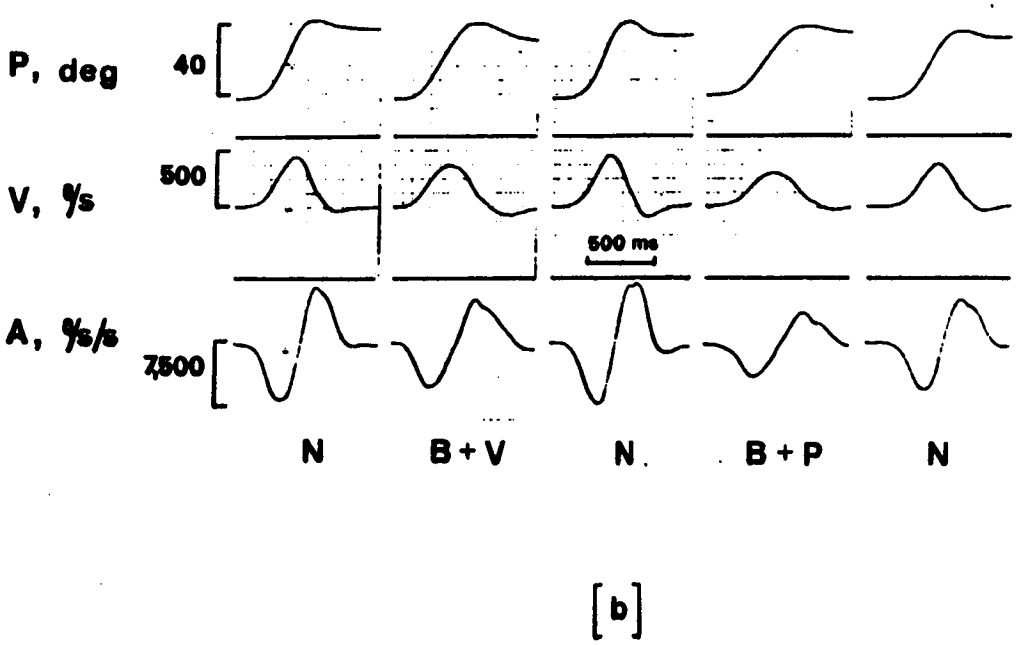


Figure 1. (a). Superposition of 16,40 degree movements. The records are aligned in time using criterion (i): first change of position (See Text)
(b). Ensemble average of movements shown in (a).
(c). Overplot of ensemble averages using criterion (i) (thick line) and criterion (ii), the zero crossing of acceleration trace (thin line), for time alignment; note good agreement.

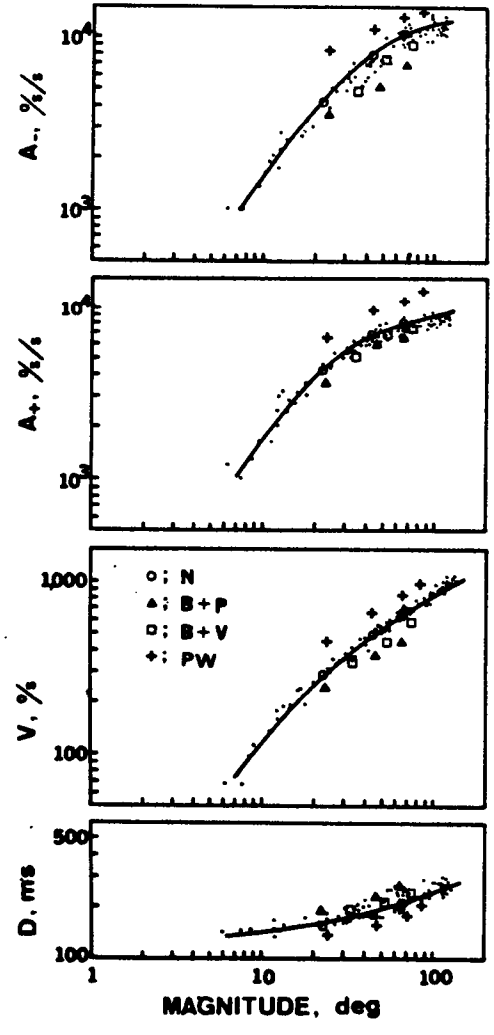


[a]

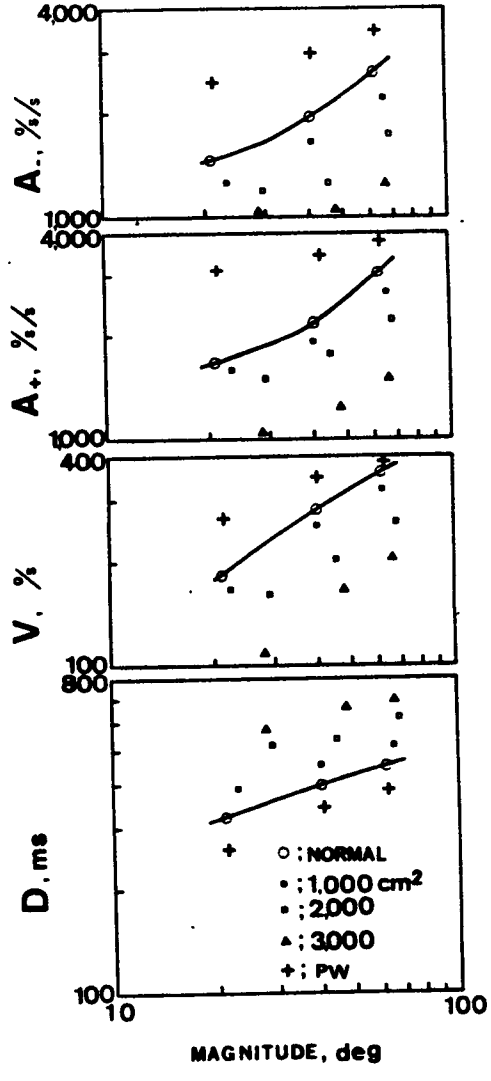


[b]

Figure 2. Head movement trajectories (a). Under normal conditions (N) and with added viscosity (B). Position P (degrees), Velocity V (°/sec), Acceleration A (°/sec/sec) and Time (seconds) Note decreased velocity and acceleration, especially negative acceleration with added B. (b). Importance of subject's intent: B+V and B+P conditions (see text). Note overshoot in B+V condition and reduced velocity in B+P condition.



(LEFT)



(RIGHT)

Figure 3. Main sequence for head movements

(left panel): Effect of added viscosity under two different sets of instructions (B+P and B+V; see text) Normal main sequence and variation of a single trained subject attempting time optimal head rotation is also shown. Plus sign (+) indicates pulse width control strategy

(right panel): Effect of added viscosity with different vane areas providing three levels of viscous drag All data points are averages of five consecutive adapted state movements

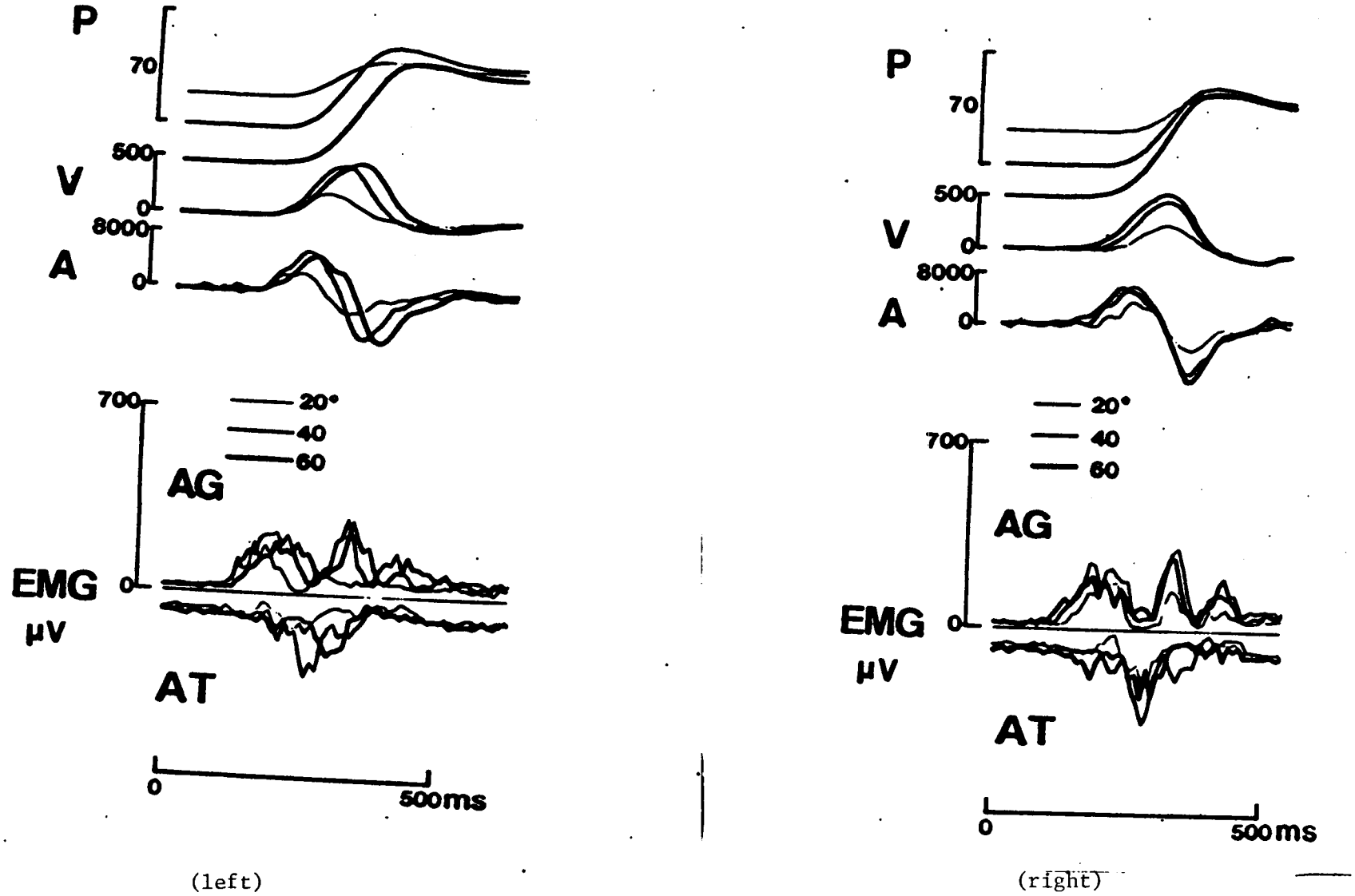


Figure 4. EMG with added viscous load
(left panel): Overplot of loaded movements aligned using criterion (i), first change of position
(right panel): Overplot of same movements aligned using criterion (ii), acceleration zero crossing
Note roughly congruent EMG envelopes indicating a controller signal that is stereotyped and simply scaled to program a movement's amplitude

Title: Effects of Loads on Time Optimal Head Movements: EMG, Oblique, & Main Sequence Relationships

Authors: Blake Hannaford, Robert Maduell, Moon-Hyon Nam, Vasudevan Lakshminarayanan, and Lawrence Stark

ABSTRACT

Head movements can be parameterized in terms of dynamical variables which are functions of movement magnitude when the movements are made as fast as possible by trained, cooperating human subjects. Electromyographic activity of neck muscles can be measured to characterize the nervous controller signals regulating such movements. Studying oblique head movements reveals the extent to which the horizontal and vertical systems are coupled or independent. In these investigations, all of these methods are used to study the effects on head movements of added inertial, viscous, and elastic loads.

INTRODUCTION

Head movements result from a variety of stimuli, both voluntary and involuntary: proprioceptive sensors in the neck and vestibular apparatus provide spatial orientation feedback; the ears provide audible feedback; and the eyes provide visual feedback through their probing of the visual world. Gaze directed head movements are much slower and smoother than the accompanying eye movements due to the large viscosity and inertia of the head-neck system (Viviani & Berthoz, 1975; Zangemeister, & Stark, 1981a,b). The head positioning system plays an important role in orientation with the external world.

In studying head movements, we use two techniques to elucidate the central and reflexive control strategies underlying the movements and their dynamics (time course of position, velocity, and acceleration). These are, measurement of EMGs from an agonist/antagonist pair of muscles contributing to the movements and adding various mechanical loads to the head, allowing the subject to adapt his movements to the load, and recording the resulting movements and EMGs.

Detailed analysis of the timing of the EMG have shown role of these multi-pulse control signals and revealed important features of the control of movement: multiple burst envelopes

(Wadman,VanderGon,Geuze & Mol 1979, Bizzi,Kalil & Tagliasco 1971, Bizzi,Kalil, Morasso & Tagliasco 1972, Wadman, VanderGon & Derksen 1980, Ghez & Martin 1982, Terzuolo, Soechting & Viviani 1973, Lestienne & Bouisset 1971,1974); occurrence and non-occurence of reciprocal innervation of agonist-antagonist pairs (Bouisset, Lestienne, Maton 1977, Lebozec, Maton, Cnockaert 1980, Wadman, Vandergon, Derksen 1980); qualitative increase of EMG with size and speed of movement (Lestienne & Bouisset 1972,74, Wadman,VanderGon,Ghez & Mol 1979, Wadman,VanderGon,Derksen 1980); and change in the braking process with added inertial loads (Lestienne 1979, Gauthier,Martin & Stark 1981).

Quantitative correlation of EMG with movement requires measurement of position, velocity, acceleration, and forces or torques simultaneously with EMG (Bouisset & Goubel 1968, Danoff 1979, Bouisset, Lestienne & Maton 1977, Ghez & Martin 1982, Maton, Lebozec, Cnockaert 1980, Zangemeister, Stark, Meienberg, & Waite 1982). Often these quantitative studies have accompanying models that explain and make consistent the observed relationship between these variables (Zangemeister, Lehman & Stark 1980a,b , Wadman, VanderGon & Derksen 1980).

Both vertical and horizontal head rotations have been studied and modeled (Viviani & Berthoz 1975, Zangemeister & Stark, 1981c,d) to characterize the parameters of these movements and to provide some insight into their control systems. A useful approach is to consider time-optimal trajectories; approximately time-optimal trajectories are obtained from a subject voluntarily rotating the head from one position to another as fast as possible. Then it becomes possible to parameterize the dynamics of head trajectories in terms of peak velocity, peak acceleration, peak deceleration, and movement duration as functions of movement magnitude and to display these in the Main Sequence diagram for study (Stark, Zangemeister et al. 1980, Zangemeister & Stark, 1981a,b, Zangemeister, Lehman & Stark, 1981a). From these diagrams, general trends can be studied and identified.

To understand the neurological control systems behind these types of movements, we have measured intended time optimal movements, their dynamics, and electromyographic activity from neck muscles in normal movements, and also in movements made with external loads (viscosity, inertia, and elasticity) applied to the head. In studying oblique head movements, we have observed the effects of load applied in one dimension on the dynamics of the resulting movements in both the same and the orthogonal direction.

METHODS

Head Movements

Horizontal head position was measured by a precision potentiometer attached with a universal joint between a bicycle helmet frame worn by the subject, and the laboratory frame. For oblique experiments, vertical rotation of the helmet frame was measured by a second potentiometer. A viscous load could be added to the subject's head by attaching large, light weight vanes to the helmet frame. The drag viscosity added by these vanes is a non-linear function of the velocity (Nam, Lakshminarayanan, Lipiansky & Stark 1982) and calibration procedures were used to calculate the exact added viscous resistance. Inertial load was added by attaching lead weights to the helmet.

In both the oblique and horizontal experiments, elastic force was applied tangentially in the horizontal plane by a the addition of a toriodal wooden pulley rigidly mounted on top of the bicycle helmet. The added force was provided by three parallel springs with spring constant, K = 10.6 gr-force/deg.

Three channels of data were recorded from the horizontal head movement experiments. These were head position, and two channels of surface EMG from the left and right splenius muscles. In the oblique experiments, EMGs were not recorded. The data was converted at 100 samples per second from analog to digital form with a Data Translation DT2762 12 bit analog to digital converter installed in a LSI 11/23 micro-computer. Position data was calibrated by recording with the subject's head pointing successively in known directions. Head velocity and acceleration were calculated digitally by the computer after data recording was complete. Zero phase shift between position and its higher derivatives was assured by the use of a non-causal differentiation algorithm which estimated the slope by:

$$\dot{x}(n) = \left(\sum_{i=1}^{3} x(n+i) - \sum_{i=1}^{3} x(n-i) \right) / 6$$

Two methods were used to give the subject supplementary visual (non-proprioceptive) feedback of actual head position during and immediately after movement. In horizontal experiments, a small light bulb projected the shadow of a pointer attached to the head onto the target screen. In the oblique experiments, the subject closed one eye and sighted the targets with the other eye through cross hairs projecting from the helmet in front of him.

EMG Recording

EMG potentials were recorded by computer at 100 samples per second with techniques described elsewhere (Lakshminarayanan, et. al., these proceedings).

Sampling Rate

One hundred hertz is an unusually low sampling rate for EMG work. We believe that it is justifiable to use a rate as low as this when the object of the experiment is to observe envelopes of the EMG signals, and not to analyze the frequency content or short time responses of the signals themselves. If the object of the sampling system is to estimate only the envelope shape as a function of time, and an ensemble of records are taken, then the signal need only be sampled at a rate sufficient to characterize the signal envelope, regardless of whether the low pass filtering is done before sampling (as when sampling the rectified, integrated EMG), or after sampling (as when the rectified members of the ensemble are averaged). Excellent, state of the art papers have used similar bandwidth in their analyses i.e. Lestienne rectified and low-pass filtered at 50 hz (Lestienne 1979).

Calibration

To calibrate the EMG signals, a 1 mv square wave from an oscilloscope's calibration output was connected to the electrode terminals after they were removed from the subject and the resulting digitized data recorded on disk. This data was used to calculate the parameters of an affine transformation with which the recorded EMG data was converted into microvolts. The EMG signals were then rectified around their average value.

Ensemble Averaging

The idea of envelope estimation by ensemble averaging depends on the assumption that each member of the ensemble consists of an underlying deterministic signal (the same for all members of the ensemble) which is multiplied by a noise signal having zero mean and mean absolute value of 1 at each sampling instant. In order for this to apply to our data, the deterministic components of our data (position, velocity, and acceleration) must be preselected so that averaging only takes place among records of movements whose dynamics are as nearly identical as possible. In computing such an ensemble average in practice, the records are selected so that throughout the movement, position, velocity, and acceleration fall within a range that is small compared to the observed phenomena. This insures that any controller signal that is deterministically related to the movement dynamics will be reliably represented by the ensemble average but does not presume the existence of such a signal because selection is based only on the movement dynamics and not on EMGs.

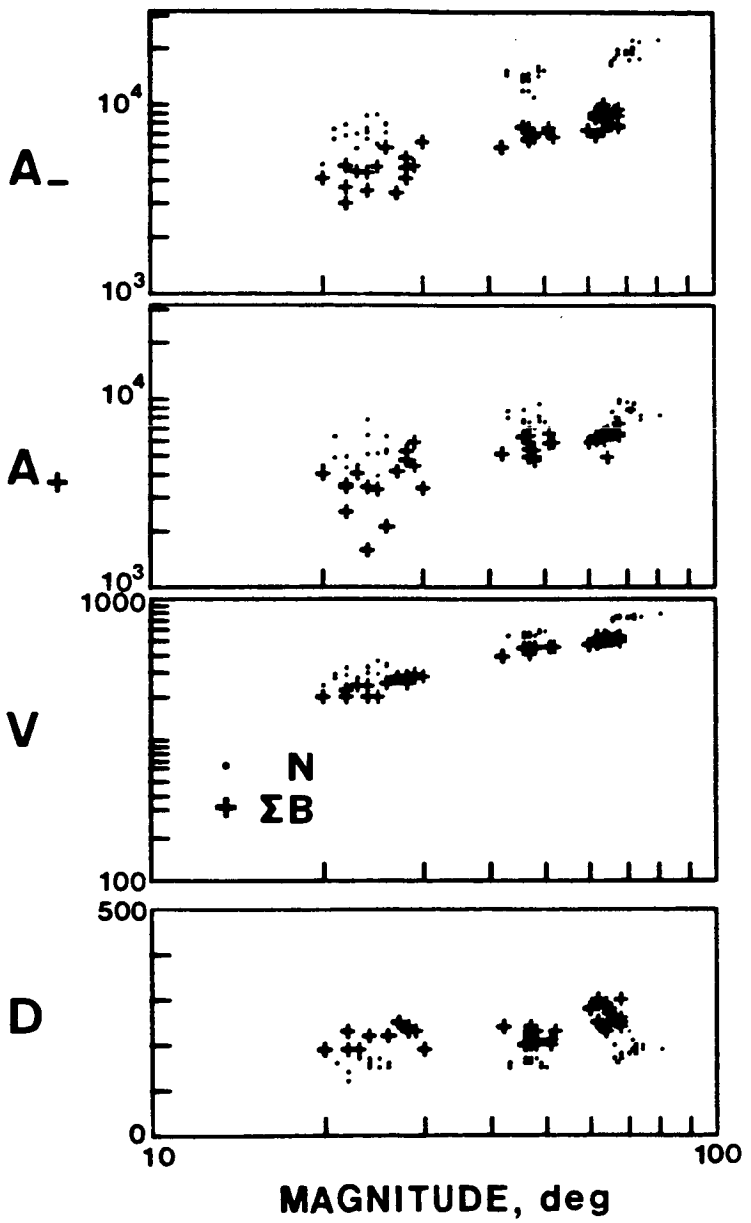


Fig. 1

Main Sequence Diagram of horizontal head rotations under conditions of normal and viscous loading. Four ordinates are movement duration (in milliseconds), velocity (deg/sec), peak starting acceleration (deg/sec/sec), and peak stopping acceleration (deg/sec/sec). Abscissa is magnitude of head rotation in degrees. Dots(.) are normal movements and plus signs(+) denote movements with added viscous load.

RESULTS

Main Sequence

A main sequence diagram of the observed head movements of one subject (M.N.) was plotted to show the magnitude, duration, velocity, and acceleration of the observed time optimal movements (dots (.)) and movements under conditions of viscous loading (plus signs (+)) (Fig 1). As might well be expected, the data show that movements of all magnitudes measured are reduced in velocity and acceleration and take longer to complete when performed with a viscous load. The data are clustered into three groups in position because they were made in response to target displacements of 20, 40, and 60, degrees. Thus, when a subject attempted to make time-optimal movements in response to a constant target displacement, the resulting movements exhibited a variability.

Oblique Main Sequence

Main sequence diagrams for selected oblique movement paradigms are plotted in Figure 2. When pure horizontal movements (elastic load: relaxing) (Fig. 2c) are closely compared with the horizontal component of oblique movements (same loading conditions) (Fig. 2d), a slight reduction in peak velocities and accelerations can be seen in the horizontal components of oblique movements with respect to purely horizontal movements of equivalent magnitude. In the relaxing direction, a similar reduction is discernible (Fig.s 2a,2b).

Oblique Trajectories

Normal (unloaded) oblique movements exhibit a wide range of trajectories (Fig. 3a). Overshoots (primarily vertical) and undershoots (horizontal) are prominent in this record of 67 degree movements from one subject (the angle refers to inclination as opposed to magnitude). When an elastic load is added in the horizontal direction, records of movements made in the direction of relaxing the spring (Fig. 3b) show reduced variability. This can be interpreted as an adaptation to the horizontal load by both the vertical and horizontal components.

Normal, Horizontal Movements

Ensemble averages of time optimal movements made without added loading were computed for target displacements of 20, 40, and 60 degrees (Fig. 4, thin lines). The EMG envelopes had three distinct pulses. We have used the terms P1, P2, and P3 to designate the first agonist pulse, the antagonist pulse, and the second agonist pulse respectively. Other pulses may also be

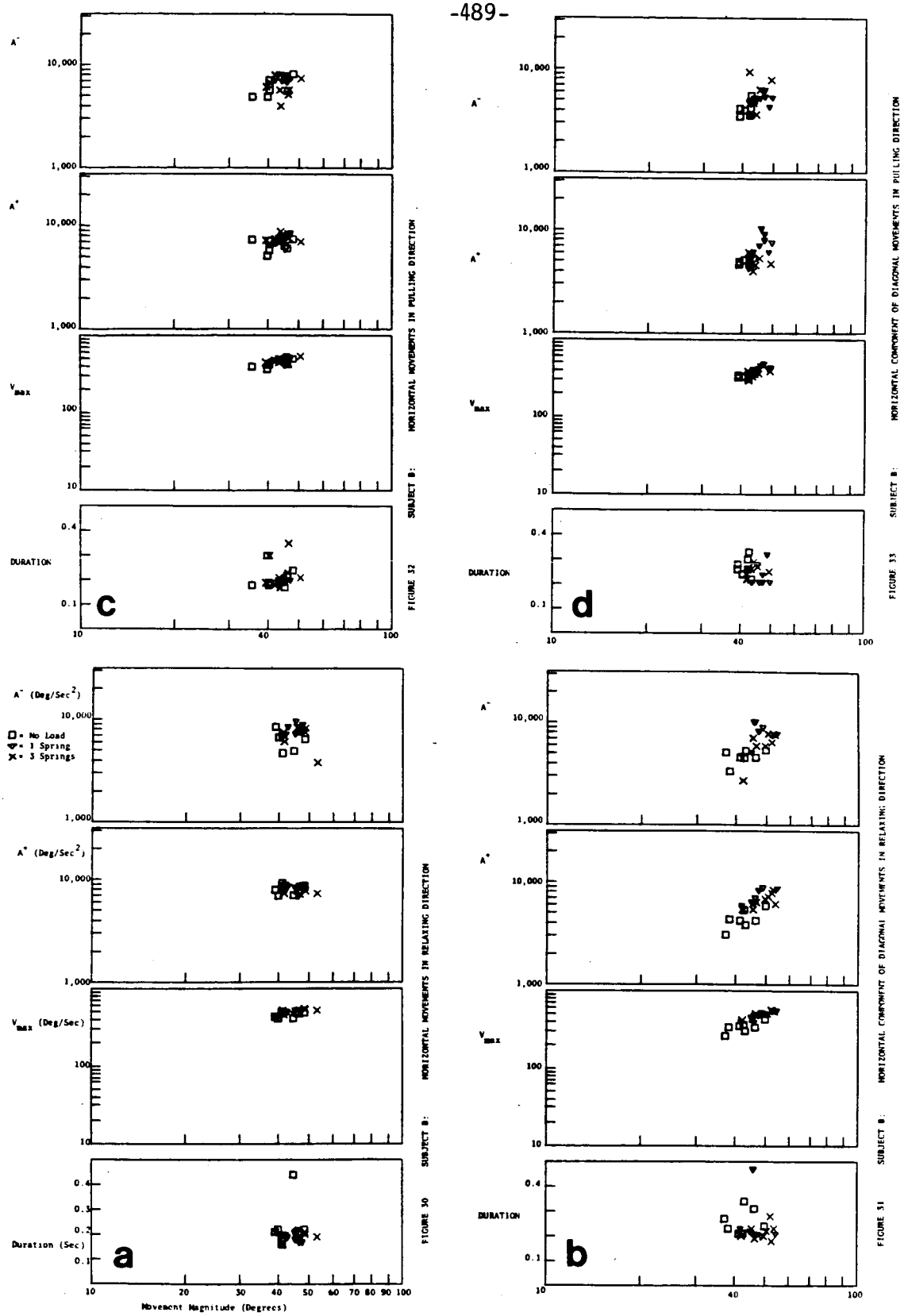


Fig. 2

Oblique Main Sequences. Main Sequence diagrams for four types of movements are displayed: a) Pure horizontal movements in the direction of relaxing the added load (if present); b) Horizontal component of oblique movements (also in relaxing direction); c)Pure Horizontal movements in the direction of stretching the added elastic load; d) Horizontal component of oblique movements in stretching direction.

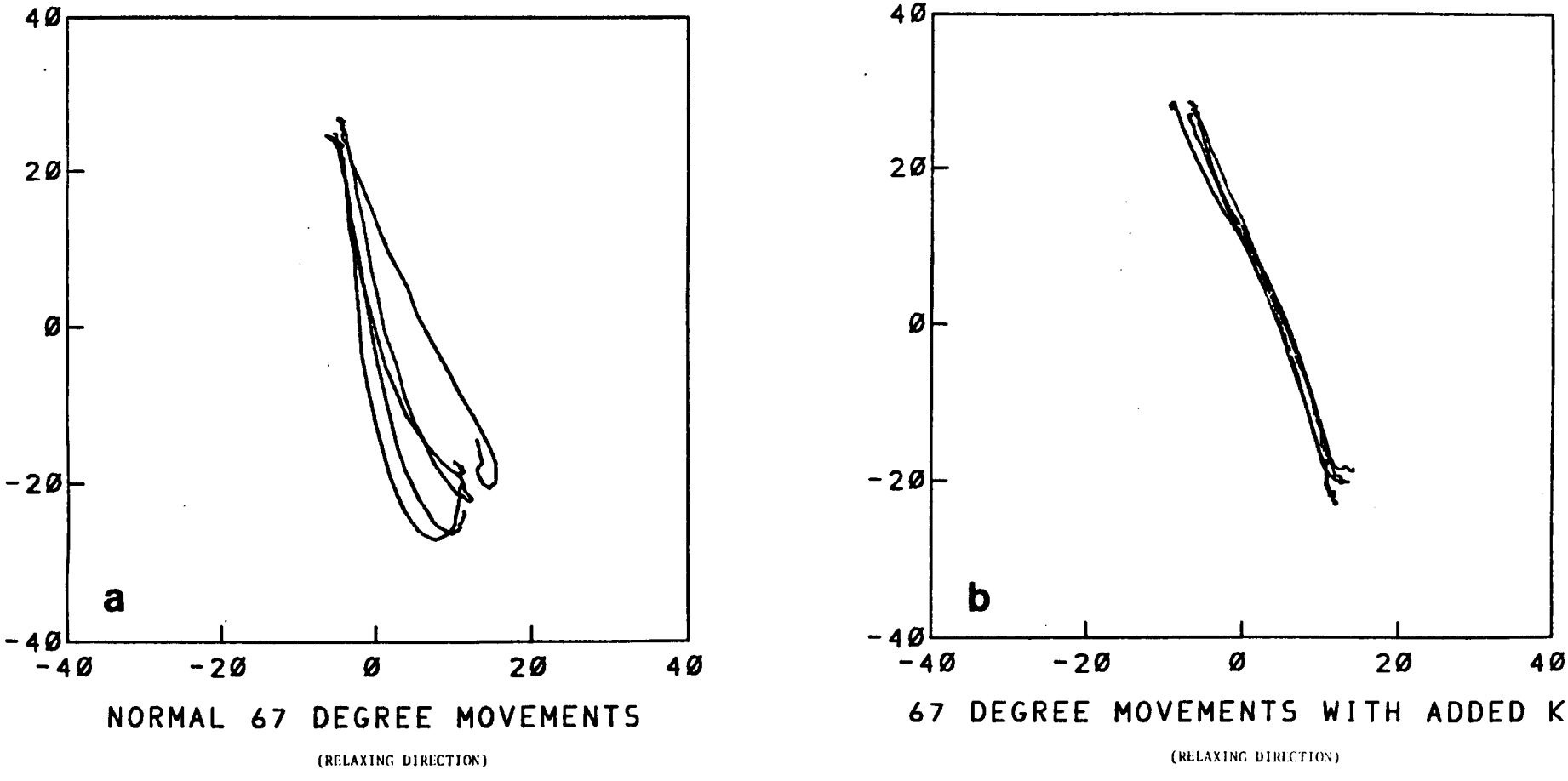


Fig. 3

Oblique Movement Trajectories. Trajectories of oblique movements at an inclination of 67 degrees. a) Unloaded movements; the direction of motion is from the upper left to the lower right; b) Movements made with added elastic load in the horizontal direction; the movements displayed are in the direction of relaxing the added spring.

present but they are hard to distinguish with our current resolution and will be considered to have secondary importance in the following discussion. P3 is of markedly higher amplitude and is more sharply peaked in the 60 deg. movements than P1 and P2 (Ghez & Martin 1982,Wadman, Van der Gon & Derksen 1980, Lestienne & Bouisset 1971,1974).

Added B

Addition of a viscous load to the head produces changes in the movement dynamics and EMG signals that demonstrate their close relationship (Fig. 4, thick lines). Comparison of loaded versus unloaded 60 deg movements shows a reduction in velocity and acceleration. In the loaded condition, the subjects were still attempting to complete the movement as fast as possible and had fully adapted to the loading by making practice movements before recording. The main difference observed is the reduction in height of P3 and a corresponding decrease in the magnitude of the stopping acceleration pulse. This suggests that the subject has developed an alternate control strategy with which to make time optimal movements in the loaded case.

Effects of Amplitude with added B

At 20 degrees, addition of viscous load results in almost total elimination of P3 when compared to 60 degrees (Fig. 7). In the 40 deg. case, P3 is also somewhat reduced and is similar in shape to that of the 60 degree average. P2 appears to be about the same in the 20 and 60 degree cases, and somewhat larger with viscosity in the 40 degree case (Lestienne 1979).

Another important effect of viscous loading is seen on the timing of the various pulses. Consistently observed in the 20 and 40 degree records, is an earlier onset of P1 when the two traces are synchronized at the first change of position. Since the EMG onset is causally first in the chain of events, synchronizing the records at the onset of P1 would reveal an extra 20ms dynamical lag between force onset and movement Zangemeister, Lehman & Stark 1981a,b) with the viscous load.

Our systematic variation of loads produced the array of averages displayed in Figure 5. Fig. 5a displays movements made to the right. Fig. 5b shows leftward movements. Across, are records of normal, viscous, inertial, and elastic loads. The addition of elasticity was accomplished by the attachment of springs in such a way that movements to the left relaxed the spring while those to the right stretched it. In the stretching direction, where the spring aids in stopping the movement, the antagonist pulse is almost completely suppressed. A note of caution must be sounded in making these observations however because of the asymmetry of this added load. Examination of the

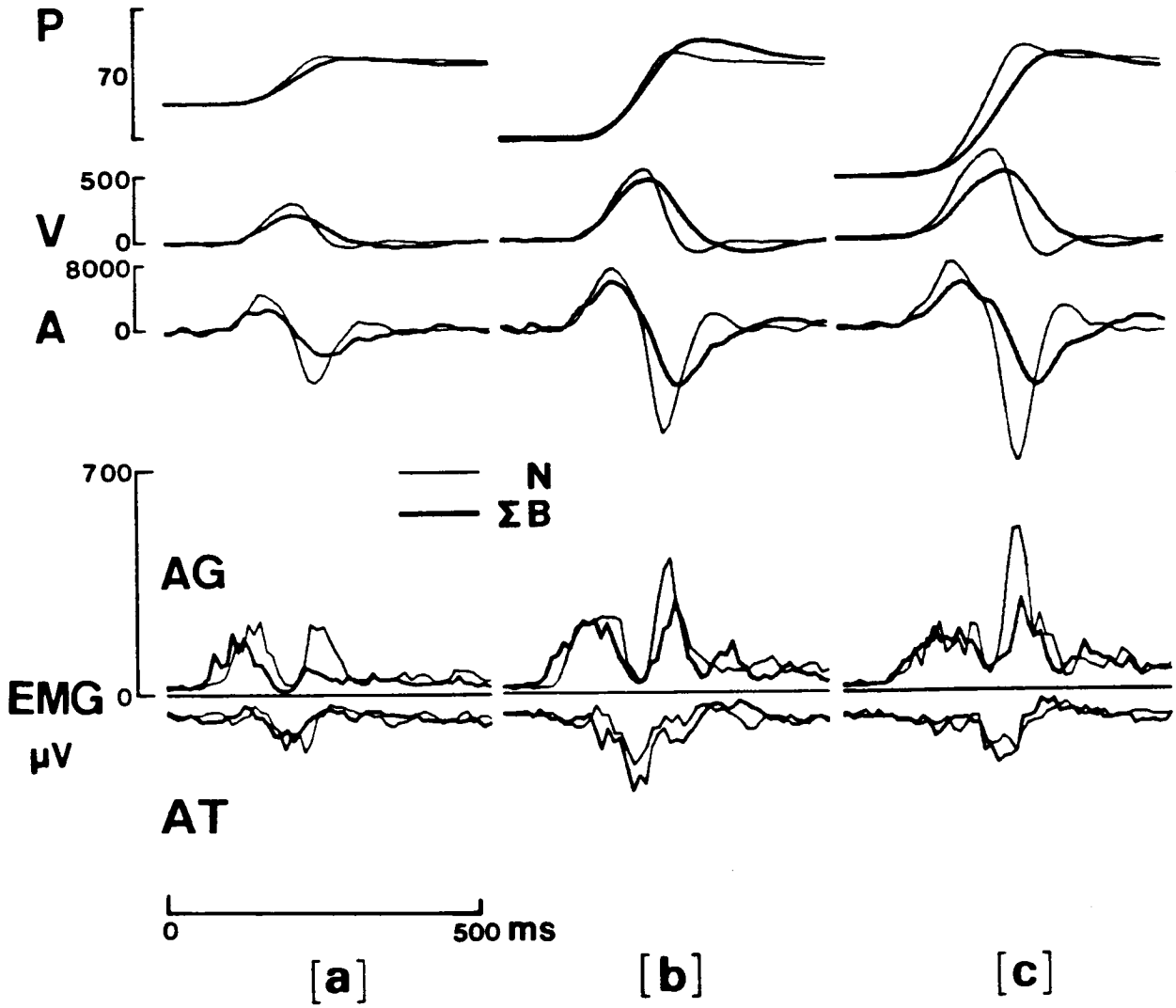


Fig. 4

Effects of Added Viscosity. Comparison of normal movements with those made with a viscous load device attached to head. Traces are ensemble averages of position, velocity, acceleration, and rectified EMG from agonist and antagonist muscles. EMGs are plotted oppositely to suggest a controller signal. a)20 deg., b) 40 deg., c) 60 degree target displacements.

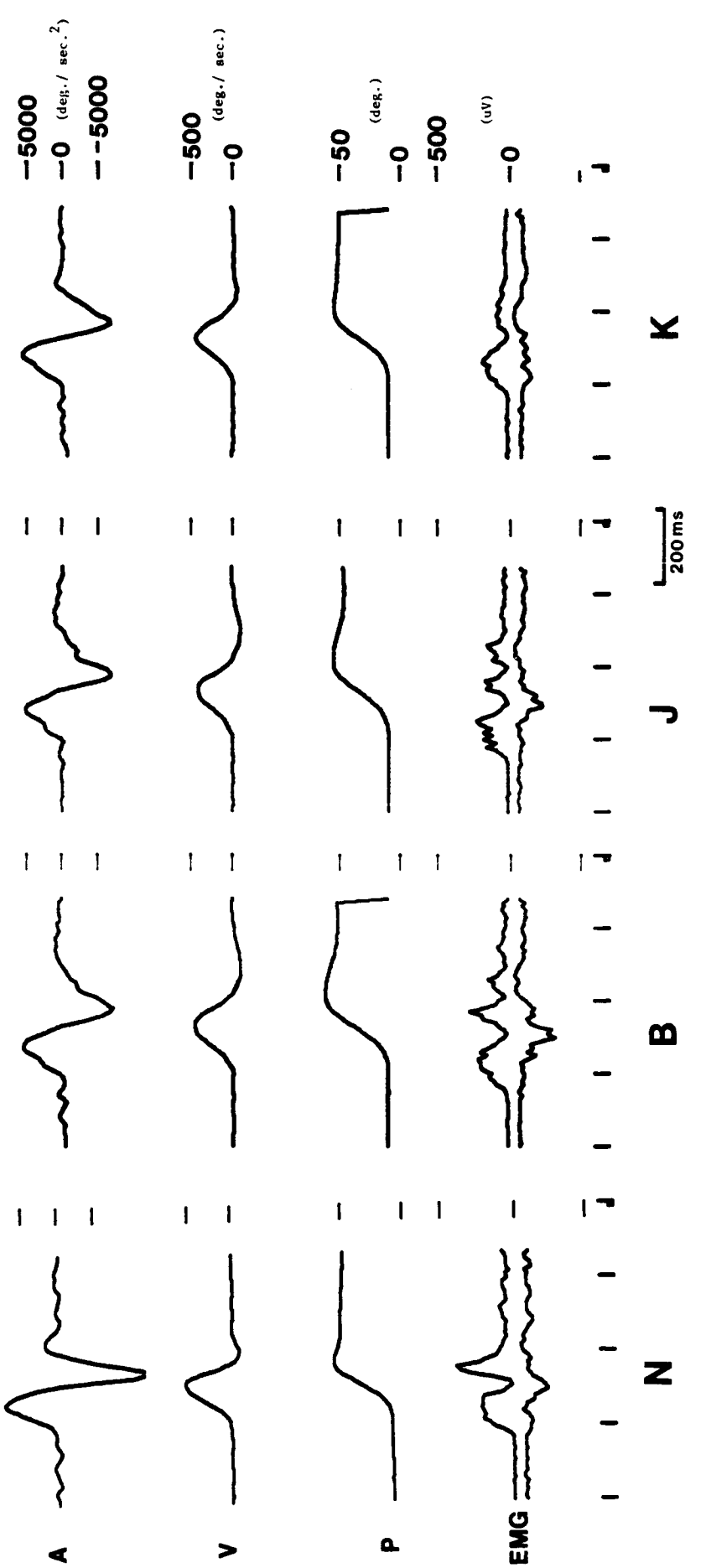

A
V
P
EMG
N
B
J
K
−5000
−0
−−5000
(deg./ sec.^{2})
−500
−0
(deg./ sec.)
−50
−0
−500
(deg.)
(uV)
−0
200 ms

Fig. 5a

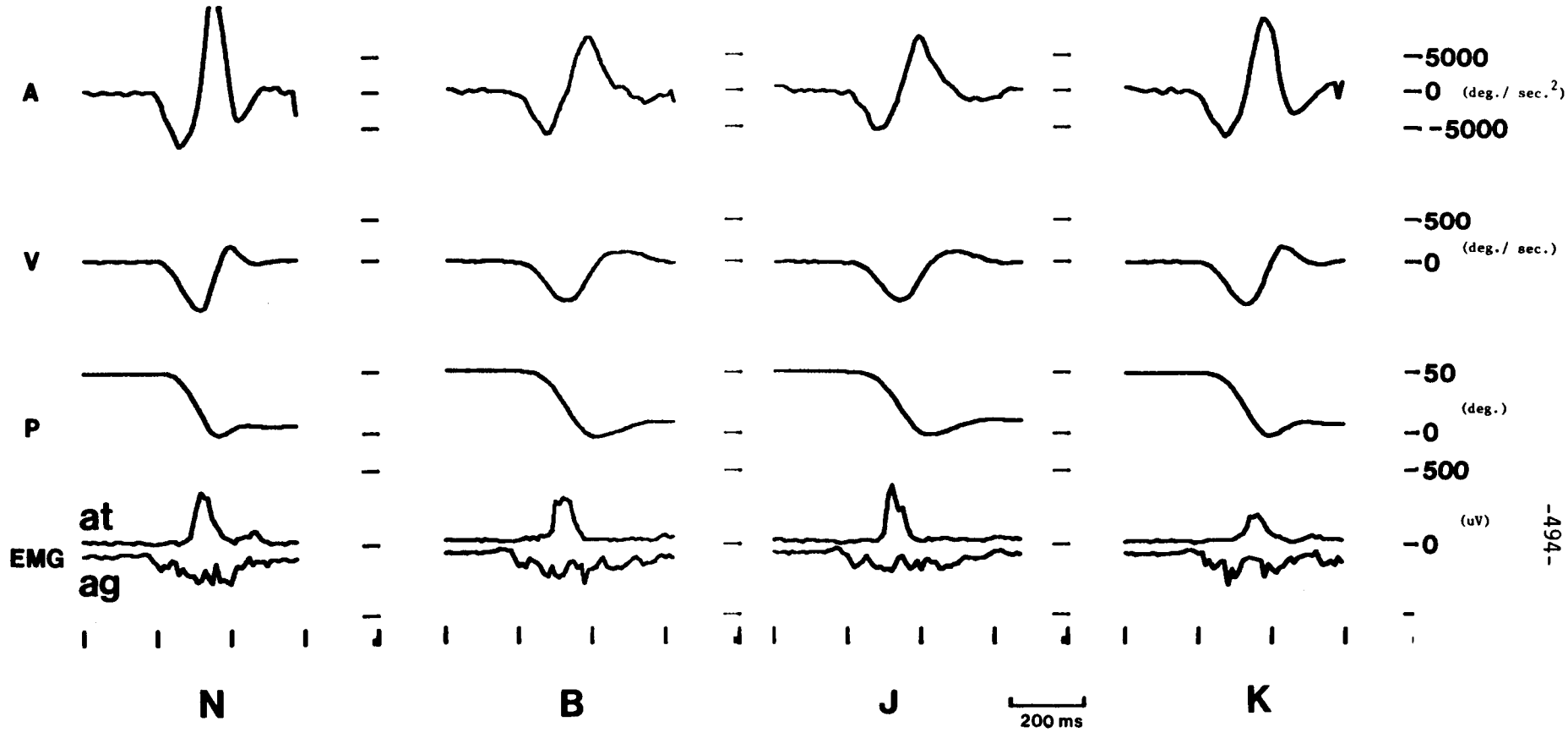


Fig. 5b

Fig. 5

Variation of Added Loads. Head movement records under loading conditions of Normal (N), added viscosity (B), added inertia (J), and added elasticity (K). a) movements made to the right (elastic load is relaxed). b) movements made to the left (elastic load is stretched).

records from the symmetrically loaded cases (normal, added inertia, and added viscosity) shows asymmetry in many of the EMG averages. Thus, more experiments should be done in which elastic load is applied from both directions in order to control for this effect.

Inertial loading is symmetrical and its effects are seen in EMGs from both the right and leftward movements in the form of an additional agonist pulse at the end of the movement. Asymmetry is evident here as well however, chiefly in the large difference in amplitude between the antagonist pulses in the two cases.

DISCUSSION

Horizontal head rotations are generated via the bilaterally symmetrical head neck muscle array composed of the sterno-cleido-mastoid, splenius capitus, semispinalis capitus, trapezius, and other, less prominent muscles (Zangemeister & Stark 1981b, Gray, H. 1858). Vertical head rotations are generated by a different utilization of these same muscles complexly rotating the head about a moving axis of rotation(Viviani & Berthoz, 1975, Zangemeister & Stark, 1981b, Grey 1858). As noted by Zangemeister (Zangemeister & Stark, 1981b), vertical rotations often synergistically utilize muscle pairs which are reciprocally innervated during horizontal rotations as agonist-antagonist pairs. This is complicated by the reorganization of these same muscles when the lateral tilt of the head varies.

EMG Causally Related to Trajectory Details

Perhaps the most obvious relationship between the muscle's electrical activity and movement dynamics appears in Fig. 6 in which the EMG complex (P1,p2,p3) appears to have a basic shape that is scaled to match the magnitude of the movement. Another striking feature is the apparently paradoxical relationship between the magnitude of the negative acceleration peak and that of the third pulse P3 (the second agonist pulse) in the averaged EMG signal. At first glance, a control strategy seems apparent that scales the width of the first agonist pulse and the height and width of the second agonist pulse according to the magnitude of the desired time optimal movement.

Bang-Bang Control

In spite of our systematic variation of target displacement and external loading, the height of P1 remained relatively constant. Although different control strategies were required of the subject, variation of pulse height was apparently not an available option when the subject had to maintain time optimality. Pontryagin's maximum principle (Lehman & Stark 1979, Zangemeister, Lehman & Stark 1981b, Pontryagin, Boltyanskii, Camkrendze, Mischenko 1962) asserts that for time optimality, the resulting controller signal must be of the "bang-bang" type, that

is, either fully on or fully off. Interestingly, the height observed for all P1's is approximately equal to the maximum averaged EMG when measured isometrically with the largest resistible force (about 8 kg) being applied to the head (Hannaford, Nam, et. al. 1983). The power of the bang-bang model to explain the height and width of P1 as a function of movement dynamics is strong evidence that EMG envelopes reflect an underlying controller signal at least during the initial phases of the movements.

Before we jump to the conclusion that these results prove a bang-bang control strategy, we must account for considerable height variations of EMG in P3 under many of the experimental conditions. Two possibilities are apparent: 1), A higher force is actually being exerted by the muscle during the later phase of the movement (voiding the constant pulse height assumption of the bang-bang model). and 2), The relationship between EMG and muscle force changes with some dynamical variable such as position, velocity, or acceleration. Modeling studies should help to distinguish between these two possibilities.

Effects of Load

With the addition of a viscous load, changes in EMG reveal a corresponding change in control strategy. The main change observed in EMG with added viscous load (again with time optimal subject intent), is the reduction in height or complete suppression of P3. It has been suggested (Lestienne 1979, Bizzi, Dev, Morasso, Polit 1978) that the third EMG pulse in conjunction with the muscle's non-linear viscosity provides extra viscosity to dampen the final stages of the movement. It is apparent from Fig, 5 that any viscous damping function provided by P3 is not needed when extra passive viscosity is added. This relationship is further strengthened by observing the records of movements made with the added inertia. Added inertia increases the kinetic energy in the moving mass which must be dissipated in order to stop the head. The clearly evident fourth pulse in the added J records (Fig. 5a,b) can be interpreted in this light as an additional damping pulse required to dissipate kinetic energy that was beyond the capacity of P3 (Gauthier,Martin & Stark, 1981).

Data from oblique experiments illuminate the effects on one component of the head's dynamics of loads applied about the orthogonal axis. Two prominent effects are those on overshoot, and extrema of dynamics (Main Sequence parameters). Added horizontal elastic load (both stretching and relaxing) eliminated overshoot in the vertical direction. In the stretching movements (where the agonist contracts against the spring force), horizontal loading had no effect on vertical Main Sequence parameters supporting the independence of vertical and horizontal control strategies and kinematics. However, in the relaxing movements reduced peaks in vertical velocity and acceleration

suggest interdependence of vertical and horizontal control strategies and kinematics.

These phenomena illustrate the kinematic, mechanical, and possibly neurological interaction between horizontal and vertical head movements arising from the complex anatomical arrangements of the neck musculature. For example, increased tone in the left splenius muscle working against a horizontal spring would cause the head to tilt back unless a compensatory force is developed in its vertical antagonist, the right sterno-cleido-mastoid. Note that if the muscle were a linear actuator, this added tone would serve only to exactly cancel the vertical component of the spring balancing force. In a more realistic, non-linear muscle, this increased tone will modify parameters of the vertical movement even though the load is only applied in the horizontal direction. Thus, the multiple roles of neck muscles in horizontal, vertical, and side to side, movements, combined with their non-linearity causes modification of movement dynamics by applied loads in the orthogonal direction.

Horizontal rotation experiments in which loads were added showed the adaptation of the nervous control signal (both central and reflex components) to the added loads. The oblique experiments showed that a load applied to the horizontal direction can influence the normally independent dynamics of the vertical movements and their associated control system. This is explained by the fact that both horizontal and vertical rotation systems utilize the same anatomical structures and are coupled through the non-linearity of the muscles.

ACKNOWLEGMENTS

We are pleased to acknowledge partial support from the NCC 2 - 86 Cooperative agreement, NASA-Ames Research Center, from the National Institutes of health training grant in Systems and Integrative Biology, and a grant from the Ministry of Education, Government of Korea. M. H. Nam is presently with the Dept. of Electrical Engineering, Kon-Kuk University, Seoul 133, Korea.

REFERENCES

1. E. Bizzi, R.E. Kalil & V. Tagliascc, "Eye-head coordination in monkeys - Evidence for centrally patterned organization," Science, vol. 173, pp. 452-454

2. E. Bizzi, R.E. Kalil, P. Morasso & V. Tagliasco, "Central programming and peripheral feedback during eye-head coordination in monkeys," Bibl. Opthal. vol. 82, pp. 220-232, 1972.

3. E. Bizzi, P. Dev, P. Morasso & L. Polit, "Effect of load disturbances during centrally initiated movements," J. Neurophysiol. vol. 41, pp. 542-556, 1978.

4. S. Bouisset, F. Lestienne & B. Maton, "The stability of synergy in agonists during execution of a simple voluntary movement" Electroencephal.Clin.Neurophysiol. vol. 42, pp. 543-551, 1977.

5. S. Bouisset & F. Goubel, "Influence of the termination of movement on the relationship between mechanical work and IEMG of the principal agonist," Electroencephal. Clin. Neurophysiol., vol. 25, pp. 395-402, 1968.

6. J.V. Danoff, "The integrated electromyogram related to angular velocity," Electromyogr.Clin.Neurophysiol., vol. 19, pp. 165-174. 1979.

7. G.M. Gauthier,B. Martin & L. Stark, "Effect of inertial loads on head-eye movements," OMS'81, Calif. Inst. Technology, Pasadena Calif., January 1981

8. C. Ghez & J.H. Martin, "The control of rapid limb movement in the cat III- Agonist-Antagonist coupling," Exp.Brain Res., vol. 45, pp. 115-125, 1982.

9. H. Grey, "Grey's Anatomy", 24th Edition, Running Press, Philadelphia, 1974.

10. B. Hannaford, M.H. Nam, V. Lakshminarayanan & L. Stark, "EMG as Controller Signal with Viscous Load," Submitted, 1983

11. S. Lehman & L. Stark, "Simulation of linear and nonlinear eye movement models: Sensitivity analyses and enumeration studies of time optimal control," J. Cyber. Info. Sci., vol. 4, pp. 21-23, 1979.

12. F. Lestienne, "Effects of inertial load and velocity on the braking process of voluntary limb movements," Exp. Brain Res., vol. 35, pp. 407-418, 1979.

13. F. Lestienne & S. Bouisset, "Quantification of the biceps triceps synergy in simple voluntary movements," in Visual Information Processing and Control of Motor Activity, Proc. of the Int. Symposium 1969, Bulgarian Acad. of Sciences, Sofia, pp. 445-448, 1971.

14. F. Lestienne & S. Bouisset, "Role played by the antagonist in the control of voluntary movement," in Advances in External Control of Human Extremities, vol. 1, pp. 12-21, Gavrilovic and Wilson, eds. Etan Press, Belgrade, 1974.

15. S. Lebozec, B. Maton & J.C. Cnockaert, "The synergy of elbow extensors muscles during dynamic work in man I. Elbow extensions," Eur. J. Appl. Physiol., vol. 44, pp. 259-269, 1980.

16. B. Maton, S. LeBozec & J.C. Cnockaert, "The synergy of elbow extensors muscles during dynamic work in man II-Braking of elbow flexion," Eur. J. App. Physiol., vol. 44, pp. 271-278, 1980.

17. M.H. Nam, V. Lakshminarayanan, E. Lipiansky, & L. Stark, "Effect of Viscous Load on Head Movements," IEEE Trans. Biomed. Eng., To Be Published 1983

18. L.S. Pontryagin, V.C. Boltyanskii, R.V. Camkrendze & E.F. Mischenko, "The mathematical theory of optimal processes," Trans. K.N. Trirogoff, Interscience Publishers, John Wiley & Sons , New York, 1962.

19. L. Stark, W. Zangemeister, et.al. "Head rotation trajectories compared with eye saccades by main sequence relationships," Invest. Opthal. Vis. Sci., vol. 19, pp. 986, 1980.

20. C.A. Terzuolo, J.F. Soechting & P. Viviani, "Studies on the control of some simple motor tasks I. Relations between parameters of movement and EMG activities," Brain Res., vol. 58, pp. 212-216, 1973.

21. P. Viviani, & A Berthoz, "Dynamics of the head-neck system in response to small perturbations," Biol. Cyber., vol. 19, pp. 19, 1975.

22. W.J. Wadman, J.J.D. van der Gon, R.H.Geuze & C.R. Mol, "Control of fast goal directed arm movements," J. Hum. Move. Stu., vol. 5, pp. 3-17, 1979.

23. W.J. Wadman, J.J.D. van der Gon & R.J.A. Derksen, "Muscle activation patterns for fast goal directed arm movements," J. Hum. Move. Stu., vol. 6, pp. 19-37, 1980.

24. W. H. Zangemeister, & L. Stark, "Active head rotations and eye head coordination", Annals N.Y. Acad. of Sci., pp.540-559, 1981a.

25. W.H. Zangemeister, & L. Stark, " Dynamics of head movement trajectories: Main Sequence Relationship," Exp. Neuro., vol. 71, pp.76-91, 1981b.

26. W.H. Zangemeister, L. Stark, O. Meienberg, & T. Waite, "Neurological control of head rotation: electromyographic evidence," J. Neurological Sci., vol. 55, pp. 1-14, 1982.

27. W.H. Zangemeister, S. Lehman & L. Stark, "Simulation of head movement trajectories: model and fit to main sequence," Biol. Cybern., vol. 41, pp. 19-32, 1981a.

28. W.H. Zangemeister, S. Lehman & L. Stark, "Sensitivity analysis and optimization for a head movement model," Biol. Cybern., vol. 41, pp. 33-45, 1981b.

A CHANNEL CAPACITY MODEL FOR THE ENCODER OF A MOTOR NEURON

William D. O'Neill
Department of Industrial and Systems Engineering
University of Illinois at Chicago
Chicago, Illinois 60680

Abstract

In any manual task, the information sequence that eventually controls muscle contraction, whether generated in the brain or processed through reflex arcs, must be transmitted through motor neurons. This paper derives a rigorous information channel model for the encoder of such neurons under the assumption that the neuron is of the integrator type. The encoder is shown to fulfill the definition of a finite state, indecomposable channel subject to intersymbol interference and fading caused by endogenously generated neuron threshold noise. The finite state of the channel is shown to be the state of an ergodic Markov chain. Numerical results are presented that illustrate how channel capacity of the encoder model can be calculated when the threshold noise is Gaussian. The frequently observed phenomenon of neuron "bursting" is given the interpretation of being an encoder transition to a non-ergodic Markov chain possibly caused by neurologicaly induced membrane depolarization. The model clearly indicates the central role played by the noise observed in the transmembrane potential of all neurons and quantifies its effect on motor neuron capacity.

Introduction

Any manually controlled task is limited in its position and velocity precision by the information capacity of those tandem elements processing the information required of the task. One such element is the neuron which converts dendritic excitatory and inhibitory analog inputs into action potentials of an essentially digital nature. Figure 1 is a schematic dichotomy of a general neuron into encoder, channel and decoder. We are concerned with the encoding process from the sum of the dendritic inputs $x(t)$ to the action potential $y(t)$.

Our analysis will result in a statistical description of the dependence of $y(t)$ on $x(t)$ under the assumption that the encoder is the classical integrator neuron (1). We further assume that the frequently observed threshold noise is also present and, as an exogenous input to encoder, is the limiting factor in information transmission through the encoder (2). The integrator model admits a single state variable $s(t)$ which, in the statistical description , becomes the state of a Markov chain.

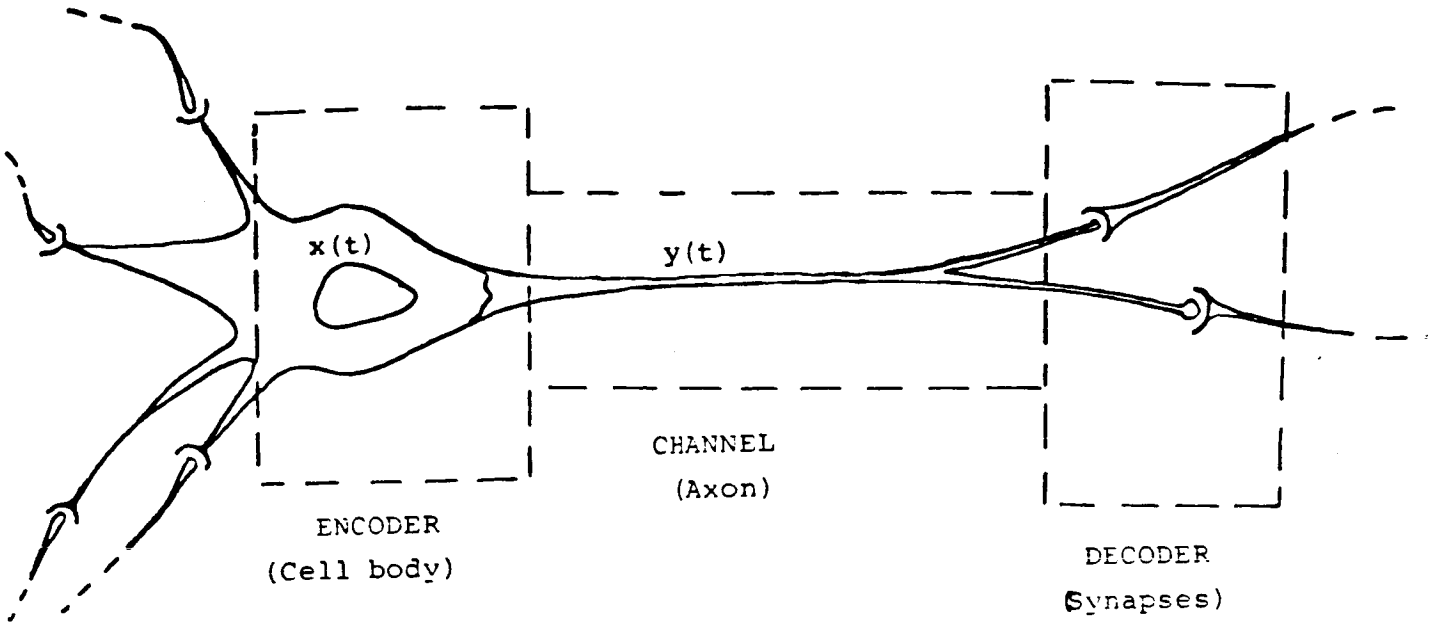
In a following section we assume that the threshold noise is approximately Gaussian distributed and compute the input distribution that $x(t)$ must follow in order to achieve encoder capacity. Unlike frequent approximate analyses, this distribution is decidedly non-Gaussian (3,4). Indeed, we show that the distribution achieving capacity is bimodal.

The Encoder Model

The integrator model relates $x(t)$ and $y(t)$ by the equation

$$\frac{dy}{dt}=x(t) \ , \ y(0)=0 \ , \ t>0.$$

When $y(t)$ exceeds the cell threshold, a brief action potential is generated


x(t)
y(t)
CHANNEL
(Axon)
ENCODER
(Cell body)
DECODER
(Synapses)

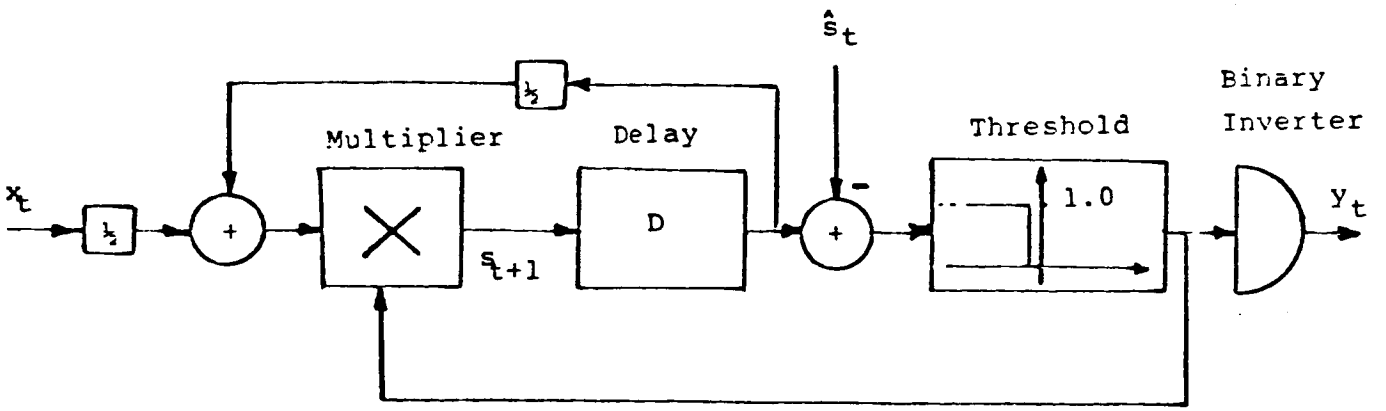

$\hat{s}_t$
Binary
Inverter
Multiplier
Delay
Threshold
1.0
x_t
s_{t+1}
y_t
D

FIGURE 1.

and y(t) is quickly reset to zero, the resting potential. We denote the time course of the action potential as t_r. Our analysis will use discrete time, measured in units of $t_r/2$, and the approximation

$$\frac{dy}{dt} \approx \frac{1}{\Delta t}(y(t+\Delta t)-y(t))=y_{t+1}-y_t$$

We define the discrete- time, state of the system as $s_t \equiv y_t$ for $y_t < \hat{s}_t$, the neuron threshold. The equations

$$s_{t+1}=\tfrac{1}{2}(s_t+x_t)[1-\mathrm{sgn}(s_t-\hat{s}_t)]$$

$$y_t=\tfrac{1}{2}[1+\mathrm{sgn}(s_t-\hat{s}_t)]$$

summarize our assumptions. Note from the system diagram, Figure 1, that y_t has been scaled to a zero-one binary action potential.

The information model of the neuron consists of two elements. The first is a Markov chain whose probability vector gives the probability of all possible values the state s_t can take on. The second is the probability transition matrix relating the range of possible values x_t can assume to the 0 and 1 values of y_t. This transition matrix depends on the probability vector of the Markov chain.

Let the range of $\hat{s}_t$ be 0,1,2,...,M mv. and suppose x_t assumes values in 0,1,2,...,N mv. Define

$$g_t=\tfrac{1}{2}[1-\mathrm{sgn}(s_t-\hat{s}_t)].$$

Then

$$s_{t+1}=(s_t+x_t)g_t.$$

If $g_t=1$ when $\hat{s}_t=M$ then $s_t<M$ and the maximum value for s_t is

$$s_{t+1}=(M-1+N)(1)=M+N-1.$$

Therefore the range of s_t is 0,1,2,...,M+N-1 mv. We assume that x_t is identically and independently distributed with proabilities

$$q_i=p(x_t=i) \qquad i=0,1,2,\ldots,M$$

and that $\hat{s}_t$ is similarly distributed independently of x_t with

$$p_j=p(\hat{s}_t>j) \qquad j=0,1,\ldots,M-1$$

If p_t is the vector of probabilities $p_t^i=p(s_t=i)$, i=0,1,2,...,M+M-1 then p_t satisfies the Markov chain

$$p_{t+1}=Qp_t$$

where the M+N rows of Q are

$$r_1=1+p_0(q_0-1)\ ,\ 1-p_1\ ,\ldots,\ 1-p_{M-1}\ ,\ 1\ ,\ldots,\ 1$$

$$r_2=q_1p_0\ ,\ q_0p_1\ ,\ 0\ ,\ldots,\ 0$$

$$r_3=q_2p_0\ ,\ q_1p_1\ ,\ q_0p_2\ ,\ 0\ ,\ldots,\ 0$$

$$\vdots$$

$$r_M=q_{M-1}p_0\ ,\ q_{M-2}p_1\ ,\ldots,\ q_0p_{M-1}\ ,\ 0\ ,\ldots,\ 0$$

$$\vdots$$

$$r_{N+1}=q_Np_0\ ,\ q_{N-1}p_1\ ,\ldots,\ q_{N-M+1}p_{M-1}\ ,\ 0\ ,\ldots,\ 0$$

$$r_{N+2}=0\ ,\ q_Np_1\ ,\ q_{N-1}p_2\ ,\ldots,\ q_{N-M+2}p_{M-1}\ ,\ 0\ ,\ldots,\ 0$$

$$\vdots$$

$$r_{N+M}=0\ ,\ldots,\ 0\ ,\ q_Np_{M-1}\ ,\ 0\ ,\ldots,\ 0.$$

In this representation we have assumed that N>M. If we assume that none of the q_i or p_j are zero then this chain is irreducible and eliminating the deterministic cases q_i of p_j=1 implies the chain has period 1. Thus the chain is, under these assumptions, also ergodic (5). Consequently, the chain has a unique equilibrium vextor p_e which is approached exponentially in time and satisfies

$$p_e=Qp_e \qquad (1)$$

We note here that if the chain were not ergodic then there exists the possibility that it might have a periodic solution of period greater than one. Such a situation might well be related to neuron bursting which is usually accompanied by gross variations in threshold. Since the range of $\hat{s}_t$ is significantly increased during such episodes, our analysis shows that the possible states of the chain are simultaneously increased. Thus, the probability of establishing a cyclic set of states of period greater than one would appear to be high.

To establish the transition probabilities $P(y_{t+1},s_{t+1}|x_t,s_t)$ we observe that s_t is independent of y_t so that the encoder fulfills the definition of a finite state, indecomposable channel. This follows because the chain ergodicity insures that for some fixed t and any sequence of x_t, $P(s_t|x_t,s_0)$ is positive for all s. The probabilities $P(y_{t+1}|x_t,s_t)$ are defined by $P(0|x_t,s_t)$ and $P(1|x_t,s_t)=1-P(0|x_t,s_t)$. In terms of the above notation we get

$$P(0|x,s)=p_{s+x}p_s+p_0(1-p_s) \qquad \begin{matrix} s=0,1,2,\ldots,M+N-1 \\ x=0,1,2,\ldots,N \end{matrix} \qquad (2)$$

A line diagram interpretation of the encoder transition matrix is shown in Figure 2. Only $P(0|x,s)$ are shown since $P(1|x,s)$ are the corresponding complements.

Encoder Information Capacity

The information capacity of the encoder in terms of the mutual information $I(x;\ y|s_0)$ is given by

$$C=\lim_{t\to\infty} 1/t(\max_{q} \min_{s_0} I(x;\ y|s_0))$$

In the limit, the Markov chain comes to equilibrium at p_e so that in

$s_t = s$; $s = 0, 1, \ldots, N+M-1$; $p_i = 0$ for $i > M-1$

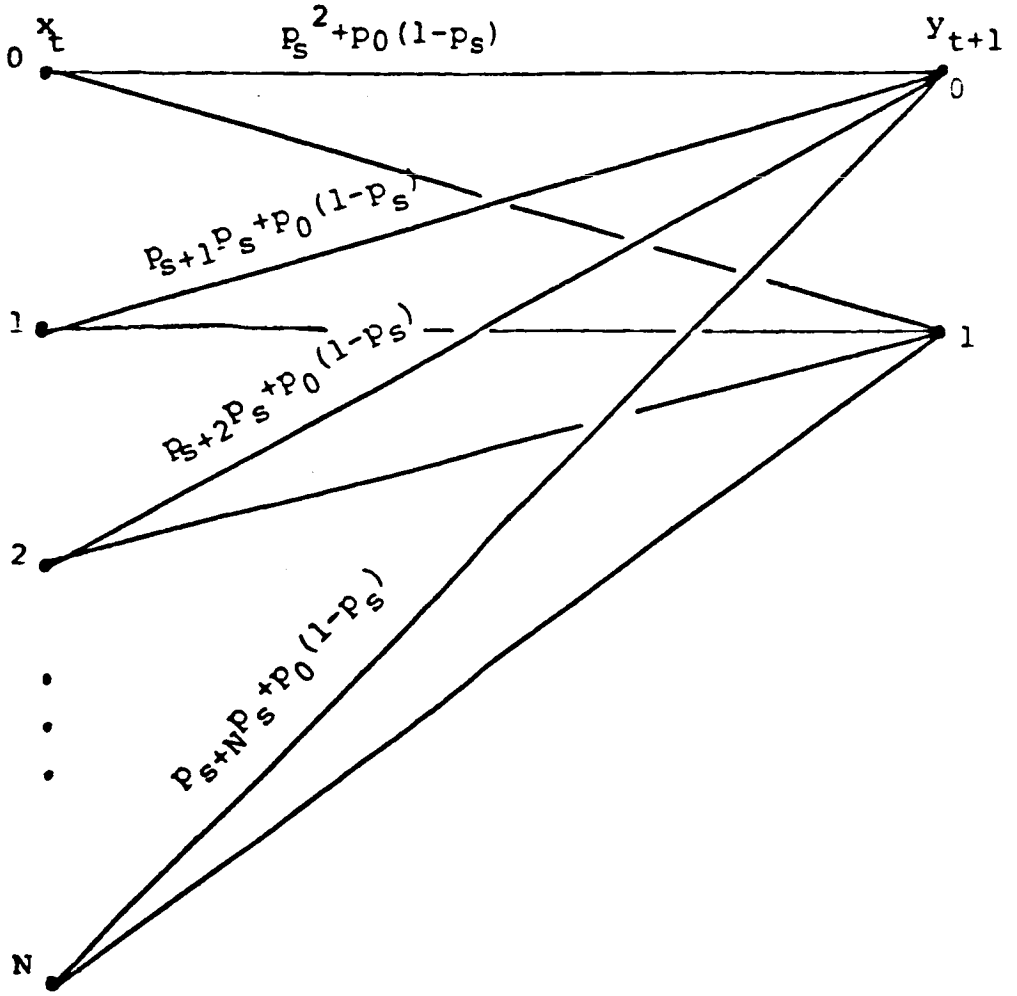


FIGURE 2.

equilibrium

$$P(y|x) = \sum_{s=0}^{M+N-1} P(y|x,s)p_e^s .$$

This implies that capacity is

$$C = \max_{q} \sum_{x=0}^{N} \sum_{y=0}^{1} q_x P(y|x) \log \frac{P(y|x)}{\sum_{r=0}^{N} q_r P(y|r)}.$$

This maximization problem must be solved recursively since the equilibrium equation (1) cannot be solved for the p_e vector without knowledge of the q vector. For a given q distribution of x_t and cumulative distribution of the threshold noise p_j, j=0,1,2,...,M-1, p_e can be found and therefore $P(y|x)$ computed and $I(x;y)$ calculated. To find C we use the Blahut algorithm (6) and a matrix inversion for p_e and start the algorithm by taking x_t to be uniformly distributed. A numerical example follows.

Encoder Capacity with Gaussian Threshold Noise

We choose a Gaussian distribution for the distribution of the threshold noise in order to compare our model with several of the approximate neuron channel capacity models we have cited previously. In the cited cases, Gaussian threshold noise produced a Gaussian distribution of x_t to achieve capacity. None of the published models to our knowledge have accounted for the Markovian character of the encoder state. In our numerical results we have taken N=20, M=13 mv. and assume that $\hat{s}_t$ has a Gaussian distribution with expected value of 8 mv. and standard deviation of 2 mv. Since our model is discrete we approximated this density with $p(\hat{s}_t)$, $\hat{s}_t$ = 0,1,2,...,13 mv. and

$$p(\hat{s}_t)=\int_{\hat{s}_t-0.5}^{\hat{s}_t+0.5} N(8,2)d\hat{s}_t$$

where N(8,2) is the Gaussian density of mean 8 and standard deviation 2. With this distribution as input, Figure 3, the Blahut algorithum produced the capacity yielding input distribution also shown in Figure 3. Capacity was found to be 0.4629 bits/encoder use. Our assumed time base is in units of $t_r/2$ so that the encoder is used once every t_r seconds. Thus, a neuron with a refractory period of 10 ms. would have an encoder capacity of 46.29 bits/second.

Simulation and Sample Results

It is instructive to interpret the neuron model and its channel capacity distribution in the time domain. A random number generator can be used to generate a sample function input x_t from the capacity achieving distribution. Similarly, a sample function for $\hat{s}_t$ from the normal distribution assumed for the threshold can be generated. Taking $s_0=0$, iteration of the model equations will then produce a typical sample function of the binary action potentials of the neuron responding to a sample function from the capacity achieving input distribution. A 100 sample record (0.5 seconds real time) of the input x_t and output y_t are shown in Figure 4. Note that an encoder capacity of 46 bits per second, corresponding to a neuron refactory period of 0.1 ms., means that a typical sample of y_t should exhibit 46 binary ones per second on average. The sequence shown in Figure 4 shows 25 ones in 0.5 seconds suggesting that the sample results shown are close to being typical.

Observe, as noted below, that the sample of x_t clearly exceeds the capacity of the encoder to transmit the input information. This is a property

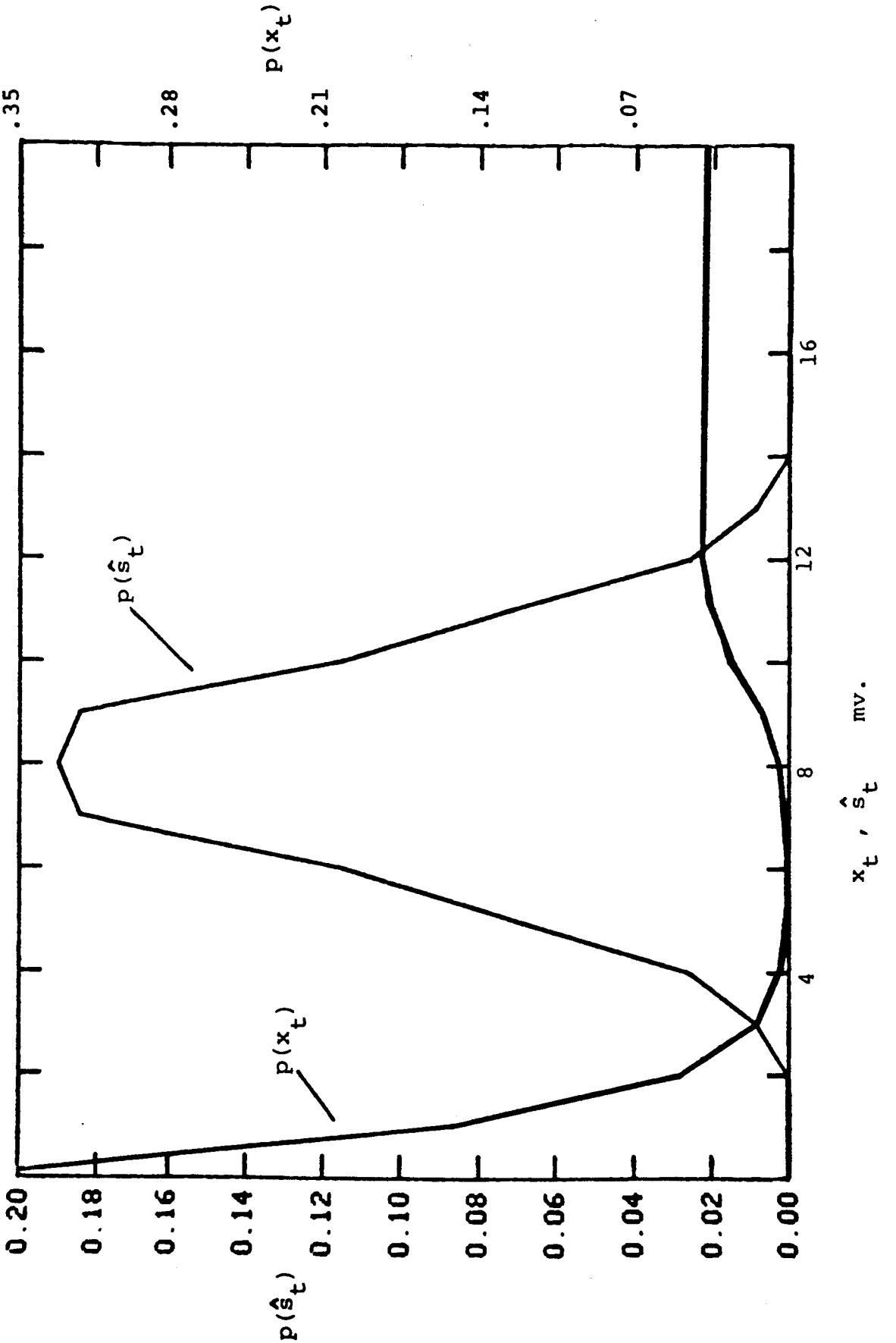
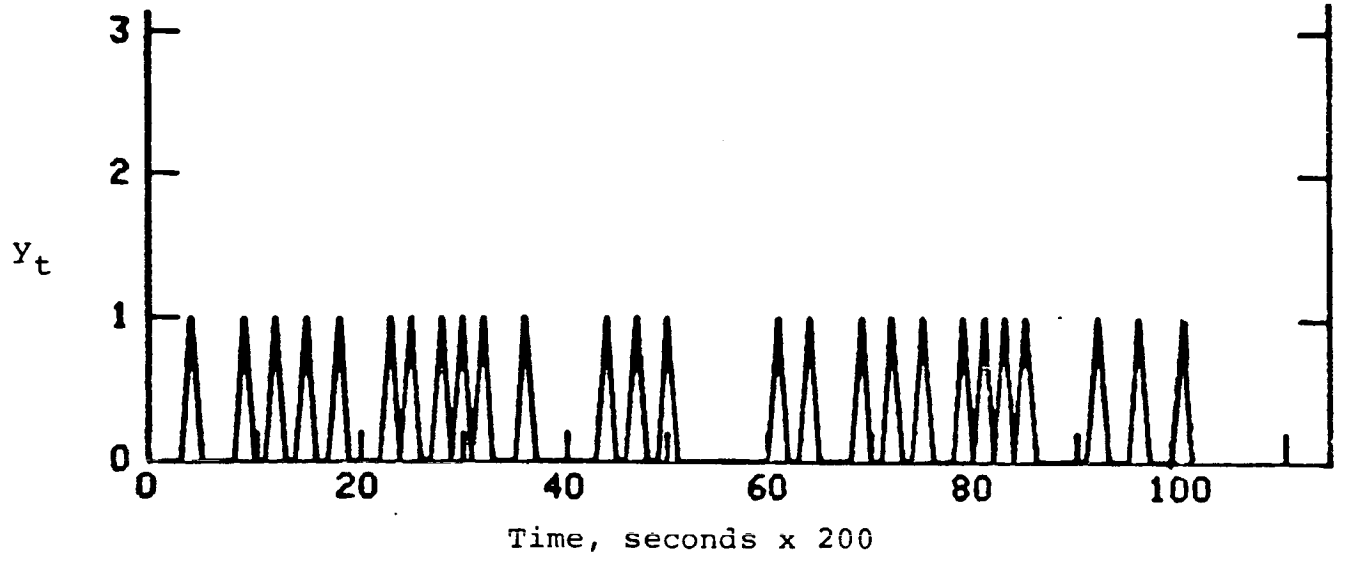

.35
.28
.21
.14
.07
p(x_t)
p(ŝ_t)
p(x_t)
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
p(ŝ_t)
4
8
12
16
x_t , ŝ_t mv.

FIGURE 3.


3
2
y_t
1
0
0
20
40
60
80
100
Time, seconds x 200

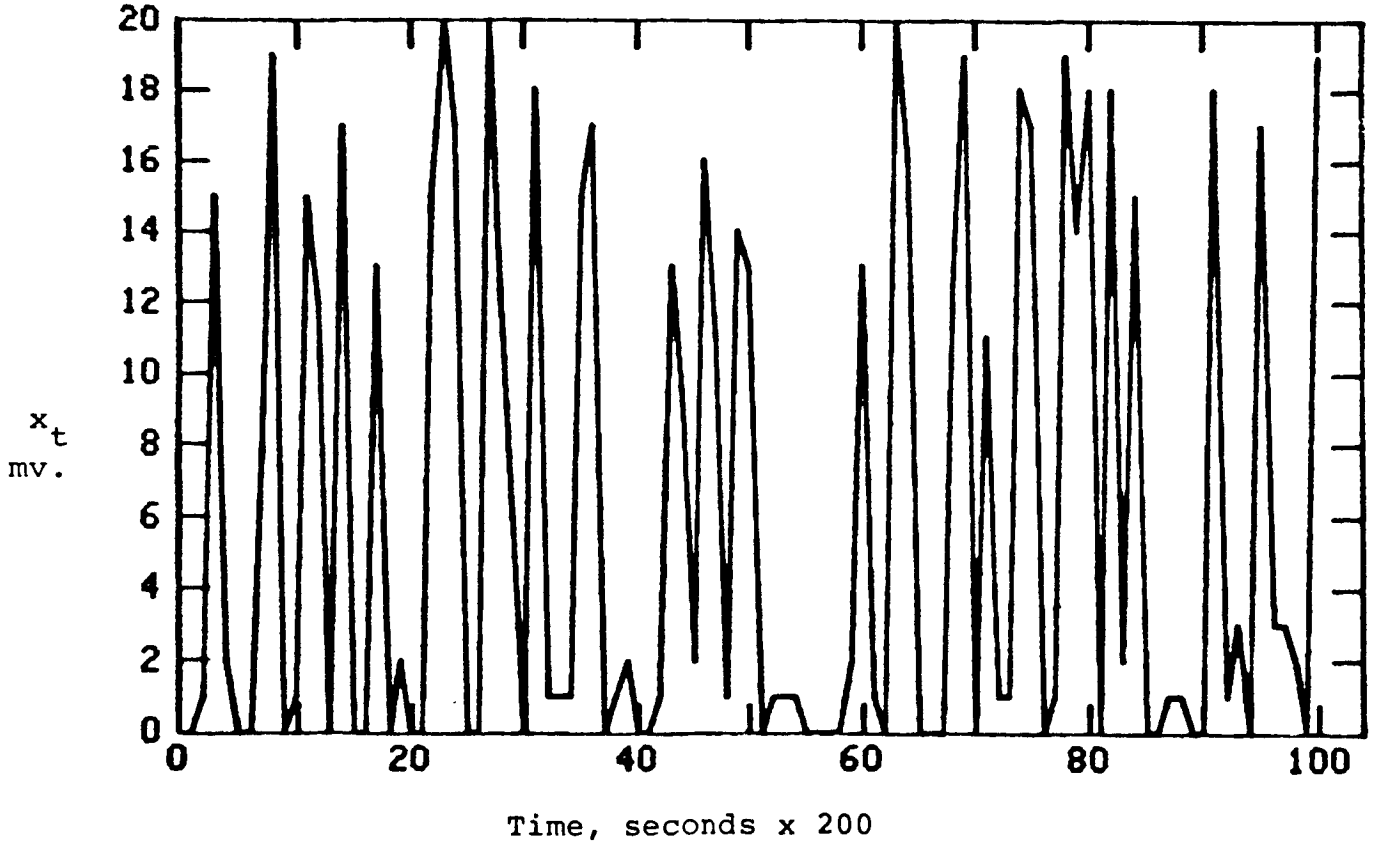

20
18
16
14
12
10
x_t
mv.
8
6
4
2
0
0
20
40
60
80
100
Time, seconds x 200

FIGURE 4.

of the optimal input distribution since it was selected to achieve capacity without regard to transmission errors. A capacity of 46 bits per second for the neuron means that for any input distribution with entropy less than 46 bits per second, a neuron code exists such that the information from the input distribution can be transmitted with arbitrarily small probability of error by means of the aforementioned code. Whether neurons actually employ such codes seems to be an open question. Perhaps a more interesting question for motor control is whether or not reflex arcs and other neurological feedback are employed to increase channel capacity since it is well known that feedback channels can have almost twice the capacity of similar single path channels (7).

Concluding Comments

Our model and capacity calculations would indicate that the integrating neuron encoder achieves capacity by selecting inputs that are, in probability, not in the range of the threshold noise. This result differs significantly from those obtained by other authors for the integrator neuron. A reading of their work indicates that their degree of approximation in the frequency domain has reduced their neuron models to memoryless encoders that ignore the Markov chain property of the state. Under this condition it is likely that Gaussian threshold noise would require a Gaussian input to achieve capacity.

From the input distribution that achieves capacity we can calculate the self information, or entropy, of x_t by the formula

$$H(x_t)=\sum_{i=0}^{20} q_i \log_2\left(\frac{1}{q_i}\right)=3.371 \text{ bits}$$

Therefore, on the average, the encoder input x_t would have $3.371/t_r$ bits of information per encoder use. For a t_r of 10 ms, this represents 337.1

bits/second. Since encoder capacity for a similar t_r is 46.29 bits/second, it is clear that the encoder would be subject to frequent transmission errors. Thus, the capacity achieving input distribution is a very unlikely distribution to be found in a functioning neuron encoder. In an *in vivo* neuron a far more likely state of affairs would be an input distribution that is multivariate in the lags of x_t, that is, one exhibiting a high degree of correlation between x_t and it lags. With regard to correlations existing in the threshold noise distribution, there appears to be ample evidence from inter-spike interval histogram studies that threshold noise is uncorrelated (8).

References

1. R.B. Stein, A.S. French, A.V. Holden: "The Frequency Response,Coherence, and Information Capacity of Two Neuronal Models", Biophysical Journal, Vol. 12, pp. 295-322, 1972.

2. A.V. Holden: Models of the Stochastic Activity of Neurones, Springer-Verlag, New York, 1976.

3. R. Eckhorn, B. Popel: "Rigorous and Extended Application of Information Theory to the Afferent Visual Stem of the Cat I. Basic Concepts", Kybernetik, Vol. 16, pp.171-200, 1974.

4. M. Abeles, Y. Lass: "Transmission of Information by the Axon II. The Channel Capacity", Biological Cybernetics, Vol. 19, pp. 121-125, 1975.

5. R.G. Gallager: Information Theory and Reliable Communiation, John Wiley and Sons, New York, 1968.

6. R.E. Blahut: "Computation of Channel Capacity and Rate-Distortion Functions", IEEE Transactions on Information Theory, Vol. IT-18, No. 4, pp. 460-473, July, 1972.

7. S. Butman, "A General Formulation of Linear Feedback Communication Systems with Solutions", IEEE Transactions on Information Theory, Vol. IT-15, pp. 392-400, May, 1969.

8. J.W. DeKwaadsteniet: "Statistical Analysis and Stochastic Modeling of Neuronal Spike-Train Activity", Mathematical Biosciences, Vo. 60, pp. 17-72, July 1982.

A GENERALIZED TREATMENT OF THE INTERACTION BETWEEN FAST AND SLOW MOVEMENTS

Jack M. Winters, Moon Hyon Nam* and Lawrence Stark

Departments of Engineering Science, Physiological Optics and Neurology,
University of California, Berkeley, CA 94720, USA

Acknowledgements: The authors are pleased to acknowledge partial support from the NCC 2-86 Cooperative Agreement, NASA Ames Research Center, the National Institute of Health Traning Grant in Systems and Integrative Biology and a grant from the Ministry of Education, Korea.

* Visiting professor from the Department of Electrical Engineering, Kon-Kuk University, Seoul, 133 Korea

INTRODUCTION

Of the wealth of types of human movement, eye movements have become one of the most widely studied neuromotor systems in the human and primate. The reasons for this are numerous. Of interest here is the fact that the eye, with its relatively simple mechanical system and its stereotyped movement types, is easier to understand and model than other more elaborate biomechanical systems such as the limbs. From a mechanical standpoint, the eyeball can be approximated as a rigid sphere of constant inertia that is rotated in the horizontal direction by "actuators" at either side. From a biocontrol standpoint, the four or five main types of eye movements - some fast, some slow -are distinct and very stereotyped and can be investigated independently under controlled experimental conditions. Some data also exists both for the nature of the neuro-control to the system and for the behavior of eye muscle. Furthermore, system "output", i.e. the eye angle, is observable.

In this presentation we will concentrate on eye movement interaction, since we have tremendous confidence in this model. Some preliminary results for head and limb fast/slow interaction will also be presented. However, since we are still actively refining some higher order, more representative models for these more complex neuromotor systems, we have chozen not to emphasize these preliminary findings which may soon become outdated.

A. The Interaction Problem

Eye movement interactions. Under natural conditions, it is well known that different types of eye movements, generated by specialised subsystems of the central nervous system (CNS), may interact spatially and temporally in a complex manner. Since, to a good first approximation, it appears that all types of eye movements use the same motoneurons, and consequently the same macroscopic neuromuscular plant (Keller (1981), Robinson and Keller (1972)), we can think of this interaction of eye movement types as being a convergence of neuro-controller signal types at the level of the oculomotor neuron and/or immediate supra-nuclear structures. However, it is evident that such interactions are not simply linear. Thus, the inherent nonlinearities need to be investigated - including studying the same type of interaction under a range of movement magnitudes.

Previous eye movement interaction studies. In recent years, the interactions between these different types of eye movements have been experimentally studied in humans and primates in a number of different ways. Some of work has been done in examining the interaction of "slow" movements with each other, such as the vestibul-ocular reflex (VOR) with the optokinetic reflex (OKR) (see Schmid et al (1980) for a review). Here, however, we will be considering the phenomena of a "fast" eye movement paired with a "slow" eye movement type. These interaction combinations include the saccade with: the VOR in monkeys (Morasso et al (1973)) and in the human (Jurgens et al (1977, 1980, 1981), Nam et al (1981)), vergence in the human (Ono et al (1978), Kenyon et al (1980)) and smooth pursuit in the human (Jurgens et al (1975)). The interaction between the fast and slow phases of vestibular nystagmus has also been investigated (Jurgens et al (1977, 1981), Chun and Robinson (1978)). The present paper is basically an extension of a previous work (Winters et al (1983)).

"Linear Summation" concept. These various eye movement investigators have

suggested a number of hypotheses to account for their respective findings. Central to the theoretical analysis is typically the idea of "linear summation", or the "additivity hypothesis", of controller signals. Here linear summation, which we will refer to in our figures by the "lim sum" line, refers to the result that would be obtained if the trajectory of the two movements, performed independently, were simply added when they occur together, i.e. summed for the duration of the fast movement. Note that, for such ideal superposition to apply, either the system is linear or any nonlinearities must somehow cancel. One of the main objectives of this paper is to show that, even for the eye movement system, there are nonlinear mechanical interactions at the plant that are significant and thus superposition does not hold. For the head and the limb, the functional nonlinear behavior of the force-velocity and length-tension properties of muscle are even more pronounced. Simple linear summation has been supported by Morasso et al (1972), and partial linear summation (for instance 70% of ideal superposition) has been found by the Jurgens group (1975,1981) for saccade/smooth pursuit interaction and by Nam et al (1981) and Jurgens group (1975,1981) for saccade/VOR interaction. Both Kenyon et al (1980) and Ono et al (1978) found nonlinear interaction for saccade/vergence interaction - more than simple superposition, and the former was the first to put forward evidence for some of this nonlinear interaction being due to the mechanics of the system (the "Kenyon effect"). "Squelching", or saccadic shut-off of the slow movement controller signal during the saccadic pulse, had been postulated by Chun and Robinson (1978) and Jurgens et al (1977, 1981) for vestibular nystagmus. Of the above, only Kenyon's group attempted to model the eye plant in order to help explain the observed behavior. Our present modelling efforts will make it possible for us to reinterpret some of these previous experimental data on this type of interaction.

B. The Simulation Approach

Models. Because of the ideal mechanical and control nature of the eye movement system, bio-control engineers have become deeply involved in trying to understand and model the various aspects of this system. In the present study, the sixth order reciprocally innervated model of the human orbital mechanics described in previous papers (Cook and Stark (1968), Clark and Stark (1974), Lehman and Stark (1979)) was modified and used to simulate the dynamics of the eye. Notice that, for the purposes of our model, the "system" is of a mechanical nature (the eye dynamics), the "input" is the neurological controller signal at the level of the oculomotor nuclei, and the "output" is eye position and its derivatives. Since we found it desirable to have the capability to just specify a saccadic magnitude (system output) without explicitly setting the controller input, we needed to create the "correct" causal controller signal inputs for our desired range of saccadic outputs. This was accomplished by modifying an existing optimization algorithm (Bremermann (1970), Lehman and Stark (1979)) and then using it to determine the time-optimal second order "bang-bang" neural controller signal input for the range of saccadic eye movements between 2 and 40 degrees (Winters et al (1983). Similar efforts have now been used for the head and limb systems. Thus, we see that simulation has enabled us to separate any nonlinear mechanical system phenomena from nonlinear behavior due to the neuro-control input.

Definition of simulation conditions. In this study, we reexamine the interaction of generalized "fast" and "slow" movements by computer simulation by having the fast, multi-pulse "saccadic" neuro-controller signal interact with a "base" sinusiodal signal that represents the neuor-controller signal for a "slow" type of eye movement, such as the VOR, or a slow prepatory limb movement. Furthermore, for the eye movement system, we do so both for a wide range of

"slow-movement induced" initial conditions at saccadic implimentation and for the wide range of time optimal saccadic sizes. We then compare these model results with our experimental data generated by having a subject voluntarily rotate his or her head (in approximately a sinusoid) while simultaneuosly following a target with the eyes. The resulting VOR/saccadic interactions seemed to be an ideal system for gaining insight into the general interactions of slow and fast eye movements over a wide range of initial conditions and saccade sizes. Since, to a first approximation, when the head rotates, the eyes, as part of the head rigid body, must also rotate, an equal and opposite compensatory eye movement is required to keep fixation on a given target. For our simulation studies, we assumed a VOR gain of -1 and a negligible time delay between head rotation and the VOR-induced neuro-controller signal - assumptions that appear to be justified (Winters et al (1983)). We could then concentrate on the eye system, and on the two possible levels of interaction: 1) those due to the known nonlinearities of muscle-plant mechanical system and 2) those due to interaction between the two neuro-controller input signals that generate the two types of eye movements.

Controller signal interaction possibilities. One of the main advantages of modelling is that a known controller signal is the input to the system. Thus, different types of controller signal interaction inputs can be given to our model allowing neuro-controller signal effects can be investigated separately from the mechanical system effects. It is important to look in detail into the nature of the interaction between the saccadic and the VOR neuro-controller signals, or, more generally, between "fast" and "slow" signals. At the simple interaction extremes, one can distinguish between, on the one hand, an ideal neurological summation of the two controller signals, which would result in the saccadic agonist and antagonist burst magnitudes being shifted up or down depending on the existing VOR controller signal level. In such a case, the VOR affects the input to the model, thus allowing both neurological, as well as mechanical, interaction. At the other extreme, one can envision an ideal nonlinear "squelching" phenomena, in which the "overpowering" saccade (or quick phase) temporarily eliminates, or dominates the sinsuoidal slow phase neurological controller signal during the saccadic pulse(s). Put another way, the neuro-controller signals would alternate. However, since these alternating signals are inputs to the same mechanical plant, there will coupling effects, i.e. initial conditions, due to dynamics of the system. In such a case any observed saccadic behavior that is different from an isolated saccade of the same magnitude can only be of a mechanical origin since the saccadic controller will always be the same and only the initial mechanical conditions will be different.

C. The Goal of This Presentation

In this introduction we have attempted to put the literature in perspective and to present a systematic simulation approach for decoupling the neurological interaction possibilities from the possible role of the system mechanics. Given the nonlinear plant mechanics, what type of neuro-controller signal will explain the experimentally observed phenomena? And is the role played by the perpheral nonlinear interaction of state variables in the nonlinear plant significant? To attempt to answer these questions for eye movements, we simulated these two controller signal interaction extremes as inputs to the nonlinear mechanical model: i) neurological summation of the VOR controller signal with the saccadic controller signal, called "CS SUM" or "VOR ON" in the figures, and ii) nonlinear saccadic 'squelching' of the VOR signal at the ocular motoneurons (the final common pathway) during the saccadic pulse, called "CS ALT" or "VOR OFF" in the figures. The other general neurological possibility that must be considered would be for there to be higher level

preprogrammed changes in saccadic magnitude due to prior CNS learning. Only by this isolation via simulation of these neurological possibilities from the role played by the mechanics of the plant can we effectively tackle these questions. By combining present and past experimental results with our simulation studies, we have been able to demonstrate that, even for the eye movement system, the role played by the mechanical plant is significant and that the majority of the experimental saccadic/VOR data can be explained by the saccadic "squelching" of the VOR signal (CS ALT), although there is some neurological interaction, ranging from about Ø to 8Ø%. Furthermore, we obtained quantitative data on the velocity and (newly discovered) position initial condition effects at saccadic initiation. Finally, these results indicate that in general slow movement induced initial conditions will in fact effect any type of fast movement trajectory, whether eye, head or limb.

METHODS

The basic model algorithm, previously used for simulating both the eye globe and recti muscle behavior and the horizontal head-neck rotational system with a set of six nonlinear differential equations, has been described in detail elsewhere ((Lehman and Stark (1979), Zangemeister, Lehman and Stark (1981)). See Winters et al (1983) for a discussion on the adequacy of this general model structure, for the details of the saccadic magnitude range optimization technique and for results indicating that the eye model, for the concerns of this paper, is robust with regard to the details of the model. The basic experimental method has also been described previously ((Nam et al (1981), Winters et al (1983)).

RESULTS

Initial condition effects: experimental. Figure 1a gives some typical head, eye and "gaze" traces. The gaze trace displayed here just represents a summation of the head and eye traces - approximately the true gaze. The resulting gaze trajectory should be proportional to the target change for the ideal "linear summation" case. Thus the gaze trace presented is reasonable but more ideal than is commonly found - it is the slight deviations, that will also be present in the eye trace, that are of special interest here.

Prior observations had indicated that the initial velocity definitely played a significant role, and this observation was the initial reason for this study, and led to the early designation of saccades with the current VOR-induced direction termed "with" saccades and conversely saccades that opposed the current velocity termed "against" saccades (see Figure 2b and c). Simulation and experimental findings (to be described later) also indicated that the initial position also played a role.

The saccadic magnitude for a large number of 10 degree target jumps is plotted in Figure 2. Similar results were obtained for all 2.5, 5 10 and 20 degree targets. These will be summed up in later figures. Some of the spread in the data is simply due to the initial position condition, which is an implicit parameter in this figure, and which will be examined in more detail later. In addition to the experimental and simulation (for CS SUM) results, a line representing the ideal linear summation hypothesis is also displayed.

Initial condition effects: simulation. One can investigate the effects of these various initial conditions more explicitly via simulation. In the left and middle panels of Figure 3 "with" and "against" saccades occur at strategic times during the sinusoidal VOR, i.e. where certain initial conditions should have their peak effect, no effect, etc.. The left panel differs from the middle one in that the former simulates the "CS SUM" neural interaction hypothesis while the latter presents the results for the "CS ALT" interaction assumption. In both cases there are two separate runs superimposed on one graph, in which five saccades of the same magnitude but often different directions and/or initial positions occur at the same five instants in time. Saccade cases "a" and "b" in both panels show explicitly typical "with" and "against" behavior differences. Case "a" is additionally at a region where there is a negligible initial acceleration effect. The lack of a with-against difference change relative to similar situations in which an acceleration existed, along with other results not presented, indicates that the initial acceleration has little effect relative to other initial conditions. This might have been expected since the inertia of the eyeball is very low. In case "b" the saccade occurs at a time when the two controller signal types are approximately the same (i.e. the VOR controller signal is passing through its mean), showing that the observed phenomena cannot be due solely to the controller signal effects. In general, the "with" trace has a larger saccadic magnitude and a higher apparent peak velocity, although not necessarily a higher relative peak velocity.

Cases "c" and "d" isolate the effects of the initial position on the saccadic magnitude and the peak velocity since the other conditions are all the same for the two runs. Notice that saccades going "toward" the primary position (or "with" the elastic springs) are larger than those "away" from the midline. This phenomenon shows that the model is very sensitive to the passive plant elasticity, an observation that is supported by the earlier sensitivity analysis work of Lehman and

Stark (1979). This simulation finding on the effect of the initial position on the saccade resulted in the authors reanalyzing their experimental data to search for - and find - an initial position effect in the experimental data (discussed later).

It is apparent that, for a given frequency, a given initial velocity can occur a two times during one cycle of the waveform. This inherent coupling of the initial conditions makes the job of isolating the relative effects more difficult, and is the basis for the "direction dependence", or "hysteresis", in the "CS SUM" simulation trace of Figure 2 where VOR velocity is an explicit independent variable. A further consideration that may also explain some of this "hysteresis" is that the initial kinematic values are measured relative to the initiation of the saccade as opposed to the expected average value of the velocity during the saccade duration if no saccade was to occur. However, these figures, plus additional simulation results (unpublished observations), show that this is not a main factor. Thus, it is apparent that, in addition to the initial velocity effect, the initial position and, as we shall see next, the controller signal type of interaction, are also major causes of the observed phenomena.

Controller signal effects: simulation. The differences due to these two extreme controller signal interaction strategies are investigated in more detail in the right panel of Figure 3. Here the same input train of saccades are run for the two aforementioned extremes in controller signal strategy. The "normal" bang-bang controller signal for the model, without any VOR or "slow phase" signal, has an "on" of, say, 60 grams for 10 degrees, and an "off" of 2 grams. By definition, this is what the controller signal will be during the pulse for the "CS ALT", or squelching assumption, regardless of what level the VOR signal was at. On the other hand, for controller signal summation hypothesis (CS SUM), there will usually be both an agonist and an antagonist shift, with one in a direction opposite to the other. The resulting CS traces clearly show these differences in both agonist and antagonist muscles. In the antagonist, "CS OFF" can be thought of as a total "turning off", or inhibition, of the antagonist to the 2 grams minimum while "CS SUM" is the maximum of a summation of 2 grams and the VOR signal (typically +/- 8 grams, 18 gram mean) and two grams itself. Thus, the major finding here is that, for an appropriately placed saccade (such as in cases "c" and "e"), the CS signal difference - particularly for "away" saccades, is very significant. The fact that the position, velocity, and acceleration traces are very different for cases "a", "c" and "e" indicates that the model is sensitive to changes in the controller signal. The obvious question is: which of these extremes in controller better approximate the observed experimental behavior?

Interaction results over the wider operating range. Averaged experimental and simulation data for saccades resulting from target jumps of 2.5, 5, 10 and 20 degrees are plotted versus the initial velocities of approximately +/- 40 and 80 deg/sec in Figure 4. In the left trace the vertical axis represents the absolute saccadic magnitude while in Figure 6b the vertical axis gives the peak relative velocity. As mentioned previously, the values on these plots can change with the amplitude and frequency of the periodic VOR base signal, since this affects the initial condition "couplings" between velocity and the (implicit) position effect. The experimental points on these curves were obtained via averaging a large sample, which causes the initial position effect to be effectively "averaged out". The simulation values shown were obtained by the elimination of other initial conditions through runs at strategic locations. The results indicate that the "CS ALT", or "squelching" hypothesis best approximates the observed behavior for "with" saccades, while "against" saccades fall directly between the two extremes. An inspection of the right panel, giving the peak relative velocity, supports the findings of the left panel.

< Figure 7a, b, c >

A further look at the position effect. The position effect first became apparent from the simulation findings and was then consistently found in the experimental data. It is hard to quantify the position role explicitly because it depends on the relative amount of going "toward" the primary position versus going "away" from it - tough to specify as one independant variable. To explicitly show this effect in our experimental data, we decided to code and sort the data as follows: if over twice as much of the saccadic trajectory is going "toward" the primary position as "away" from it, i.e. over 2/3 of the trajectory was "toward", the saccade was called a "toward" position saccade. If it was between 1/3 and 2/3 "toward" (or 2/3 to 1/3 "away"), it was classified as neither (O), and if it was less than 1/3 "toward", i.e. over 2/3 "away", it was called an "away" from position saccade. All the data for a falling within a given classification for a given target jump were then averaged. The net results of this classification scheme are displayed in the right two panels of Figures 5.

As an example, a "with velocity, toward the primary position" would be called a "WI", an "against velocity, away from zero position" was coded as an "AA", and so on. In the middle panel the data is for the 2Ø degree target jump, while in right panel we present the normalized (saccadic magnitude divided by target jump size) results for all sizes for the different classification categories. For the sake of visualization, we have placed the three "against velocity" with different position conditions on the far left, the "away position only" next, the "toward position only" further right, and the "with" velocity, under different position conditions to the far right. We see that "with" velocity saccades going primarily toward the zero position were on the average 2Ø% larger than "with" saccades going primarily away from the center position. Similar findings are apparent for "against" velocity saccades. This effect was particularly significant for larger "with" saccades. These right two panels also show quite nicely that there are initial velocity and position effects in the abscence of the other, and that when the two effects are both involved they both exert some influence, resulting in a saccadic magnitude "compromise".

Main sequence relationship. The main sequence results for relative peak velocity versus saccadic magnitude for the present experimental and simulation runs are compared with the prior main sequence line in Figure 6. As in Figure 4, the target jump sizes are only 2.5, 5, 1Ø and 2Ø degrees. The somewhat continuous range of magnitudes is of course due to the wide variety of saccade sizes for a given target jump, as mentioned in detail priviously. Main sequence data gives an indirect indication of plant dynamics behavior. Of interest here is to see whether the magnitude and peak velocity, which as already discussed are both functions of the initial conditions, are affected in a similar relative scale. In other words, the initial conditions and time are implicit parameters and thus by looking at the local main sequence (that around the specific target jump) we can, among other things, get an indication of whether or not initial conditions affect the plant in a manner similar to neuro-controller signal variation effects due to different intended saccadic magnitudes. By careful inspection one can notice the 4 "groups" of data corresponding to the four target jump sizes. For the 2Ø degree cluster the slope of the experimental data is steeper than that of the main sequence "normal" line, and is in between the two simulation extremes. Notice that for these large saccades sizes "sqelching" the VOR controller signal allows the other initial conditions to move the "phase plane" right along the "normal" curve, while VOR summation causes a higher slope similar to that in lower regions of the curve. For the middle two clusters around 5 and 1Ø degrees the experimental and simulation results lie along the "normal" line. This indicates that the initial conditions tend to affect the model trajectory in a manner similar to controller signals of different saccade

sizes! Furthermore, this observation may provide indirect evidence against having to postulate a higher brain center "preprogramming" of saccade size. For the 2.5 cluster, the experimental data indicate a lower value for the peak velocity - this is probably subject dependent, and the corresponding simulation results indicate that the model bang-bang controller signal fit with amplitude is somewhat too strong for very small saccades.

Preliminary head and limb findings. Models with the same basic sixth order structure have also been used to simulate horizontal head and forearm flexion-extension limb movements. The fundamental difference between the eye model and the head and limb models is that the former is, to a first approximation, a viso-elastic system while the latter two exhibit a very significant inertia. The simplest versions of the models, which we will make use of here, have only the force-velocity nonlinearity also in the eye model. However, a plot of the viscous force versus the velocity across the dashpot representing Hill's equation during a fast movement indicates the the dynamic force-velocity relation is extremely nonlinear, much more so than that for the eye (unpublished). The parameters for the head model are similar to those of Zangemeister et al (1981), while those for the limb, which were found by a search of the surprisingly good literature, will be presented elsewhere. To a first approximation, the head is overdamped and the limb is slightly underdamped. In the left panel of Figure 7, we have plotted the effect of an initial velocity condition of +- 100 deg/sec on 20 degree fast movements, with the controller signal being the same for the three cases. Notice that this is similar to the CS OFF condition in which there is an initial velocity. The fast movements are initiated by an slightly overlapping three-pulse burst (agonist-antagonist-agonist) followed by a step. The tuned pulse heights/widths were basically similar to those found in the literature (Hannaford et al (this meeting), Freund and Budingen (1978), Lestienne (1979), Ghez and Martin (1982), Wadman et al (1979)). In the right panel we see the effect of a periodic "slow" movement on the "fast" movement magnitude, assuming neurological summation (CS SUM). It is apparent that the story is more complex for these systems, partly because the movement duration is very significant relative to the sinusoid period. However, once again we see that there is less than ideal linear summation. Notice how the "fast" movement efects the phase. Finally, from a practical viewpoint, the "against" condition is of special interest, since many limb movements involve a preparatory movement followed by a fast movement in the opposite direction.

DISCUSSION

In this paper we have analyzed in some depth the effect of slow movement induced initial kinematic conditions and controller signal protocol on the trajectory of the resulting fast movement. For the eye, we are able to do this explicitly by using our well tested model of the eye muscle and plant. For the head and some simple limb systems, we presented some preliminary findings, with final results awaiting the further refinement of our rather complex models. For eye movements, we can compare our broad range simulation results on the interaction between different "slow" and "fast" eye movements with both our own experimental data on saccadic/VOR interaction and with others' results for various fast/slow interactions. Our results allow us to reinterpret some of this past work using our present simulation insights.

Effect of initial conditions as a basic finding. One basic finding was that the initial velocity of the system when the saccade occurs does in fact affect the resulting saccadic trajectory, with "with" saccadic magnitudes being larger than normal saccades which in turn are larger than "against" saccades. This is a fundamental observation of most previous investigations of eye movements and was expected. This is true even when the the "slow" controller signal is turned off during the saccadic pulse, and furthermore is a nonlinear effect. We also saw that these basic observations also hold for head and limb movements. Another basic finding is that there is a significant initial/final position effect, i.e. a saccade is larger when moving primarily "toward" the primary position than "away" from the primary position. This effect is less apparent in the head and limb movement systems, where the elastic terms are smaller relative to the viscous and inertial parameters.

The controller signal effect as a basic finding. Finally, for eye movements, we went a step beyond what most previous investigators could do by explicitly separating interactions - linear or otherwise - at the neurological from those at the eye muscle/plant (mechanical) level via simulation. This separation allowed us to check in a rigorous manner the controller signal interaction.

Interpretation of data: saccade / VOR interaction. A comparison of our simulation and experimental results for the CS SUM and CS ALT neurological controller signal extremes showed that, for the saccadic/VOR interaction case, over half of the VOR signal is squelched by the saccade - particularily during the "with" initial velocity movements. Thus, the nonlinear mechanics explain the rest of the behavior, with the overall result being a trajectory that could, in some cases, be mistaken as a neurological summation and a negligible mechanical role. Our dynamic model shows quite clearly that there is a significant mechanical interaction even if the controller signal types were to act alternatively.

Jurgen's group (1977, 1980, 1981) indicated that the saccadic/VOR interaction was about 70% of "linear summation" - similar to our data. Our current insights suggest that, since they were not aware of the significance of the mechanical plant, their hypothesis that there must exist either some high-level predictive mechanism that takes into account what the VOR system is doing, or local feedback, may require modification. We still need to address one of their main arguements in support of their line of reasoning: that the durations of the saccades seem to be different for "with" and "against" saccades, with the former larger. Since it is generally believed (for example Jurgens et al (1975)) that the duration of a saccade is a good indication of its intended magnitude (a notion especially compatible with the "pulse width" modulation strategy assumed here), this finding needs to be considered. Their differences in duration, shown by a linear regression line in a graph with a large scatter, are only a couple of milliseconds - we have found that the duration of

saccades is hard to measure with sufficient accuracy to a millisecond for this type of experimental data - especially for "with" saccades where the VOR-induced "slope" makes it hard to define or quantify an exact "duration". Since superposed "with" and "against" simulation traces indicate that "with" saccades tend to have an "eyeballed" duration of about a millisecond or so longer, we hypothesize tht their slight duration difference findings might be due to mechanical factors and to the frustrations of measuring durations accurately. Thus, postulating a higher level "predictive" capacity or "local feedback" to control saccadic magnitude may not be necessary. Our comments are not intended to mean that the plant mechanics tell the whole story - for instance, their finding of a stronger correlation between saccadic magnitude and the final rather than the initial velocity cannot be mechanically explained.

Interpretation of data: vestibular nystagmus.

A few groups (Chun and Robinson (1978), Jurgens et al (1981)) have examined the interaction between the quick and slow phases of vestibular nystagmus. The basic finding is that the quick phase magnitude does not seem to be a function of the ongoing VOR velocity. However, the quick phase movement (which to put in perspective is always an "against" movement) is not goal directed. Thus there is no way of knowing the intended magnitude. Now consider the main sequence data of "CS ALT" displayed in Figure 5. We found earlier that it is difficult to distinguish between the "intended" magnitude and any mechanical role because both lie approximately along the main sequence! Thus, the present modelling efforts would support the view of the quick phase squelching, or alternating with, the slow phase (CS ALT). This strengthens the previous hypothesis because we have not ignored the mechanics and have shown that the mechanics are compatible with the hypothesis.

Interpretation of data: saccade / smooth pursuit interaction. Our model also appears to shed some light into the other types of interaction between a "fast" eye movement system and other "slow" eye movement systems. Jurgens et al (1975) very elegantly showed that the smooth pursuit system seemed to be turning off just before a saccade occurs and seemed to turn back on after the saccadic pulse. It is as if the saccadic system inhibited (squelched) the smooth pursuit system. However, the resulting saccadic magnitude appeared to be a compromise and to partially take into account the pursuit system. Their data is very similar to our simulation results in which the saccade is squelched (CS ALT) and the observed behavior is just due to the system mechanics. They suggested that the saccadic system predicts what the smooth pursuit system would have done and generates an appropriate magnitude saccade. Our simulation studies show that a predictive "intelligent" saccadic generator may not be needed and that the behavior found is probably mainly due to the mechanics of the system.

Interpretation of the data: saccade / vergence interaction. Finally, Ono et al (1978) and Kenyon et al (1980) have investigated the interaction between vergence and the saccadic system. This system is different from the previous type of interaction in that the vergence "initial velocity" is small - typically under 20 deg/sec. Both groups found that peak velocity and the saccadic magnitude differences were greater that which could be predicted by the "linear summation", "additivity" hypothesis. From the lower middle of the left panel of Figure 4, for 2 degree, one will notice that the the neurological summation condition (CS SUM) is also greater than the "linear summation" line slope - especially for the low velocity region of interest here. This data is blown up in Figure 5a We see that most of this "extra strong" interaction may be explained by a new definition of "summation" as having not only a neurological but also a mechanical aspect. This hypothesis of the mechanics, as well as the neural structures, playing a significant role, first put forward by Kenyon et al (1980), seems to be a much more likely explanation for most of the effect than the "fast" and "slow" muscle fiber type differences postulated by Ono's group, and

furthermore allows Hering's law to remain approximately intact.

Final Remarks. In conclusion, our findings show the elegance of using a model to help understand a biological system. Our model showed explicitly that much of the interaction between the saccadic system and any "slow" eye system can be explained simply by the mechanics associated with the eye. More generally, we see that for any "fast" and "slow" neuromusculoskeletal interaction one cannot ignor the mechanical aspects of the problem. We were able to isolate the neural and mechanical "summation" aspects and show that the result is not the same as ideal "linear summation". Our model results also led us to look for - and find - an eye position effect related to the parallel elasticity. Furthermore, since our basic simulation findings were robust we are confident that the system mechanics do in general play a significant role and we feel that those who investigate movement interaction in the future will have to consider both neural and mechanical interaction to interprete their data.

REFERENCES

A. T. Bahill, M. Clark and L. Stark, "Dynamic Overshooting in Saccadic Eye Movements is Caused by Neurological Control Signal Reversals", Exp Neurol, 48: 197, (1975).

A. T. Bahill and L. Stark, "Trajectories of Saccadic Eye Movements", Sci Amer, 240: 107-122 (1979).

O. Bock, "Non-Linear Interaction of the Vestibular and the Eye Tracking System in Man", Exp Brain Res 47: 461-464, (1982).

H. Bremermann, "A Method of Unconstrained Global Optimization", Math Biosci 9: 1-15, (1970).

K.S. Chun and D. A. Robinson, "A Model of Quick Phase Generation in the Vestibulo-Ocular Reflex", Biol Cyber 28: 209-221, (1978).

M.R. Clark and L. Stark, "Control of Human Eye Movements: Modelling of Extraocular Muscle", Math Biosci 20: 191-211, (1974).

M.R. Clark and L. Stark, "Control of Human Eye Movements: A Model for the Extraocular Plant", Math Biosci 20: 2213-238, (1974).

C.C. Collins, A.B. Scott and D. O'Meara, "Elements of the Peripheral Oculomotor Apparatus", Amer J. Ophthal, 46: 510-515, (1969).

G. Cook and L. Stark, "The Human Eye Movement Mechanism: Experiments, Modelling and Model Testing", Arch Ophthal, 79: 428-436, (1968).

H.J. Freund and H.J. Budingen, "The Relationship Between Speed and Amplitude of the Fastest Voluntary Contractions of Human Arm Muscles", Exp Brain Res, 31: 1-12, (1978).

C. Ghez and J.H. Martin, "The Control of Rapid Limb Movement in the Cat III. Agonist-Antagonist Coupling", Exp Brain Res, 45: 115-125, (1982).

A.V. Hill, "The Heat of Shortening and the Dynamic Constants of Muscle", Proc. Roy. Soc. B, 126, 136-195, (1938).

R. Jurgens, W. Becker and P. Rieger, "The Programming of Fast Eye Movements During Natural Vestibular Stimulation - Two Types of Interaction", IFAC Symp "Bio- and Eco-Systems", Leipzig, 1977, 120-129.

R. Jurgens, W. Becker and P. Rieger, "Different Effects Involved in the Interaction of Saccades and the Vestibulo Ocular Reflex (VOR)", Conference on Oculomotor and Vestibular Physiology, Sarang, 1980.

R. Jurgens, W. Becker, P. Rieger and A. Widderich, "Interaction Between Goal-Directed Saccades and the Vestibulo-Ocular Reflex (VOR) is Different from the Interaction Between Quick Phases and the VOR", "Prog in Oculomotor Res", Ed: Fuchs, A. and Becker, W., Elsevier, North Holland, 1981, 33.

E.L. Keller, "Oculomotor Neuron Behavior", in "Models of Oculomotor Behavior and Control", Ed: Zuber, B.L., CRC Press, Boca Raton, Florida, 1981, pp 1-20.

R.V. Kenyon, K.J. Ciuffreda and L. Stark, "Unequal Saccades During Vergence", Am J Optom and Physio Opt, 57(9): 586-594, (1980).

S. Lehman and L. Stark, "Simulation of Linear and Nonlinear Eye Movement Models: Sensitivity Analysis and Enumeration Studies of Time Optimal Control", Cybern and Info Sci 4: 21-43, (1979).

F. Lestienne, "Effects of Inertial Load and Velocity on the Braking Process of Voluntary Limb Movements", Exp Brain Res, 35: 407-418, (1979).

J.M. Miller, H. Ono and M.J. Steinbach, "Additivity of Fusional Vergence and Pursuit Eye Movements", Vision Res, 30:43-47, (1980).

P. Morasso, E. Bizzi and J. Dichgans, "Adjustment of Saccade Characteristics During Head Movement", Exp Brain Res, 16: 492-500, (1973).

M.N. Nam, J.M. Winters and L. Stark, "Model Simulation Studies to Clarify the Effect on Saccadic Eye Movements of Initial Condition Velocities Set by the Vestibular Ocular Reflex (VOR)", Proc. of Seventeenth Annual Conf. on Manual Control, Los Angeles, 1981, 537-548.

H. Ono, S. Nakamizo and M.J. Steinbach, "Nonadditivity of Vergence and Saccadic Eye Movement", Vision Res, 18: 735-739, (1978).

D.A. Robinson, "Models of the Mechanics of Eye Movements", In: "Models of Oculomotor Behavior and Control", Ed: Zuber, B.L., CRC Press, Boca Raton, Florida, pp 21-42, (1981).

D.A. Robinson, "Linear Addition of Optokinetic and Vestibular Signals in the Vestibular Nucleus", Exp Brain Res, 30: 447-450, (1977).

D.A. Robinson and E.L. Keller, "Behavior of Eye Movement Motoneurons in the Alert Monkey", In: Cerebral Control of Eye Movements and Motion Perception", Ed: Dichgans, J. and Bizzi, E, S. Karger, Basel, 5, (1972).

R. Schmid, A. Buizza and D. Zambarbieri, "A Non-Linear Model for Visual-Vestibular Interaction during Body Rotation in Man", Biol Cybern 36: 143-151, (1980).

M. Takahashi, T. Uemura, and T. Fujishiro, "Studies of the Vestibulo-Ocular Reflex and Visual-Vestibular Interactions During Active Head Movements", Acta Otalaryngol, 90: 115-124, (1980).

W.J. Wadman, D. van der Gon, R.H. Geuze and C.R. Mol, "Control of Fast Goal-Directed Arm Movements", J. Human Movement Studies, 5: 3-17, (1979).

J. M. Winters, M. H. Nam, and L. Stark, "Modelling Dynamical Interactions in Eye Movements: Saccadic Behavior in the Presence of the VOR", Math Biosci, (in press)

W. H. Zangemeister, S. Lehman and L. Stark, "Simulation of Head Movement Trajectories: Model and Fit to Main Sequence", Biol Cybern, 41: 19-32, (1981).

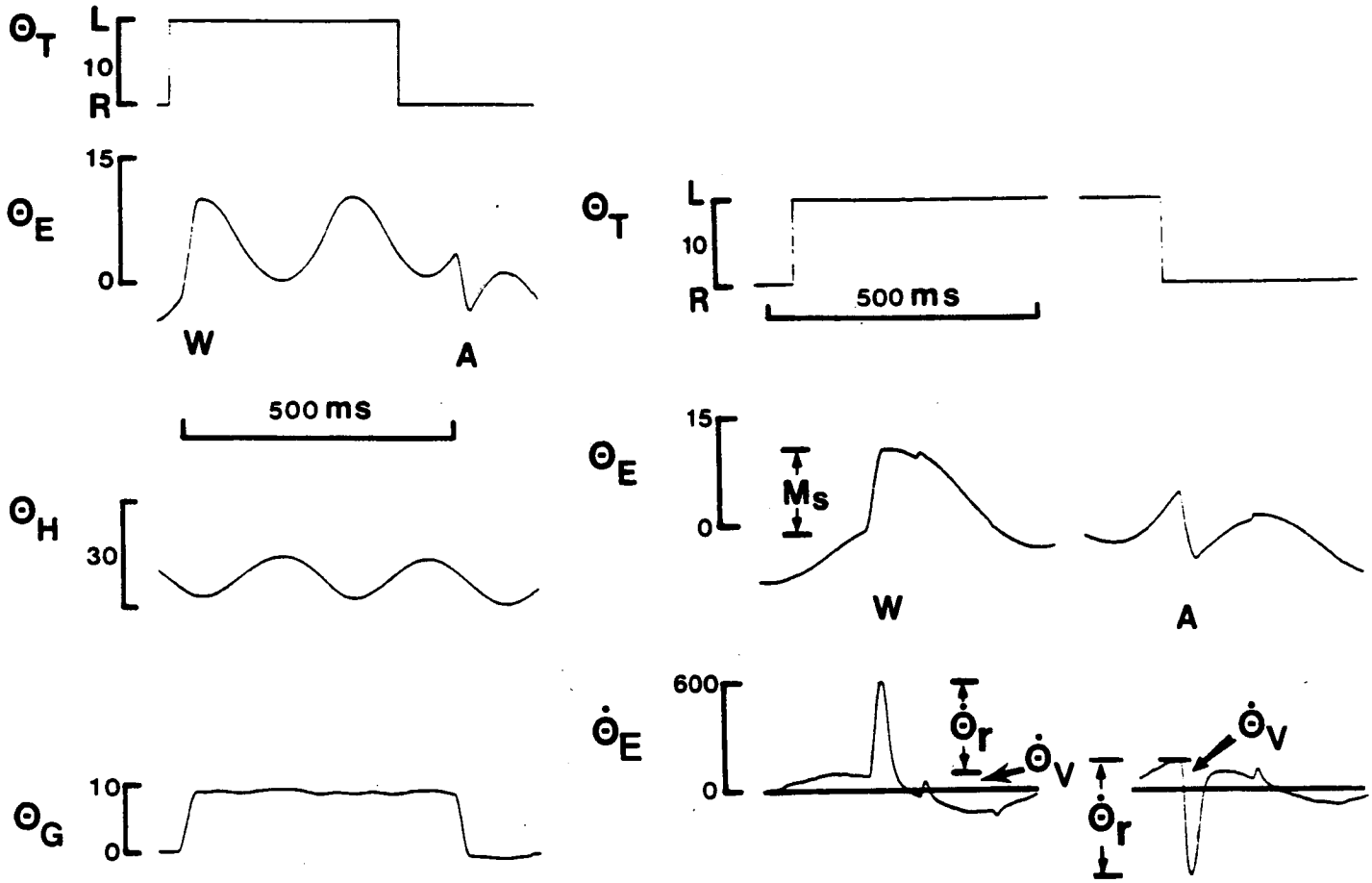


Figure 1: a) Experimental saccades going "with" and "against" VOR. In this typical case the target (top trace) jumps from right to left and back while the head rotates in approximately a periodic waveform (third trace). The eyes, via the VOR, consequently rotate in the opposite direction while simultaneously following the target jumps (second trace). The resulting gaze position, which is a linear summation of the eye and head positions, approximately follows the target (fourth trace). b) and c) Definition of the saccadic magnitude and velocity measurement conventions. $\dot{\theta}_r$ is the relative peak velocity, i.e. the velocity measurement difference between the initial velocity $\dot{\theta}_v$ and the apparant peak velocity relative to zero $\dot{\theta}$.

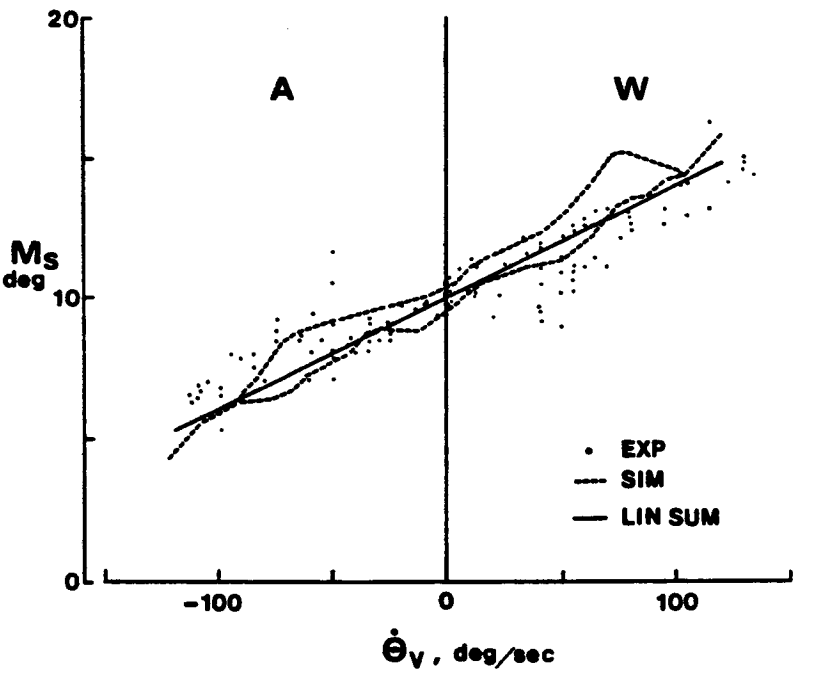


Figure 2: Experimental and simulation saccadic magnitude versus the initial velocity due to the VOR for the 1Ø degree target jumps. The simulation results (dashed) are for the neurological controller signal summation hypothesis (CS SUM) driven saccades occurring at different times during a VOR sine wave of magnitude +/- 15.5 degrees and frequency 1 Hz. Note the direction dependance (hysteresis) in the simulation curve (see text for explanation).

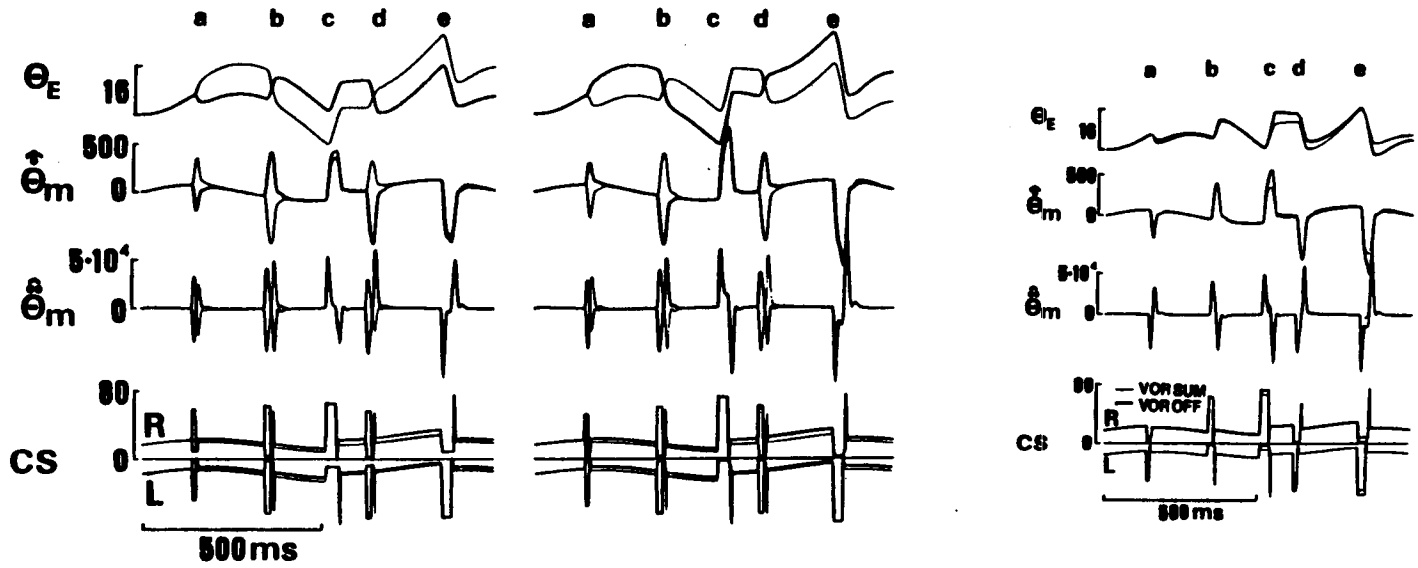


Figure 3: Simulation results for 5 "strategically" placed saccades on top of the VOR induced sinusoidal base signal. In all three panels there are two superimposed runs per trace, with each run having the 5 saccades of the same magnitude occuring at the same times. The only difference is in the saccadic direction or the controller signal hypothesis. Traces, from top to bottom, are the eye position, velocity, acceleration, and the left and right controller signals. Left panel: the "CS SUM", or "VOR ON" summation hypothesis. Middle panel: the "CS ALT", or "VOR OFF" hypothesis. Right panel: the two controller signal types (CS ON, CS ALT). See text for additional details.

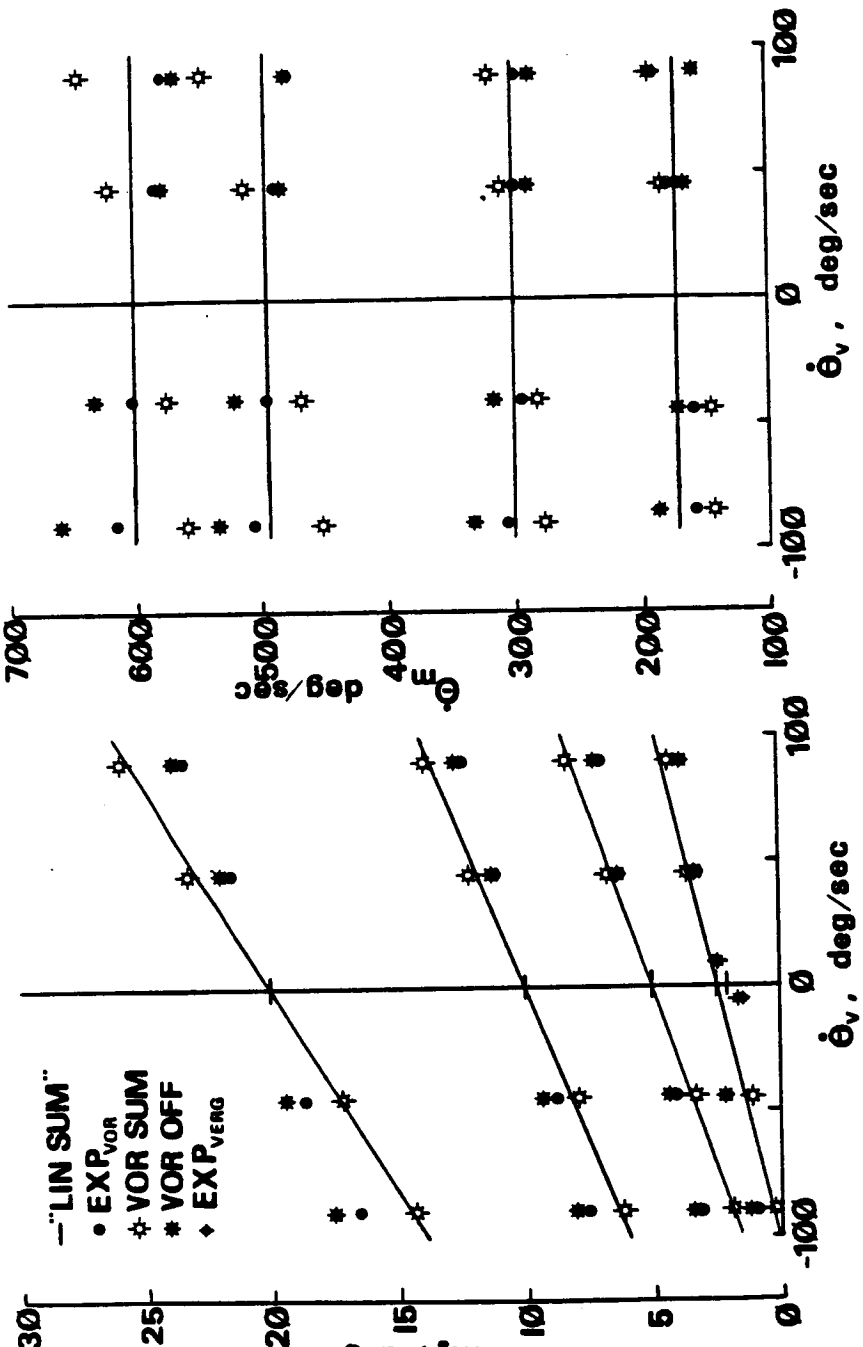


Figure 4: a) Summary of the averaged saccadic magnitude as a function of the initial velocity for experimental and simulation data for 2.5, 5 10 and 20 degree target jumps. Both the "summation" (CS SUM) and "squelching" (CS ALT) hypotheses for the neurological controller signal is presented. Also shown in the lower middle are experimental data for 2 degree vergence and the simulation results when one assumes neurological summation (CS SUM). See text for more details. b) Summary of the averaged saccadic peak relative velocity as a function of the initial VOR velocity for experimental and simulation for 2.5, 5, 10 and 20 degree target jumps, again for both the "CS SUM" and "CS ALT" neurological interaction assumptions.

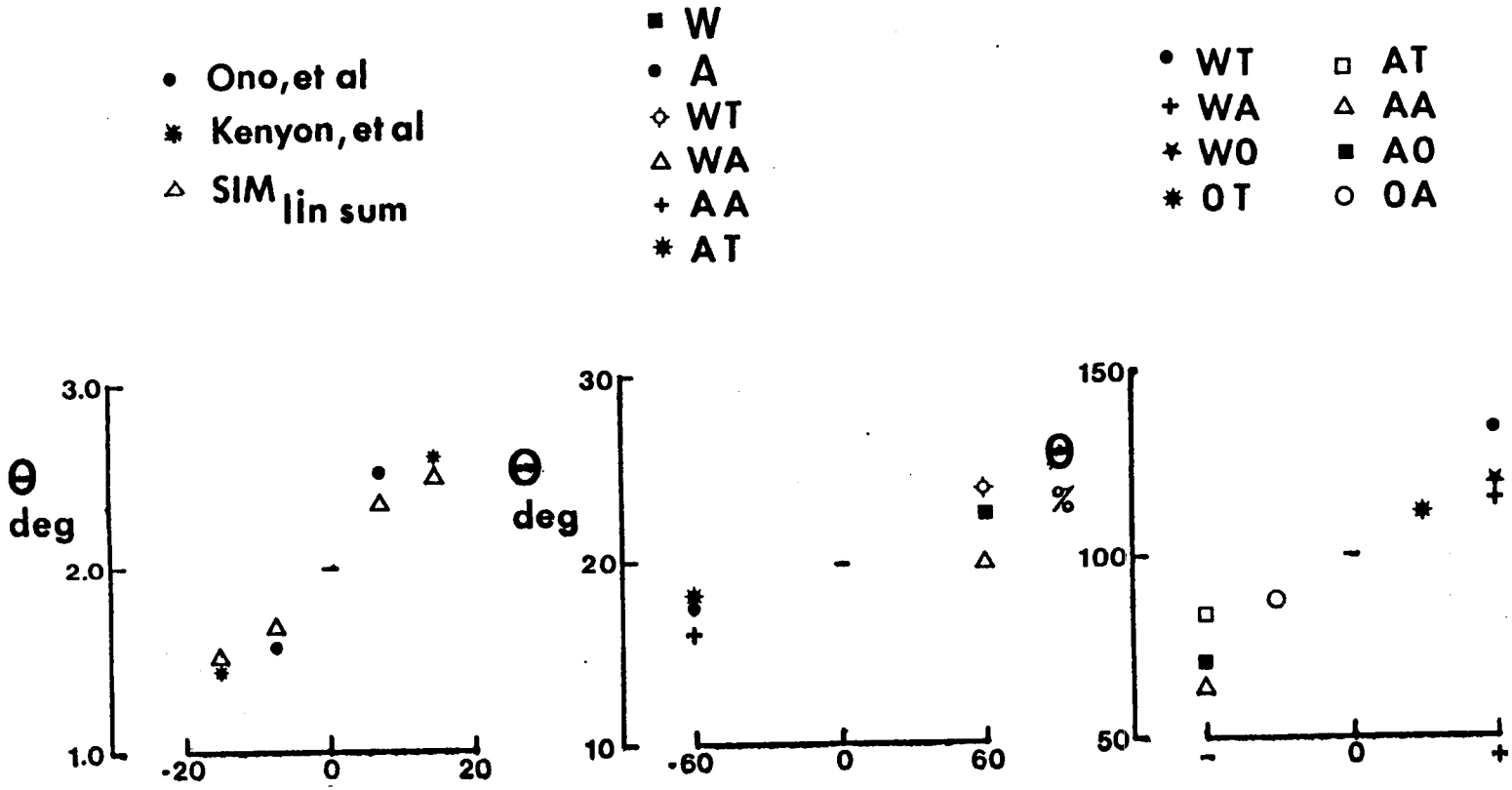


Figure 5: Left Panel: Experimental data for vergence and saccade interaction from Ono et al (1978) and from Kenyon et al (1980) and simulation results for the neurological summation condition (CS SIM) for this low initial velocity, low saccadic magnitude region; Middle panel: The effect of the position and velocity effects on 20 deg target movements (see text for more details and below for an explanation of the lettering); Right Panel: The normalized average effect of velocity and position on saccadic magnitude, given a target jump size. Explanation of lettering: The first letter represents the initial velocity condition, either "with" (W), "against" (A), or neither significantly (0). The second letter represents the initial position condition, either "toward" (T), "away" (A) or neither significantly (0). The outer points (far left, far right) up from the abscissa of the right panel represent the points where there is an "against" (left) or "with" (right) initial velocity condition, the inner two points where the initial position is the only significant parameter. (See text for more details).

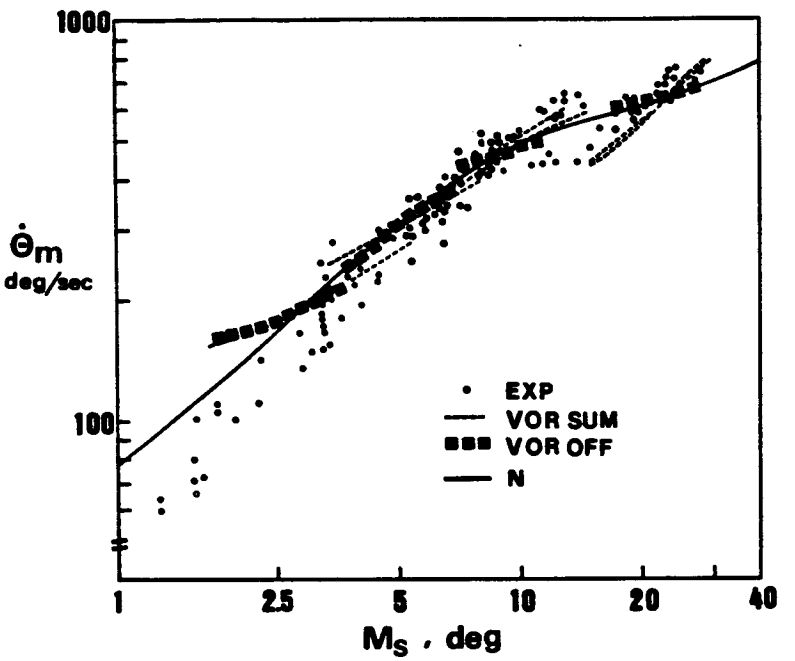


Figure 6: Main sequence for experimental and simulation data. The four clusters of data are for 2.5, 5 10 and 20 degree target jumps. See text for further explanation.

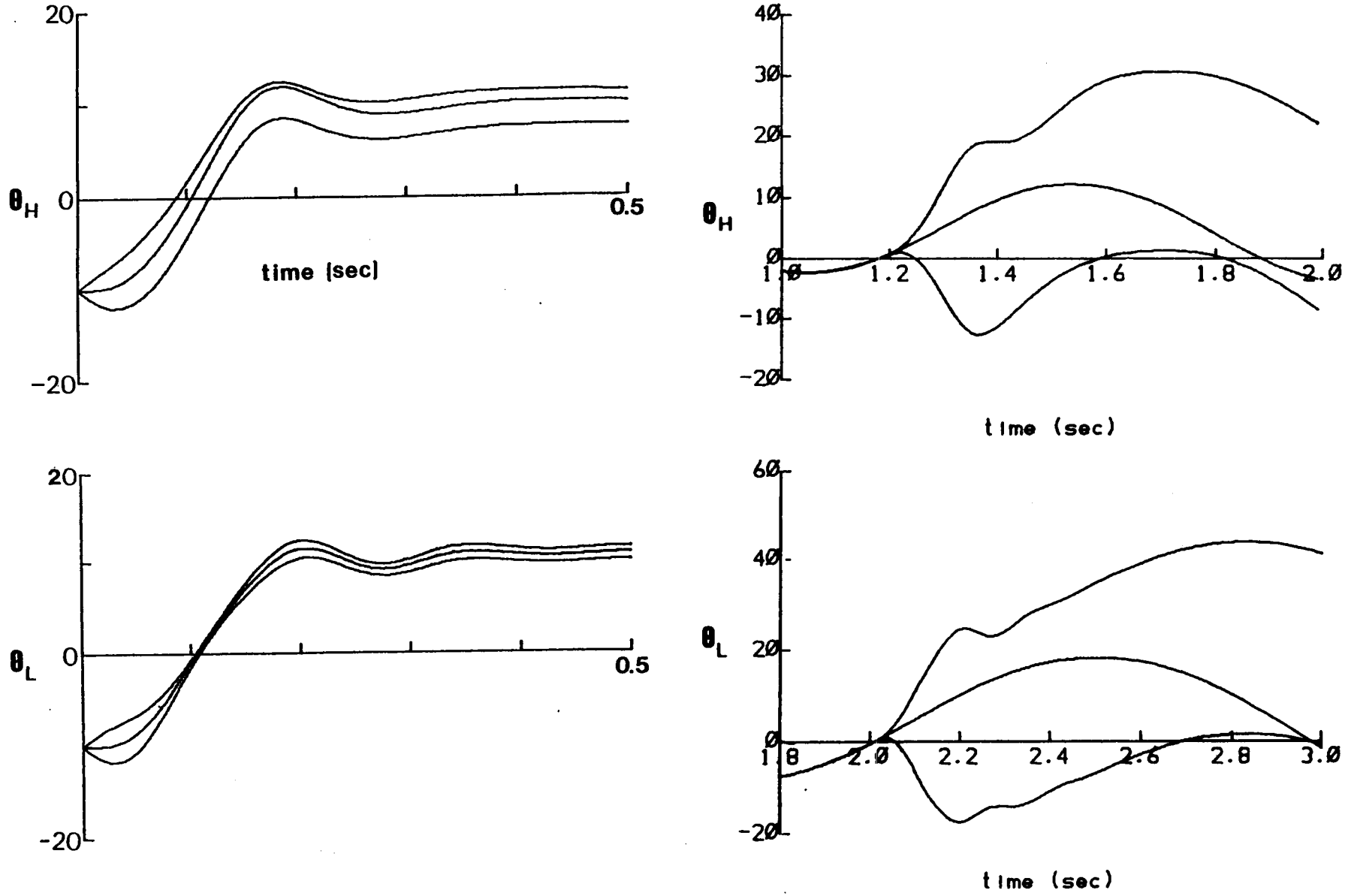


Figure 7: Simulation of head and limb "fast-slow" interaction. Left Panel: Fast 20 degree head and limb movements for initial conditions of 0, -100 and 100 deg/sec. The controller signal is the same for all cases. Right Panel: Fast "with" and "against" 20 degree movements in the presence of a slower periodic type movement. Neurological summation (CS SUM) is assumed.

Visual Near Response is Mediated by Asymmetric Dual-Interaction Control System

Glenn Myers and Lawrence Stark

Department of Physiological Optics and Engineering Sciences, University of California, Berkeley. 94720.

Key Words: Pupil, near response, synkinesis, latency, block diagram, computer analysis.

Running Title: Pupillary Near Response

Acknowledgements: The authors are pleased to acknowledge the assistance of James Wong, Junell Anchetta, and Paul Peng. One of the authors (GM) was supported by NIH Training Grant #GM07379-04. We are pleased to acknowledge partial support from the NCC 2-86 Cooperative Agreement, NASA-Ames Research Center.

ABSTRACT

The near response complex comprises three elements: accommodation, convergence, and pupillary constriction. The synkinesis between vergence and accommodation has been studied since the time of Donders (1864) and Meuller (1826), and is well understood functionally, if not neuroanatomically. To investigate the topology of this complex, the latency of the near response components were investigated in four healthy trained subjects. Both vergence and accommodation respond with shorter latencies to disparity than to blur stimuli; by contrast, pupillary latencies to blur and disparity stimuli are essentially equal. This evidence suggests that the near response complex is mediated by an asymmetric system of neural control. A computer search confirms that this model fits the data better than competing models.

INTRODUCTION

The near response comprises three elements (the near response triad): As an object of regard approaches a viewer, the crystalline lens rounds up (accommodates) to keep the object in focus, the eyes turn in (converge) to keep the object on the fovea, and the pupil constricts to increase the depth of field, helping the accommodative system keep the object in focus. Following Krishnan, Shirachi, and Stark (1977), Figure 1 shows these three systems and the hypothetical synkinetic pathways linking them. The synkinesis between accommodation and vergence is well understood functionally, if not neuroanatomically. How is the pupillary near response mediated? There are three hypothesis in the literature:

1. Vergence-Pupil Synkinesis. Path VP in Figure 1 illustrates a hypothetical link between the vergence motor controller and the pupillary motor controller.

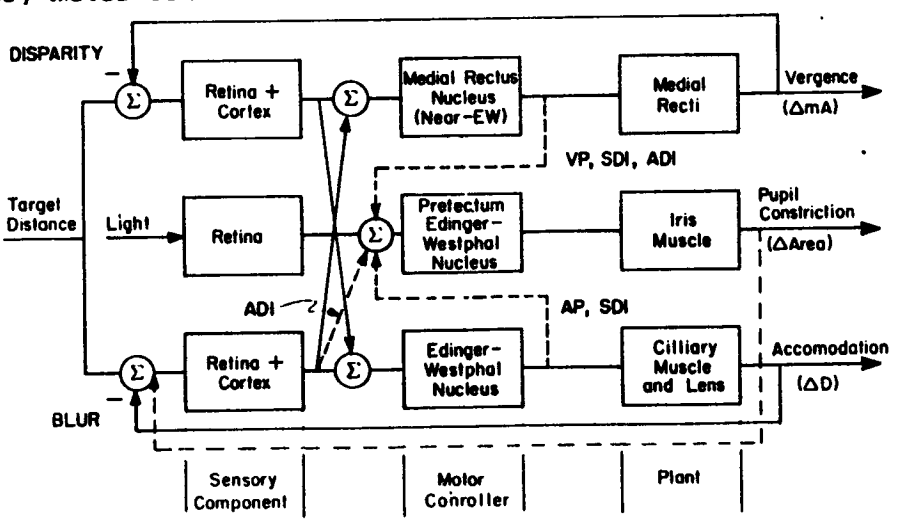


Figure 1. Block diagram representation of visual near response system showing synkinesis between vergence (top path), accommodation (bottom) and pupillary (middle) systems. Lines labeled AP,VP,SDI,ADI correspond to hypothetical synkinetic pathways described in the text. This block diagram representation is not intended to be interpreted neuroanatomically. Rather, it should be regarded as a tool to aid in expressing and testing our hypotheses.

2. Accommodation-Pupil Synkinesis. Path AP in Figure 1 illustrates a hypothetical link between the accommodation motor controller and the pupil system motor controller.

3.Dual Interaction Theory The paths labled SDI in Figure 1 together comprise a symmetric dual interaction mechanism.

To determine how the pupillary near response is mediated, we studied latencies to onset of response of these three systems to blur and near stimuli. Most prior investigators (see discussion) studied the magnitude of the responses of these systems. Since accommodation and vergence are inextricably linked through their synkinetic interactions, response magnitudes are not reliable indicators of this system's topology. Our results lead us to propose a new version of the Dual Interaction Theory, with asymmetric synkinesis among the near response systems, as illustrated by the paths labled ADI in Figure 1.

EXPERIMENTAL METHODS

Four healthy, trained subjects were studied to determine their responses to three kinds of stimuli: Light, blur, and near stimuli (blur plus retinal disparity). Three kinds of responses were recorded: Pupillary responses (constriction to both light and near stimuli), accommodation, and vergence eye movements. The experimental setups are shown schematically in Figure 2. In order to investigate the topology of the systems of interest, stimulus configurations were chosen to minimize all of the latencies measured, within the constraint that prediction not be present.

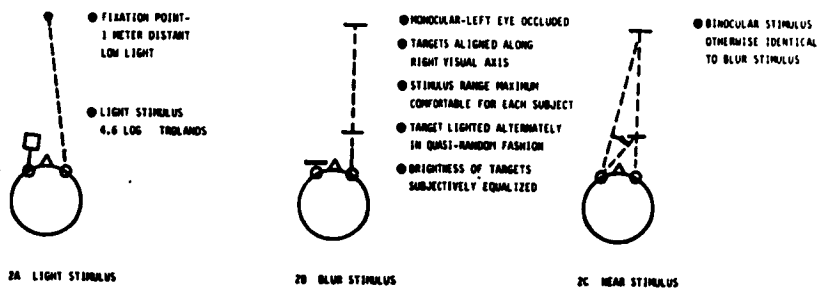


FIGURE 2. EXPERIMENTAL PARADIGMS
2A: LIGHT STIMULUS
2B: BLUR STIMULUS
2C: NEAR STIMULUS

Light-pupil responses: A Maxwellian view stimulus subtending a rectangular solid visual angle of 20 degrees by 20 degrees was presented centrally and the consensual pupillary response monitored by a Video Eye Monitor described elsewhere (Myers et al,1981). A fixation point served to control accommodtion and to minimize eye movements. Light intensity was 4.6 log Trolands, (15,000 milli-foot Lamberts) against a dark background.

Blur-accommodation responses: A Stark-O'Neill optometer (1968) was used to measure accommodtion in response to blur stimuli. Two targets of equal size and subjectively equal brightness were presented alternately along the visual axis of the right eye. The left eye was occluded to eliminate disparity stimuli. The times between changes in stimulus were varied randomly to eliminate prediction. Artifacts due to small movements of the right eye-the Kenyon effect (Kenyon et al, 1978)-were noted in the record, but were easily distinguished from the accommodative responses.

Blur-vergence responses: Using the same stimulus apparatus as above, and infrared limbus eye monitors (described by Stark, Vossius and Young, 1962), vergence eye movements in response to blur stimuli were monitored.

Blur-pupil responses: Using the same stimulus apparatus as above, and the TV pupillometer, changes in pupil size in response to blur stimuli were monitored.

Near stimulus-pupil: The subject viewed binocularly a pair of targets which were presented alternately along the visual axis of the right eye. Pupil size was monitored as above.

Near stimulus-accommodation: Accommodation in response to a near stimulus is much more difficult to measure than blur accommodation since the eye movements involved are greater. The values presented in Table I for disparity accommodation are modified from Krishnan, Shirachi and Stark (1977).

Near stimulus-vergence: Using the stimulus apparatus described above, and the limbus eye trackers, vergence eye movements in response to near stimuli were monitored. Krishnan, Shirachi, and Stark (1977) showed that disparity vergence latency is much less than accommodative vergence latency. Thus near response-vergence latency is essentially equivalent to disparity-vergence latency.

Previous investigators have employed a variety of methods to detect the onset of response. The three most popular methods are: (i) Reading directly from the record, the investigator makes a judgement as to when the response begins. (ii) Using a computer to match a prototype to an averaged actual response. (c.f. Lee 1964). (iii) Setting some threshold (e.g. 10% of the maximum magnitude of response), beyond which the response is considered to have begun. The second method is unsatisfactory for this application because it depends upon a fit to averaged data, taking response latencies to be constant for any given stimulus magnitude. Although Lee found this to be the case for the light-pupil reflex, it is not the case for the pupillary near response. The third method is perhaps more objective than the first, but incorporates some delay due to the response dynamics of the system being measured. Thus it is unsuitable when comparing the responses of different systems (with different dynamics) to different types of stimuli. Thus the method of choice for this study was the first, i.e. the experimenter made a judgement as to when the response began, based upon visual inspection of the

record. Both pupil size, and it's derivitive were recorded. Latency due to the video eye monitor was subtracted. Outlier points were discarded, as were points with excessive noise. It was particularly important to eliminate points which were suspected of representing anticipation (prediction) on the part of the subject.

EXPERIMENTAL RESULTS

Figure 3 shows typical experimental responses. Table I summarises the experimental results. Latency of pupil response to disparity plus blur is appproximately equal to that to blur alone. Latency of eye movement (convergence) to disparity plus blur is 50 msec shorter than that to blur alone ($p < .01$).

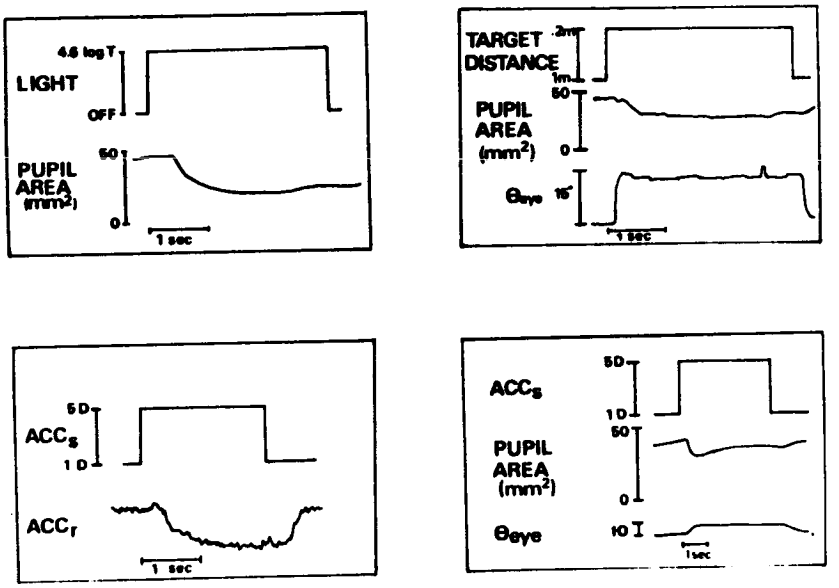


Figure 3. Typical responses to accommodative, near and light stimuli. Upper left: Pupil constriction to light stimuli. Upper Right: Pupillary consriction and eye movement (convergence) to a near (blur plus retinal disparity) stimulus. Lower left: Accommodative response to blur stimulus. Lower right: Pupillary constriction and convergence eye movement to blur stimulus.

TABLE I SUMMARY OF NEAR RESPONSE LATENCIES
Latencies are in msec. ACC=accommodation
VG=vergence eye movements

SUBJECT	LIGHT	MONOCULAR (BLUR)		BINOCULAR (BLUR+DISP)		BLUR-ACCOMMODATION
		PUPIL	VG	PUPIL	VG	
NE-F20	210	323	228	305	163	328
GM-M31	259	321	188	306	163	406
PP-M20	220	320	303	305	203	336
JA-M23	292	360	192	354	184	461
AVERAGE	245	331	228	317	178	383

ANALYTICAL METHODS

Addition of disparity stimulus to blur (yielding a near stimulus) reduces the latency of vergence eye movements by an amount which is significantly greater than the corresponding reduction in pupillary latency. Consider Figure 4. Component latencies in this block diagram were apportioned according to topological constraints (i.e. the component latencies must sum approximately to the experimentally determined system latencies) and physiological constraints (e.g. vergence is mediated by the striated extraocular muscles, which are known to be much faster than the smooth ciliary and iris muscles which mediate, respectively, accommodation and pupil movements). If the pupil system were driven by the vergence system (path VP-Figure 4), or by the accomodation system (path AP), one would expect these reductions to be the same. Clearly, they are not.

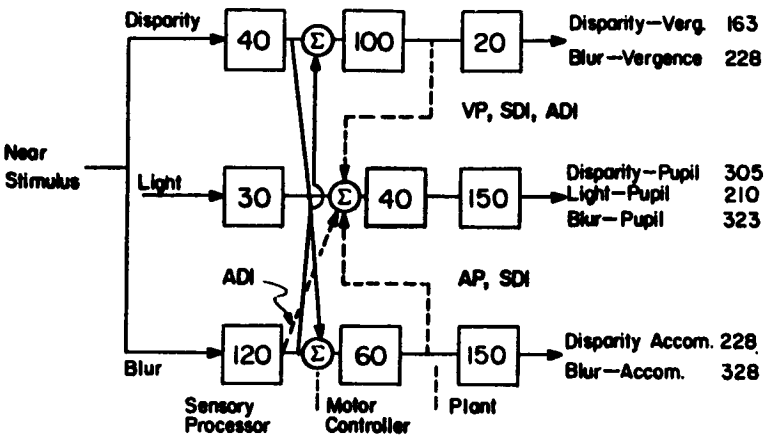


Figure 4. Detailed latency distribution for one subject (NE). Average response latencies in milliseconds given at the right. Numbers in the blocks represent the "component" latencies in milliseconds. Component latency assignments are based on topological and physiological considerations (see text).

Consider the Dual Interaction hypothesis with symmetric accommodation-pupil, vergence-pupil synkinesis (Figure 4, paths SDI). Again, the vergence-pupil latency should be less than the accommodation-pupil latency by about 50 msec, and it is not.

In order to resolve this inconsistency, the topology of the accommodation vergence synkinesis must be modified so that the accommodation-pupil path exits the accommodation system before the disparity-accommodation synkinesis path enters the accommodation system. This may be achieved in two ways: The disparity-accommodation path may join the accommodation system at the level of

the plant, or the accommodation-pupil signal may leave the accommodation system at the level of the accommodation sensor. Either or both of these changes may be realised. In any case, it is necessary that there be direct vergence-pupil and accommodation-pupil pathways to mediate the corresponding responses. One can rule out the vergence sensor to accommodation-plant path on physiological grounds, since a motor controller is required in order to form the controller signal.

To confirm and quantitiate this finding, we conducted a computerised search of all possible combinations of latency components. The range of latency components was constrained to those physiologically reasonable, as discussed above. The increment size was 10 msec. A goodness of fit metric (total square error) was calculated for each possible combination of latencies. The total system latency was calculated for each of the seven stimulus-response subsystems. The difference between the calculated and measured values were squared, and the sum of the seven squared differences formed. This number is a measure of goodness-of-fit of the topology used in the calculation.

ANALYTICAL RESULTS

Table II summarises these values for each subject, for each hypothetical topology. A low value means that the topology fits the data well.

TABLE II SUMMARY OF COMPUTED GOODNESS OF FIT

Subject (Age-sex)	Hypothesis			
	VP	AP	ADI	SDI
NE F-19	2100	1900	395	1800
GM M-31	3300	2500	530	2210
PP M-21	2800	2500	75	2860
JA M-23	14000	3300	2280	3000

Key: VP Vergence-Pupil (Hypothesis 1)
AP Accommodation-Pupil (Hypothesis 2)
SDS Dual Interaction-Symmetric (Hypothesis 3)
ADA Dual Interaction-Asymmetric

Clearly, the dual interaction theory with asymmetric synkinesis (paths ADI, Figure 1) provides the best fit to the data.

DISCUSSION

Latency of vergence eye movements in response to near stimuli (disparity plus blur), is less than that to blur alone by an average of 50 msec for the four subjects. This difference is significant both physiologically and statistically ($p < .01$) for each of the four subjects. Latency of accommodative responses to near stimuli are less than those to blur alone by about 100 msec (Krishnan et al,

1977-we have used these data in modified form, in Figure 4). Latency of pupillary responses to near stimuli are only slightly less (14 msec) than those associated with response to blur alone. This difference is statistically significant in two of the four subjects ($p < .01$), but is probably not physiologically significant. This difference may be due to the Ogle Effect (Backer and Ogle-1964), to the Semmlow Effect (1981)-enhanced accommodative vergence due to binocular accommodative stimulus, as distinct from vergence, or it may be non-artifactual, possibly representing a "near response center".

Numerous investigators have studied the relative magnitudes of pupillary changes in response to changes in accommodation and vergence. Burian and Schubert (1937), Renard (1947), and Schaknovich (1977) found that pupillary diameter follows vergence more closely than accommodation. Magnitude studies suffer from several weaknesses in addressing questions of system topology, viz: (i) The synkinetic interactions between accommodation and vergence require that these two systems move together, so that it is difficult to distinguish between cause and effect. (ii) These methods often implicitly assume linear summation of neurological signals. This condition is rarely met. (iii) Even if the system is rendered open loop (c.f. Stark, 1962) with respect to input parameters, the system may have internally set resting levels. (iv) Most magnitude studies have employed monocular accommodative stimuli, and (necessarily) binocular disparity stimuli. Semmlow and Heergma (1979) have demonstrated the existence of a convergence component in response to binocular accommodative stimuli, in excess of monocular accommodation-vergence. (v) System nonlinearities (e.g.the "dead zone" in accommodation), may render magnitude comparisons invalid. (vi) Judge and Miles (1981) have shown that the Accommodative Convergence to Accommodation (ACA) ratio and the Convergence Accommodation to Convergence (CAC) ratio can be modified by optically altering the pupillary distance.

Jampel (1960) stimulated the cortices of macaques, and observed accommodation, vergence, and pupil constriction separately, and in every combination except accommodation and pupil constriction, without vergence.

Marg and Morgan (1949, 1950) found that pupil response associated with accommodation is greater than that associated with a comensurate vergence response. In this form, these data support the Accommodation-Pupil Hypothesis. This work has been critisised (Stark, 1981), since the accommodation and vergence responses were rendered in diopters and prism diopters respectively. If these same results are rendered in commensurate units (dioters and meter angles), the pupillary response to vergence stimuli are approximately equal to those to accommodation stimuli. In this form, the data support the "Dual Interaction Theory". Marg and Morgan also found that about half of their subjects showed no pupillary response to vergence stimuli alone. This observation strongly supports the Accommodation-Pupil Hypothesis.

Animals with lateral eyes, and lacking binocular vision, are known to have accommodative and pupillary near responses without vergence

(c.f. Walls, 1942). These animals are derived from our common ancestors and their visual systems are, in this respect more primitive than our own. If our binocular visual system is derived from a system like theirs, one could argue that the vergence system was a later development than the pupillary near response, and hence would have been added onto the accommodation-pupil near response system. This argument supports the Accommodation-Pupil Hypothesis.

Although the near response triad involves the cortex, much of it is subserved in the oculomotor nucleus of the midbrain. Furthermore, the pupillary light reflex pathway does not directly involve the cortex, and remains intact in spite of total cortical ablation. The accommodation-vergence synkinesis, which is a basic control function, is probably realized there. The putative vergence-pupil link may be realized in the III-C nucleus of Buttner-Ennever(1981). Keller (1972,73,74), Hodges (1942), and Glicksman (1980), have investigated the neurophysiology of the oculomotor complex.

Krishnan and his colleages (1973a,73b,77) studied latencies and prediction in the vergence and accommodative systems, but did not investigate the pupillary component of the near response. Phillips (1974) and Phillips, Krishnan and Stark (1975) investigated the transfer function of all of the near response components. The transfer function incorporates information as to the phase and magnitude of the response components. Phase is related, but in a nontrivial way, to latency. The stimuli used in Phillips' studies were predictable sinusoids. Krishnan (1973a) found that the vergence system has prediction, so that phase information is not directly related to latency. Nevertheless, it is instructive to note that Phillips found that the blur-pupil response had less phase lag than did the disparity-pupil response. Disparity vergence phase lag was less than that of accommodative vergence, and the phase lag of vergence accommodation was less than blur accommodation. These findings tend to support the Accommodation-Pupil Hypothesis.

Summary.

Addition of disparity stimulus to blur (yielding a near stimulus) reduces the latency of vergence eye movements and of accommodation by an amount which is significanly greater than the corresponding reduction in pupillary latency. Viewed in the context of the topological paradigm of Figure 1, these data imply that the pupillary component of the visual near response is mediated by an asymmetric dual interaction between the accommodation and vergence control systems.

REFERENCES

Alpern M., W.M. Kincaid, and M.J. Lubeck. (1959) Vergence and Accommodation. III. Proposed Definitions of AC/A Ratios. Am. J. Ophthalmol. 48(1):141.

Burian H.M. and G. Schubert. (1937) Das Wesen der Naheeinstellungsreaktion der Pupillen, v. Graefes. Arch. Ophthalmol. 136:377.

Buttner-Ennever J.A. and K. Akert. (1981) Medial Rectus Subgroups of the Oculomotor Nucleus and Their Abducens Internuclear Input in the Monkey. J. Comparative Neurology 197:17-27.

Donders F.C. (1864) On the Anomolies of Accommodation and Refraction of the Eye. The New Sydenham Society.

Glicksman M.A. (1980) Localization of Motoneurons Controlling the Extraocullar Muscles of the Rat. Brain Res. 188:53-62.

Hodes R. and H.W. Magoun. (1942) Autonomic Responses to Electrical Stimulation of the Forebrain and Midbrain with Special Reference to the Pupil. J. Comp. Neurol. 76:169.

Jampel R.S. (1946) Representation of the Near Response on the Cerebral Cortex of the Macaque. Am. J. Optom. 25:480.

Judge S.J. and F.A. Miles. (1981) Gain Changes in Accommadative Vergence Induced by Alteration of the Effective Interocular Separation. In: A.F. Fuchs and W. Becker (Eds.) Proc. Symposium on Neural Control of Eye Movements Elsevier, N.Y.

Kenyon R.V., K.J. Ciuffreda, and L. Stark (1978) Binocular Eye Movements During Accommodative Vergence.Vision Res. 18:546-555..

Keller E.L. (1973) Accommodative Vergence in the Alert Monkey: Motor Unit Analysis. Vision Res. 13:1565.

Keller, E.L. (1974) Participation of Medial Pontine Reticular Formation in Eye Movement Generation in Monkeys. J. Neurophysiol. 37:316.

Keller E.L., and D.A. Robinson. (1972) Abducens Unit Behavior in the Monkey During Vergence Movement. Vision Res. 12:369.

Krishnan V.V., F.Farazian, and L. Stark. (1973) An Analysis of Latencies and Prediction in the Fusional Vergence System. Am. J. of Optometry 50:933-939.

Krishnan V.V., D. Shirachi, and L. Stark. (1977) Dynamic Measures of Vergence Accommodation. Am. J. Optom. and Physiological Optics. 54(7):470-473.

Krishnan V.V., S.Phillips, and L. Stark. (1973) Frequency Analysis of Accommodation, Accommodative Vergence and Disparity Vergence. Vision Res. 13:1545-1553.

Krishnan V.V. and L. Stark. (1975) Integral Control in Accommodation. Computer Programs in Biomedicine 4:237-245.

Maddox E. (1886) Investigations on the Relationship between Convergence and Accommodation of the Eyes. J. Anat. 20:475-505, 565-584.

Marg E. and M.W. Morgan Jr. (1959) The Pupillary Fusion Reflex. Arch. Ophthal. 43:871-878.

Marg E. and M.W. Morgan Jr. (1949) The Pupillary Near Reflex: The Relation of Pupillary Diameter to Accommodation and the Various Elements of Convergence. Am. J. Optom. & Arch. Am. Acad. Optom. 26:183-198.

Marg E. and M.W. Morgan Jr. (1950) Further Investigation of the Pupillary Near Reflex: The Effect of Accommodation, Fusional Convergence and the Proximity Factor on Pupillary Diameter. Am. J. Optom. &. Arch. Am. Acad. Optom. 27(5): 217-225.

Myers G.A. et al. (1981) Pupillometery, a Bioengineering Overview. In: Proc. Seventeenth Annual Conf. on Manual Control. JPL-Caltech. Pub:81-95.

Mueller J. (1826) Vergleishende Physiologie des Gesichtssinnes.

O'Neil W.D. and L. Stark. (1968) Triple Function Ocular Monitor. J.O.S.A. 58:570-573.

Phillips S.R.(1974) Ocular Neurological Control Systems: Accommodation and the Near Response Triad. PhD Thesis, University of California, Berkeley.

Phillips S., V.V. Krishnan, and L. Stark. (1975) Frequency Characteristics of Eight Oculomotor Systems. Proc. IEEE Conference on Cybernetics and Society, San Fransisco.

Renard G. (1947) La synergie pupillaire à la convergence. Rev. Otoneurcophtalmol. 19:240.

Samoilov A. Y., O.N. Sokolova, and A.R. Shakhnovich. (1961) The Pupillographic Method in the Study of the Act of Convergence. Biofizika 6(1):84.

Schubert G. and H.M. Burian. (1937) Die Fusionsreaktion, eine bisher unbekannte Reaktion der Pupille. Pflügers Arch. 238:182.

Semmlow J. (1981) The Oculomotor Near Response. In: Models of Oculomotor Behavior B. Zuber Ed. Chemical Rubber Pub. Co. Bocca Raton, Fla.

Semmlow J. and R. Jaeger. (1972) Modelling the Visual Motor Triad: An Example of a Multiple Input-Output Biocontrol System. Proceedings of 25th ACEME, Bar Harbor, Fl., p.23.

Semmlow J. and L. Stark. (1973) Pupil Movements to Light and Accommodative Stimulation: A Comparative Study. Vision Res. 13:1087-1100.

Shakhnovich A.R. (1977)The Brain and Regulation of Eye Movement. Plenum, New York.

Stark L. (1962) Environmental Clamping of Biological Systems: Pupil Servomechanism. J.O.S.A. 52: 925-930.

Pupil Size Effect in the Pupillary Control System

Fuchuan Sun*, Susan Chan, William Krenz, Jiarui Lin+, Glenn Myers
Lawrence Stark, Pamela Tauchi, Joseph Terdiman

Neurology Unit
University of California, Berkeley

ABSTRACT

The pupil reflex to light has been considered the paradigm for the application of linear control theory. We now explore in depth the "pupil size effect", an important nonlinearity that determines the symmetrical behavior of the pupil in the small pupil size operating regime, and the asymmetrical pupillary escape behavior in the large pupil size operating regime. Experiments are explained in terms of nonlinear systems models that fit the data well. Finally, some general principles of nonlinear control theory are suggested.

Can we imagine utilization of these concepts from natural control systems in modern engineering design?

* on leave from Shanghai Institute of Physiology, Shanghai, China

+ on leave from Huazhong College of Engineering, Wuhan, China

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

INTRODUCTION

The pupillary reflex is an excellent paradigm of a biological control system, and thus the pupil's nonlinearities have been the subject of numerous studies. Early efforts to model the pupil dealt with linear control (Stark and Sherman, 1959); later work (Stark, 1959; Troelstra, 1960; Clynes, 1961; Baker, 1963; Varju, 1969; Webster, 1971) considered such nonlinear phenomena as stimulus saturation, pupillary escape, asymmetry, and binocular rectification. Semmlow and Stark (1973) and Semmlow and Chen (1977) developed models utilizing saturation of the iris to account for pupillary capture behavior; Shimizu and Stark (1977) proposed a model whereby the time constants of the response changed with the stimulus intensity. Each of these heuristic models fitted the experimental data for a small class of inputs and initial conditions.

Recently, a nonlinear "pupil size effect" has been demonstrated to control the dynamics of pupillary responses to step increases of light intensity (Sun and Stark, 1983). At small pupil sizes, a DC or tonic pupillary constriction occurs, called pupillary capture by Usui (1974); at large pupil sizes the pupil constricts and then redilates. This AC or phasic response, first demonstrated by Stark (1962), and named pupillary escape by Lowenstein (1964), has been related to certain binocular clinical tests. Careful experiments have shown that whether the pupil

size is set by light, or by accommodation, or even by light to the contralateral retina (Sun, Tauchi, and Stark, 1983), the pupil size effect is still evident; this demonstrates that the state of retinal excitation or adaptation is only of importance in so far as it sets pupil size. Further experiments with drugs (Sun, Kato, and Stark, 1983) have put the anatomical locus of the source in the Edinger-Westphal nucleus or the pretectal nucleus in the brain stem, since changing the pupil size with neuromuscular drugs does not produce the "pupil size effect". Even in the frequency domain, a pupil size effect was demonstrated (Sun, Myers, and Stark, 1983); when sinusoidal stimulation was employed, the small pupil region displayed low pass filter characteristics, whereas the large pupil region displayed high pass filter characteristics. Careful heuristic modelling studies (Sun, Krenz, and Stark, 1983) have simulated these effects by employing an internal parameter control nonlinearity to set the gains of a phasic AC and a tonic DC path. In this way the dynamics are controlled by the pupil size, or better, an <u>internal activity level</u> that determines pupil size.

Pupil size also determines the degree of symmetry or asymmetry in pupillary responses to positive and negative steps of light, (the latter being a decrease in light level as intensity is a positive quantity.) In addition, pupillary responses resulting from binocular stimulation also show the influence of the pupil size effect (Sun, Tauchi, and Stark, 1983).

In this paper, the general responses of pupillary escape and capture as well as the symmetry/asymmetry phenomenon are defined by experimental, modelling, and simulation studies, utilizing both monocular and binocular stimulus conditions.

METHODS

The dynamic pupillary response was measured by an infra-red TV pupillometer (Semmlow and Stark, 1973; Myers et. al., 1981).

Step, sinusoidal, and pulse stimulus signals of varying amplitudes and duration were generated by a computer or a waveform generator. These signals drove a pulse width modulator to control a glow modulator tube (Sylvania R1131C). Stimulus light was controlled not only by the modulator, but also by insertion of Wratten neutral-density filters to obtain a large range of light intensities. In our sinusoidal experiment the average light level was varied over a very large domain to obtain a range of average pupil sizes and the modulation coefficient was maintained at a constant value of m = Ø.6.

In the monocular case the light stimulus was presented in Maxwellian view in order to ensure open loop conditions (Stark, 1959). A fixation point was provided with an especially dim target (1×10^{-4} ft-lamberts) to control eye movements. When the pupil size was controlled by light level, the target was placed at a 1 meter distance from the subject's eye (the accommodation

stimulus was then 1 diopter). When pupil size was controlled by accommodation level, the target was placed at 25, 20, or 17 cm. (4, 5, or 6 diopters, respectively) to set different initial pupil sizes.

For the binocular experiments, two glow modulator tubes were employed; one tube was placed directly before the left eye, while light from the other tube was reflected onto the right eye by a cold mirror. As the cold mirror reflects visible light but is transparent to infra-red light, it enabled binocular stimulation without interference with the function of the pupillometer. Two signals from a computer wave generator supplied the two glow modulator tubes, producing alternating stimuli which could be independently regulated in intensity and in frequency.

Experimental data were gathered from five normal healthy subjects without refractive errors. All showed similar responses in each experimental run. Fatigue was checked by repeating the first stimulus at the end of every experimental run. Subjects were dark adapted for 10 to 15 minutes before the low light level experiments were begun.

The data was processed using an LSI-11/23 microcomputer with graphics. All of the simulations were also run on the LSI-11/23, which had floating point hardware and graphics display. The model was reformulated into a state-space representation and Runge-Kutta integration with multiple step input functions was used.

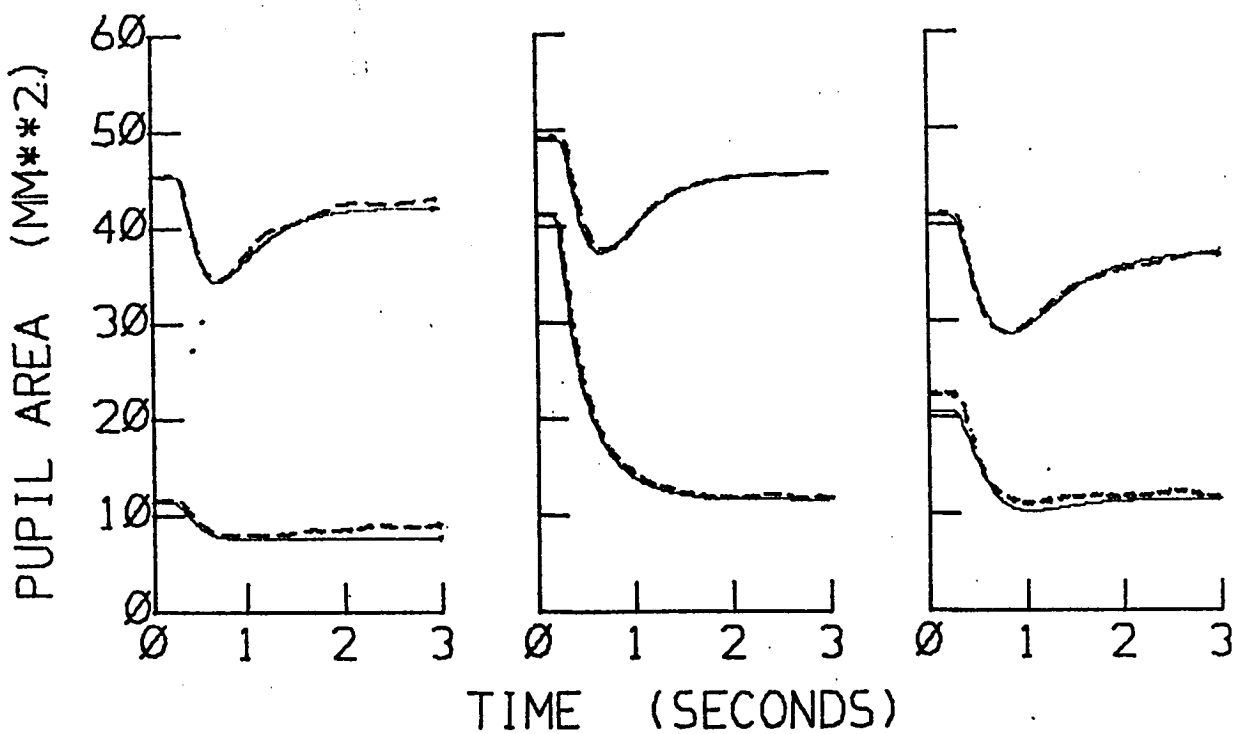


Fig.1 -- Pupil Size Controlled by Light Level or Accommodation.
Left Panel. Pupillary escape (upper trace) occurred with large pupillary sizes. Pupillary capture (lower trace) occurred with small pupillary sizes. The curves are averages of ten responses.
Center Panel. Pupillary escape occurred with a moderate stimulus (upper trace). Pupillary capture occurred with a large stimulus (lower trace). The curves are averages of ten responses. Note that DC level changed because average light with brighter inputs is greater.
Right Panel. Pupillary size controlled by accommodation level through synkinetic triadic near response mechanisms. Pupillary escape occurred with large pupils (upper trace). Pupillary capture occurred with small pupils (lower trace).
Solid lines denote model simulations. Dashed lines denote experimental data.

RESULTS

Pupil size controlled by light level

The main result (Figure 1, left panel) was that pupillary escape occurred with large pupillary sizes between 37 and 50 mm^2 (6.9 to 8.0 mm in diameter), and pupillary capture occurred with small pupillary sizes of 12 to 19 mm^2 (4 to 5 mm in diameter). This was a consistent finding with all subjects and under a variety of operating conditions. The responses shown were averages of ten single responses. The stimulus was a small one, a doubling of light intensity, so that the pupil operating level would not be significantly disturbed.

With this result in hand, we turned to the older experimental results (Lowenstein and Loewenfeld, 1969) that confounded (did not distinguish) operating level and size of pupillary response (Figure 1, center panel). Pupillary escape occurred as above with a moderate stimulus (to 0.035 ft-lamberts from a dark adaptation level of 1×10^{-4} ft-lamberts) and pupillary capture occurred with a large stimulus (to 3500 ft-lamberts from the dark adaptation level of 1×10^{-4} ft-lamberts). However, it can be seen (Figure 1, center panel, lower trace) that the pupil response was driven to the small pupil size operating level (12 to 19 mm^2 in area, or 4 to 5 mm in diameter). This result led previous investigators (Shimizu and Stark, 1977) to conclude incorrectly that pupillary escape was a small signal

phenomenon and pupillary capture a large signal phenomenon. Again the curves are averages of ten responses.

In these experiments pupil size was controlled by operating light level ranging from 1×10^{-4} to Ø.Ø35 ft-lamberts for large pupil sizes and to 35ØØ ft-lamberts to produce small pupil sizes; of course, retinal operating level was changed and retinal nonlinearities might be thought to play a role. The next experiment ruled out this conjecture.

Pupil size controlled by accommodation level

Accommodation level also controls pupillary size through well known synkinetic, triadic near response mechanisms, but via unknown neural pathways. After adjusting the dim fixation point from 1 to 6 diopters (1 m to 17 cm distance), pupillary size changed from 42 mm^2 (6.9 mm in diameter) to 27 mm^2 (2.5 mm in diameter). Now with retinal operating level in the same state but with pupillary size different, an experiment was performed using a moderate step of light (to Ø.Ø35 from 1×10^{-4} ft-lamberts). Pupillary escape occurred with large pupils (Figure 1, left panel, upper trace) and pupillary capture occurred with small pupils (Figure 1, left panel, lower trace).

Double step experiment

Saturation to large signal inputs is a common nonlinearity in biological experiments and historically was used to explain pupillary capture by Semmlow and Stark (1973) and Semmlow and

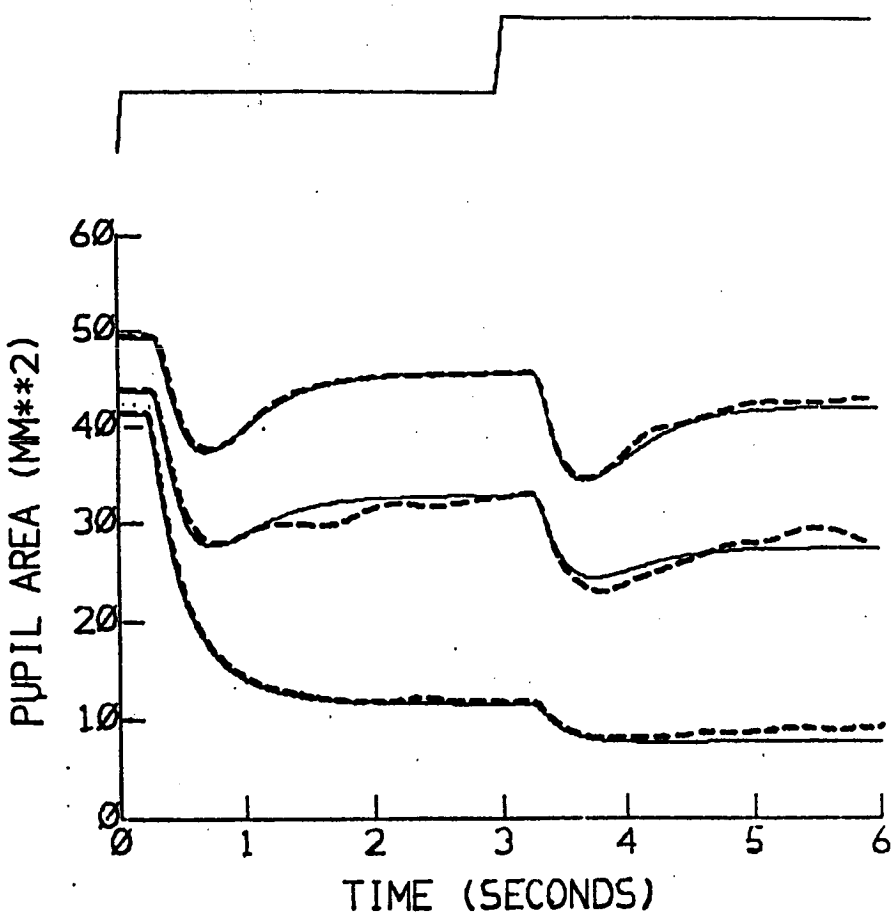


Fig.2 -- Double Step Experiment Upper graph shows double step light input. Lower graph shows responses to small inputs (upper trace), medium inputs (middle trace), and large inputs (lower trace). Note absence of hard saturation even in lower trace with pupillary capture. Solid lines denote model simulations; dashed lines denote experimental data.

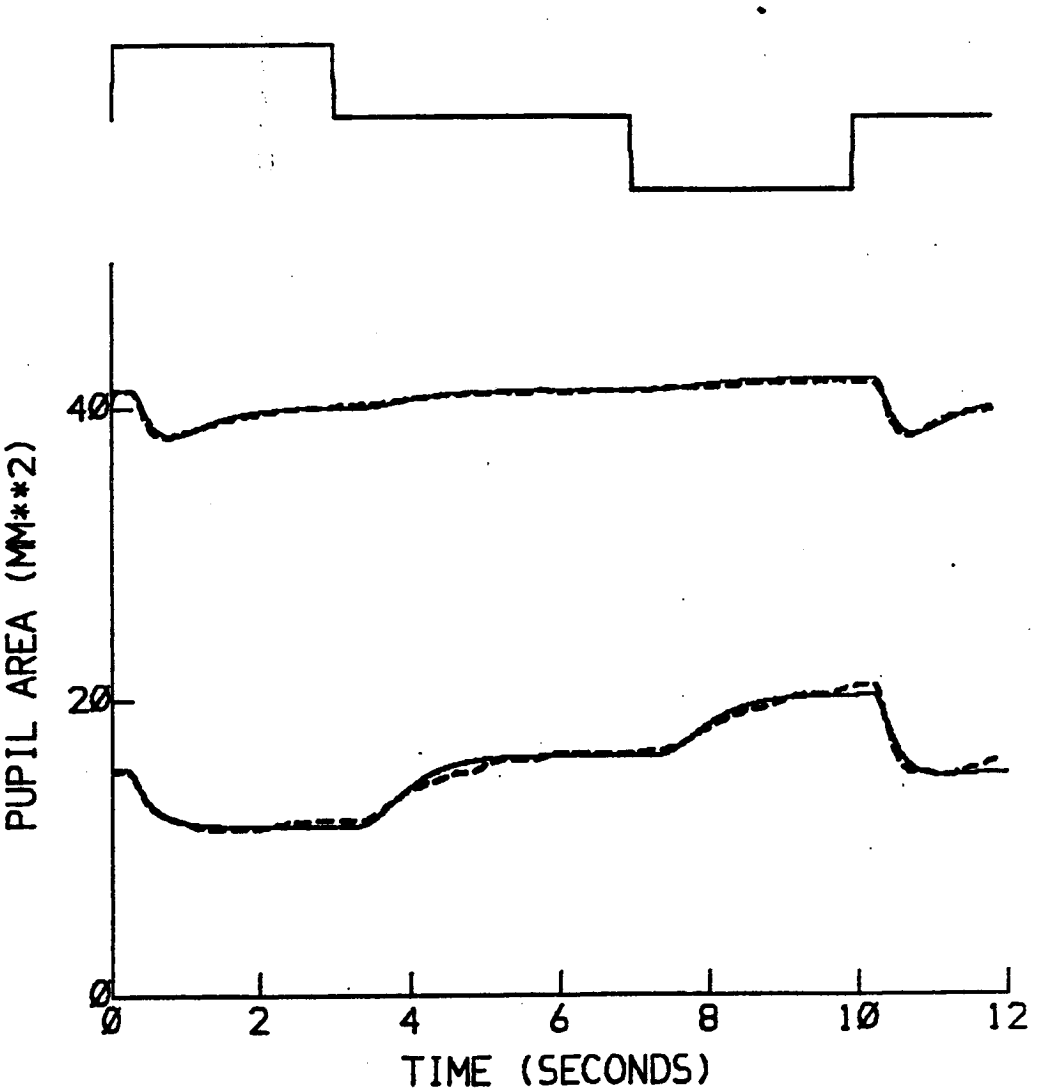


Fig.3 -- Symmetry/Asymmetry Phenomenon by Light Level Upper graph shows bidirectional light input. Lower graph shows symmetrical response of a small pupil (lower trace), and asymmetrical response of an enlarged pupil (upper trace). Pupil size was changed by differing neutral density filters in front of the stimulated eye. Solid lines denote model simulations; dashed lines denote experimental data.

Chen (1977) (see Figure 1, center panel, lower trace). A double step experiment (Figure 2) was performed to obtain evidence directly to this point. The large amplitude response drove the pupil into pupillary capture again (Figure 2, lowest trace). A second step, going from 3500 to 7000 ft-lamberts produced a second response, and demonstrated the absence of a hard saturation. The similarity of this second response (Figure 2, lowest trace) to the pupillary capture response at small pupil size operating level (Figure 1, left and right panel, lower traces) argued against an important role for soft saturation.

Symmetry/asymmetry phenomenon demonstrated by light level

When the pupil was small, 16 mm^2 in area (4.5 mm in diameter), positive going or negative going steps of light produced tonic and almost symmetrical constriction and dilatation responses of the pupil (Figure 3, lower trace, dashed line). The pupil was then made to dilate by placing a -5 log neutral density filter before the stimulated eye. This greatly reduced the light intensity; the pupil dilated over a period of 15 minutes devoted to the retinal dark adaptation and readjustment of pupil size to a stationary level. With the pupil enlarged, 42 mm^2 in area (7.3 mm in diameter), the same relative stimulus then produced a highly asymmetrical response (Figure 3, middle trace, dashed line). This response demonstrated phasic or AC pupillary escape to positive light steps, and only very small tonic dilatation to

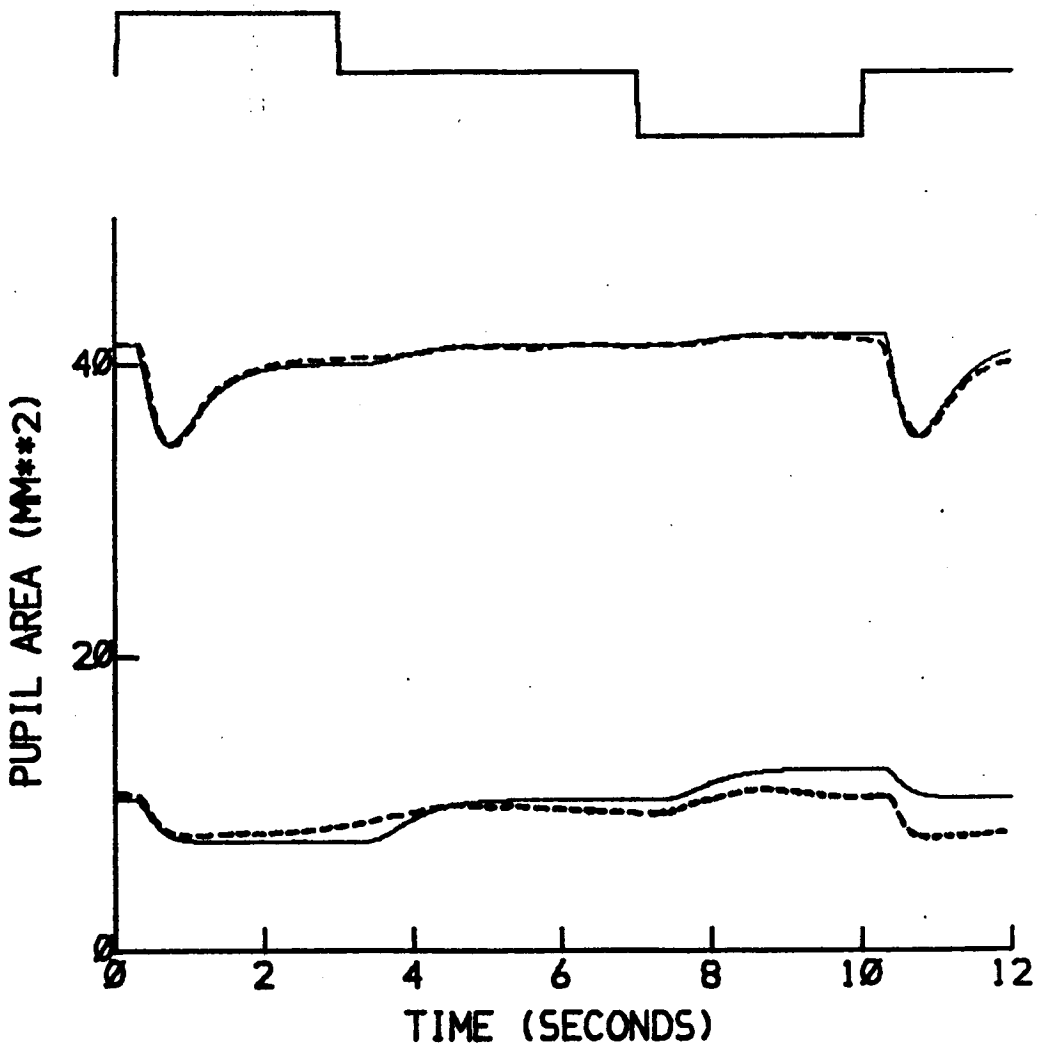


Fig.4 -- Symmetry/Asymmetry Phenomenon by Accommodation Pupil size changed by changing accommodation level. Upper graph shows bi-directional light stimulus. Lower graph shows symmetrical response of a small pupil with accommodative focus (lower trace), compared with the asymmetrical response of a normal relaxed pupil (upper trace). Experimental data were displayed using dashed lines; model simulations were displayed using solid lines.

negative going light steps.

Symmetry/asymmetry phenomenon demonstrated with accommodation

Questions again arose as to the role of retinal state (retinal excitation level or retinal adaptation level) in such an experiment. Pupil size was then changed by changing accommodation level. (When the eye focused at a near target, the pupil constricted; when the eye focused at a far target, the pupil dilated.) With the pupil small, 12 mm^2 in area (3.9 mm in diameter), symmetrical responses were again observed in another subject (Figure 4, lower trace, dashed line). The subject then changed accommodative focus from the near target to a far target; this change produced a synkinetic enlargement of pupil size to 42 mm^2 in area (7.3 mm in diameter). Again the dynamical pupillary responses demonstrated highly asymmetrical constrictions and pupillary escape to positive steps of light and small tonic dilatations to negative steps of light. In this experiment, the "open loop" light stimuli (Figure 4, upper trace) were exactly the same and therefore retinal state remained the same. The only change was in pupil size as produced by accommodation level.

Binocular pupillary response

Equal alternating binocular stimulation of low intensity revealed pupillary escape as well (Figure 5, upper trace, dashed line). Although the total amount of light presented to both eyes was kept constant, two responses were obtained per cycle, within

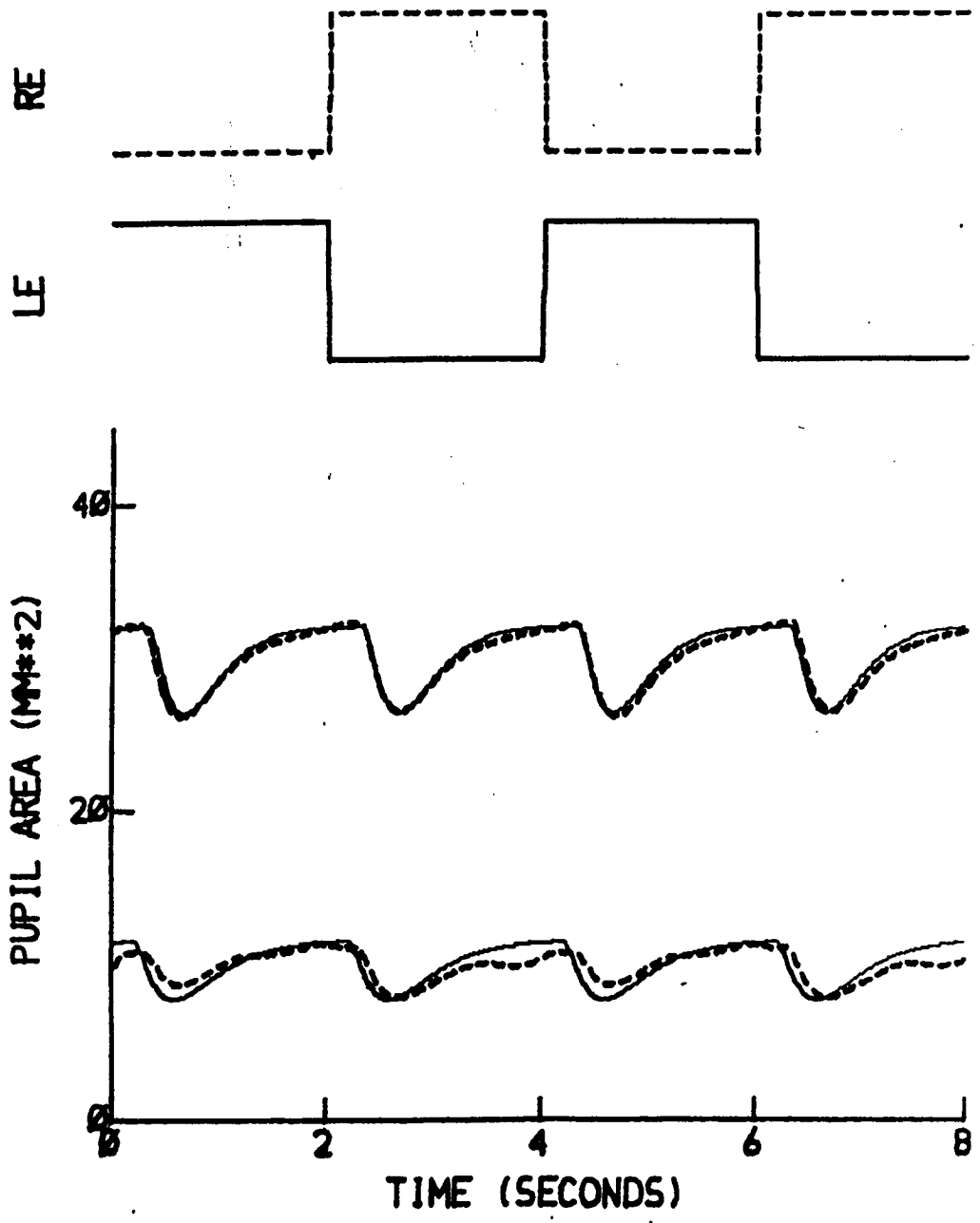


Fig.5 -- Binocular Pupillary Response Upper graph shows equal alternating binocular stimulus. Lower graph shows pupillary escape with low intensity stimulation (upper trace), and with high intensity stimulation (lower trace). Two responses were noted per cycle. Experimental data are displayed in dashed lines; model simulations are displayed in solid lines. LE denotes light stimulus to left eye; RE denotes light stimulus to right eye.

the range of frequencies tested. Similarly, alternating binocular stimulation with high intensity light resulted in phasic contrictions (Figure 5, lower trace, dashed line), that occurred at the beginning of each pulse throughout stimulation, despite a constant total light amount presented to the eyes.

MODELLING

Monocular model structure

The monocular model (Sun, Krenz, and Stark, 1983) is straightforward (Figure 6). There are two parallel pathways carrying light information: an AC or phasic with a lead-lag operator, a rectifier, and a gain; and a DC or tonic path with a lag element, a gain, and a summer to receive a tonic synkinetic control from accommodation. Both paths receive an input, logarithmic intensity of light, and after summing drive the iris neuromuscular plant with its negative gain, time delay, and multiple lag. Finally, a nonlinear internal parameter control (Figure 6, heavy lines) modulates tonic gain as a parameter of the tonic DC level. Parameter exploration both heuristically and with a global optimizer (Bremermann, 1970; Lehman and Stark, 1982; Sun, Krenz, and Stark, 1983) suggested that the DC gain, K_{dc} varies from a high value at small pupil sizes to a low value at large pupil sizes, while K_{ac} varied from low to high.

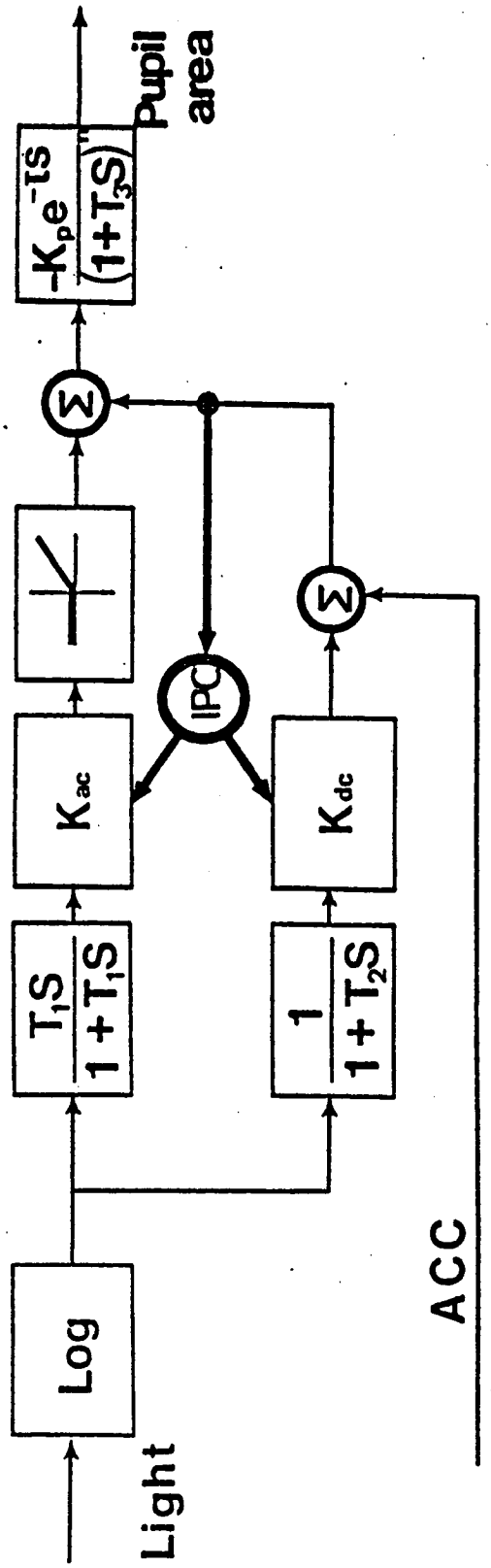


Fig.6 -- Monocular Model. Constructed by Sun, Krenz, and Stark (1982), it features internal parameter control (heavy lines) and parallel tonic and phasic pathways.

Binocular model structure

The binocular model (Sun, Tauchi, and Stark, 1983) was developed from the monocular model (Figure 7). The internal parameter control is regulated by the summed signal of the DC pathways of both eyes. AC summation occurs after seperate rectification of the individual phasic signals of the two eyes.

Model simulations

Model simulations were able to fit the data well (Figures 1-5, solid lines denote model data, dashed lines denote experimental data). Both heuristic methods and global optimization methods show close simulation of the experimental data.

DISCUSSION

It is interesting to consider what principles of control can be learned from the pupil. Can a highly asymmetrical system function as a control system? Is it stable? Can it have an operating level?

The asymmetrical pupil behavior is limited in five ways:

i) its operating regime is only at large pupil sizes

ii) it responds only over a limited bandwidth of relatively high frequencies

iii) it has a limited gain, less than unity, and so does not lead to instability

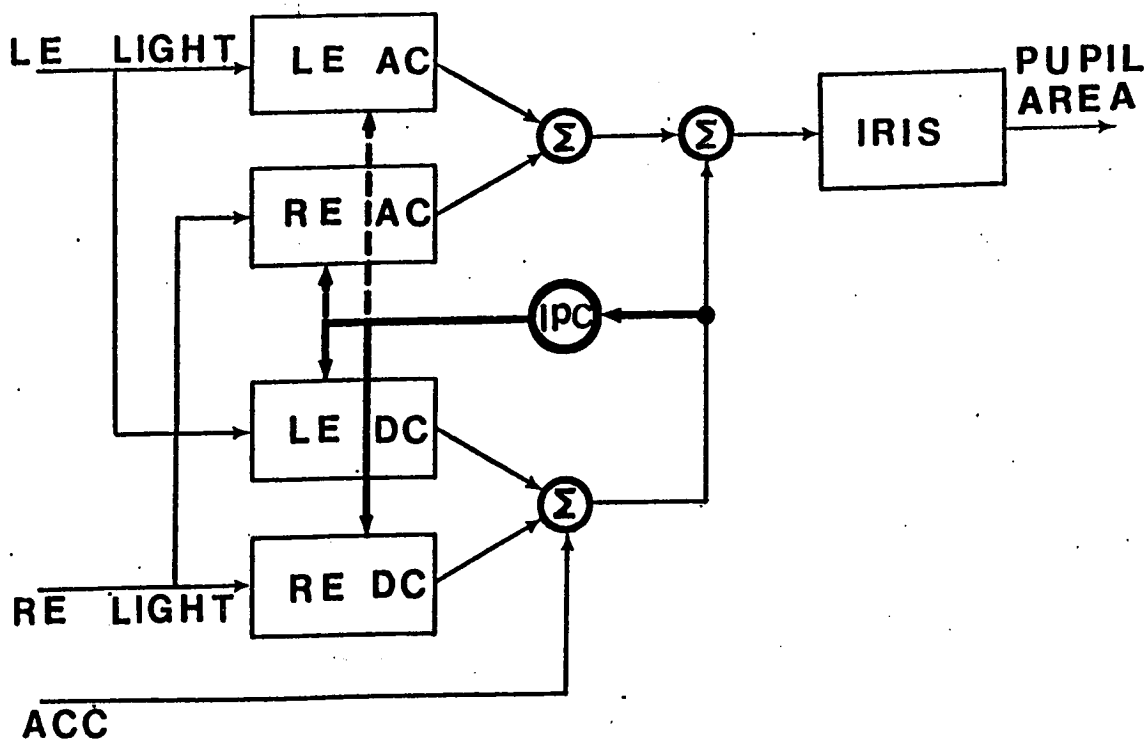


Fig.7 -- Binocular Model. Binocular tonic signals summate to control model gains before summation of tonic and phasic pathways.

iv) there is a concomitant symmetrical DC pathway, although with only a small gain in the operating regime of the AC pathway

v) even when driven by frequency matched stimulus such as high frequency sinusoids or pulse trains, it is only driven into the small-pupil-sized operating regime of the DC pathway

The symmetrical pupil behavior is basic to, and generally dominates the function of the pupil. For example, with bright lights the pupil is small and symmetric, and has a high incremental gain; it thus acts as a fairly efficient regulator of retinal flux. With near targets, the pupil is again small, and thus maintains a large depth of focus for precise viewing.

REFERENCES

1. Baker, F.: Pupillary Response to Double Pulse Stimulation; a Study of Nonlinearity in the Human Pupil System. J. Opt. Soc. Amer. Vol. 53, pp. 1430-1436, 1963.

2. Bremermann, H.: A method of unconstrained global optimization. Math. Biosc. Vol. 9, pp. 1-15, 1970.

3. Clynes, M.: Unidirectional Rate Sensitivity: A biocybernetic law of reflex and humoral systems as physiologic channels of control and communication. Ann. N. Y. Acad. Sci. Vol. 92, pp. 946-969, 1968.

4. Lehman, S., and Stark, L.: Three algorithms for interpreting models consisting of ordinary differential equations: sensitivy coefficients, sensitivity functions, global optimization. Math. Biosci. Vol. 62, pp. 107-122, 1982.

5. Lowenstein, O., Kawabata, H., and Loewenfeld, I.: The Pupil as Indicator of Retinal Activity. Amer. J. Ophthal. Vol. 57, pp. 569-596, 1964.

6. Lowenstein, O., and Loewenfeld, I.: Effect of Various Light Stimuli. The Eye Edited by Hugh Davson, Academic Press, p. 274, 1969.

7. Myers, G., Anchetta, J., Hannaford, B., Peng, P., Sherman, S., Stark, L., Sun, F., and Usui, S.: Pupillometry: A Bioengineering Overview. Proc. 17th Ann. Conf. on Manual

Control, UCLA pp. 523-536, June 16-18, 1981.

8. Semmlow, J., and Stark, L.: Pupil movement to light and accommodative stimulation: a comparative study. Vision Res. Vol. 13, pp. 1087-1100, 1973.

9. Semmlow, J., and Chen, D.: A Simulation Model of the Human Pupil Light Reflex. Math. Biosc. Vol. 33 No. 1/2, 1977.

10. Shimizu, Y., and Stark, L.: Pupillary escape. 10th Symp. on the Pupil. New York, June 1977.

11. Stark, L.: Stability, Oscillations, and Noise in the Human Pupil Servomechanism. Proc. I.R.E. Vol. 47, pp. 1925-1939, 1959.

12. Stark, L.: Biological Rhythms, Noise, and Asymmetry in the Pupil-Retinal System. Ann. N.Y. Acad. Sci. Vol. 98, pp. 1096-1108, 1962.

13. Sun, F., and Stark, L.: Pupillary Escape Intensified by Large Pupillary Size. Vision Res. Vol. 23, No. 6, 1983.

14. Sun, F., Krenz, W., and Stark, L.: A Systems Model for the Pupil Size Effect: I Transient Data. Biological Cybernetics, in press, 1983.

15. Sun, F., Tauchi, P., and Stark, L.: Dynamic Pupillary Response Controlled by the Pupil Size Effect. Experimental Neurology, accepted, 1983.

16. Sun, F., Myers, G., and Stark, L.: High pass character-

istics of the pupil controlled by operating level: pupillary escape intensified by large pupillary size. submitted to IEEE Trans. - BME, 1983.

17. Sun, F., Kato, S., and Stark, L.: Drug Effects on the Dynamic Pupillary Response. in preparation for submitting, 1983.

18. Sun, F, Tauchi, P., and Stark, L.: Binocular Pupillary Response and Modelling. in preparation for submitting, 1983.

18. Troelstra, A.: Detection of Time-varying Light Signals as Measured by the Pupillary Response. J. Opt. Soc. Amer. Vol. 58, pp. 685-690, 1968.

19. Usui, S.: Functional Organization of the Human Pupillary Light Reflex System. Ph.D. thesis, U.C. Berkeley, p. 86, 1974.

20. Varju, D.: Human Pupil Dynamics. Proc. Int. School of Physics. Course XLIII, pp. 442-464, 1969.

21. Webster, J.: Pupillary Light Reflex: Development of Teaching Models. IEEE Trans. BME-18, No. 3, pp. 187-194, 1971.

A Homeomorphic Model of the Pupil Light Reflex System

William Krenz[1], Fuchuan Sun[2], and Lawrence Stark

Neurology Unit

University of California, Berkeley

Abstract

The human pupil has been widely studied as an example of nonlinear biological systems. Many 'Black Box' models have been proposed to describe the nonlinear phenomena known as 'pupillary escape', 'pupillary capture', the response asymmetry, and, more recently, the 'pupil size effect'. A new model is given here which is simpler and more homeomorphic than previous models, while retaining the accurate description of the nonlinear responses of the pupil. A signal controlling steady-state pupil size is used to control the recruitment of neurons from an inverse "Henneman-coded" pool. Multiple configurations of this pool are explored via computer simulations of the model and comparisons with experimental data.

---

1. This project was supported in part by the NCC 2-86 Cooperative Agreement with NASA-AMES Research center

2. On leave from Shanghai Institute of Physiology, Shanghai, China

Introduction

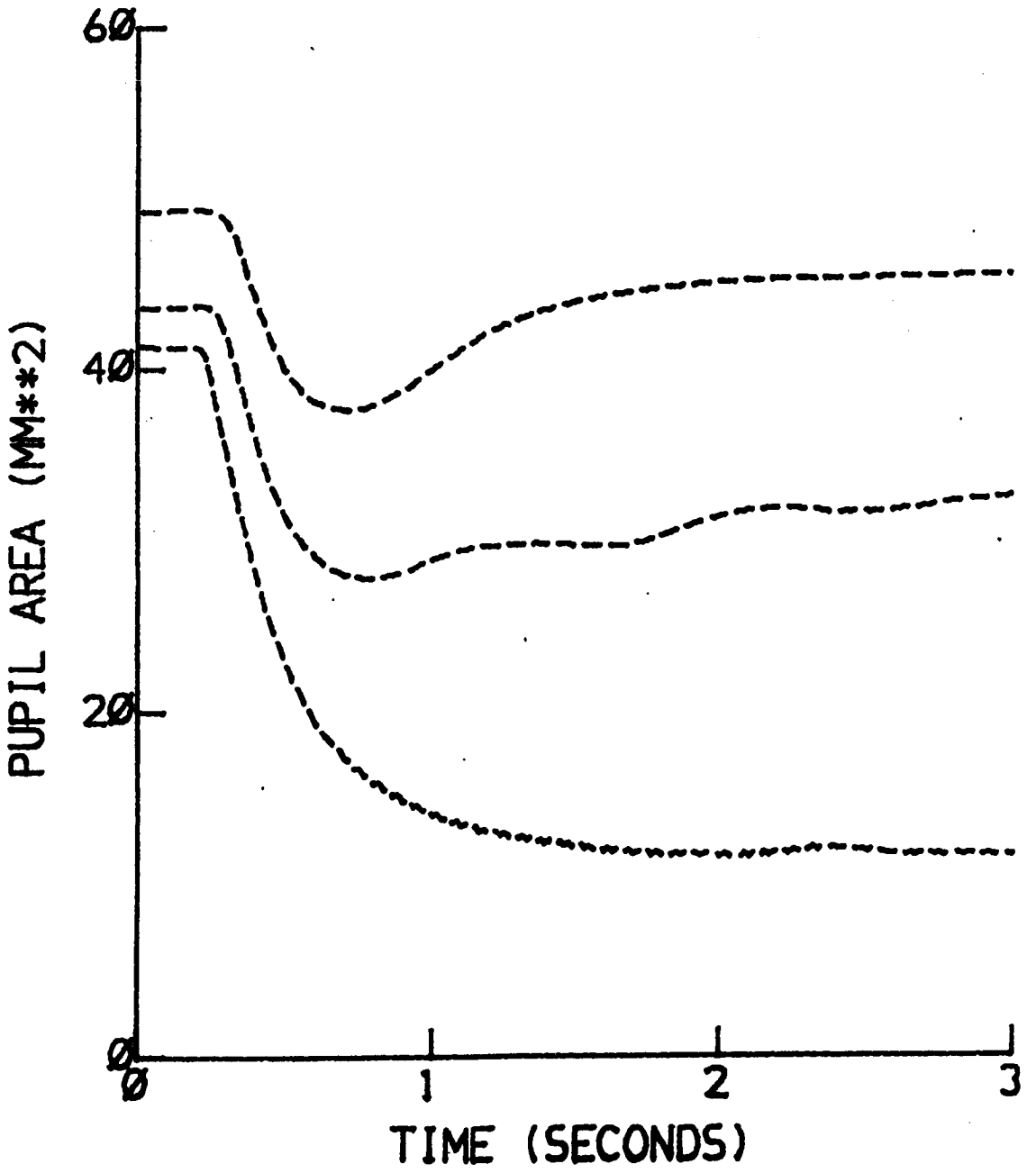
The human pupillary control system has been the object of tremendous interest to biologists and engineers as a prime example of a sensorimotor system. The pupil is an easily observable system, yet it contains a wide variety of nonlinear responses of interest to the engineer. For these reasons it has been the object of considerable investigation.

The pupil has essentially one transient input (light) and one output (pupil size), making it easy to derive simple descriptions of the system. However, other factors affect pupil size. Accommodation, the process whereby the pupil changes size to aid in focus, causes the pupil to constrict without changing the light intensity. Also, light input to the other (contralateral) eye affects the pupil size. However, since these are held constant throughout the experiments, they can be thought of as DC inputs. The pupil is also an easy system to study, unlike many sensorimotor systems. The input is easily controllable, and the output can be observed without invading the body. This allows for a wide range of experiments to be conducted without undue effort.

The pupil system exhibits many forms of nonlinear responses to light inputs. In response to steady-state levels of light, pupil size varies logarithmically with light intensity. The pupil response to varying degrees of change in light intensity is wildly different, depending upon a number of factors. In particular, when stimulated with a small step-increase of light intensity in a dark background, an initially large pupil responds with a very fast constriction. The pupil will then redilate back nearly to its original level (Fig. 1, top trace). This response has been named pupillary escape (Lowenstein & Loewenfeld, 1969). If, however, the increase in light intensity in a dark background is very large, the pupil simply constricts without redilation (Fig. 1, bottom trace), a response named pupillary capture (Usui 1974). Of note here is the fact the pupil constricts very quickly at the onset of the response. If the light intensity is decreased in the same manner, the pupil will slowly dilate. This dilation is slower than the constrictions in either pupillary escape or capture. This asymmetry in the pupil system provides still more nonlinear behavior.

A recent discovery (Sun and Stark, 1983) is the pupil size effect. The pupil, besides being affected by light, is also dependent upon the accommodation system. By changing the accommodative input (moving a target from far to near), the pupil can be made small even in a dark background. Under these

Fig. 1: Pupil response to small, medium, and large step changes in light intensity



conditions, the pupil response to a step increase in light is always pupillary capture, even if the change is small. This implies that the nonlinearities of the pupil system are not strictly dependent upon light. More detailed discussions of these phenomena appear in Sun, Tauchi, & Stark (1983) and Sun, Krenz, & Stark (1983).

The first work in the area of pupil modeling demonstrated that the pupil had some characteristics of a third-order linear control system (Stark & Sherman 1957). Subsequent investigations dealt with the problems of pupillary escape and asymmetry phenomena (Stark 1959, Clynes 1961, Troelstra 1968, Webster 1971). Thereafter, a number of models (Semmlow & Stark 1971, Semmlow & Chen 1977) utilized saturation of the iris to describe the capture response. Another model (Shimizu & Stark 1977) was proposed whereby the internal parameters of the model changed as a function of stimulus intensity. Each of the models described a subset of the actions of the pupillary light reflex system.

A model incorporating many of the ideas contained in previous pupil descriptions (Sun, Krenz & Stark 1982) suggested that separate pathways existed in the pupil controlling high- and low-frequency responses. The high-frequency (or AC) path is responsible for the escape response, while the low-frequency (or DC) path gives rise to the capture response. This model further suggested that the DC pathway controls the parameters in the model. This allows an accommodative input to have a similar effect to light input, and accurately describes the pupil size effect.

The models described above are generally systems models, which try to describe the phenomena without giving a physiological basis for them. Clearly, a model which is homeomorphic (conforms to the structure) is superior in describing the system. The model presented here attempts to do just that. It is based on the separated pathway model, but uses the idea of neural unit recruitment to control the parameters of the model resulting in the escape and capture pupil responses.

This kind of activity is not unlike the size principle in other neuromuscular systems, where some motor units are operational at all times, while others are called in as needed. This scheme also presents a simpler description of the system than does the previous separated path model. The abstract relation between the input and the parameter values is replaced with a concrete and more intuitive interdependence.

Methods

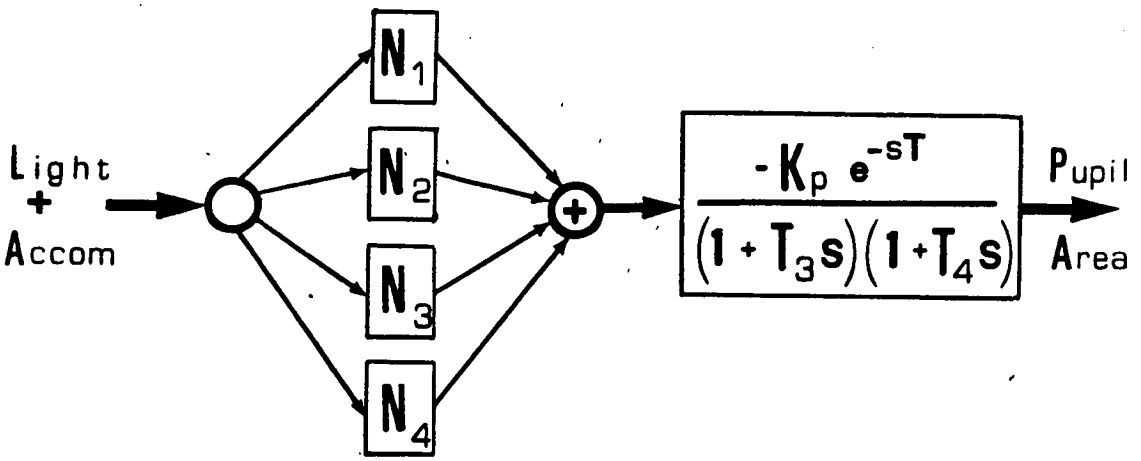
The pupil was stimulated with a spot of light small enough to be unaffected by changes in pupil size, allowing us to study the system 'open loop' (Stark 1959). The light intensities were varied by insertion of neutral density filters in front of the eye. These filters attenuated the signal by -5 log units for small inputs, -3 log units for intermediate inputs, and 0 log units for large inputs. The data pupil size was measured with an infrared TV pupillometer (Myers, et al 1981) which samples and filters pupil size at 60 hz. The waveforms used were the average of 10 data records from each of 5 normal subjects. Each of the waveforms consisted of approximately 3 seconds of transient data after application of the stimulus. These were used to determine the accuracy of the pupil models.

The model simulations were run on an LSI 11/23 computer with floating point hardware and a graphics display. A simulation package was written to provide interactive optimization and display of the model responses. Each model was transformed to its state-space representation and Runge-Kutta integration was used to simulate the responses to step inputs. An optimization technique, a modification of an algorithm by Premermann (1970), was used to provide the best fit to the data. The various data analyses were performed on this same computer.

Model Development

The model developed (Fig. 2) was a variant of the parallel path model of Sun, Krenz, & Stark (1983). This new model had the advantages of being much simpler to understand. It can be thought of as the recruitment of neural units from a pool, with different units becoming active based on the size of the input signal. Thus there exists a more homeomorphic and simpler model which should lend more insight into the mechanisms of the pupillary light reflex system.

The input to the model is assumed to be a signal related to the summation of the logarithm of light and the accommodative input. This signal then feeds into each element of a pool of neural units (N1-N4). Each neural unit has two separate pathways -- an AC (or high-pass) and a DC (or low-pass) path. The AC



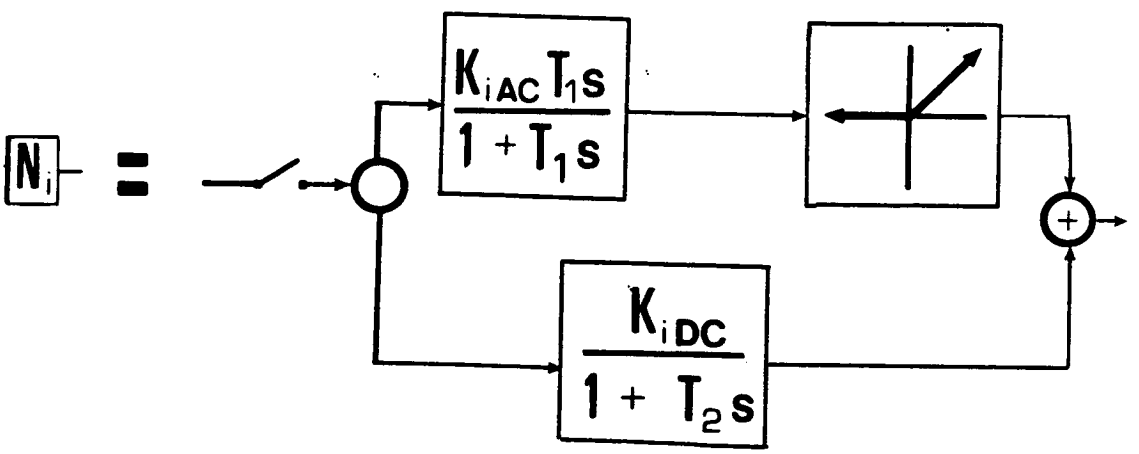


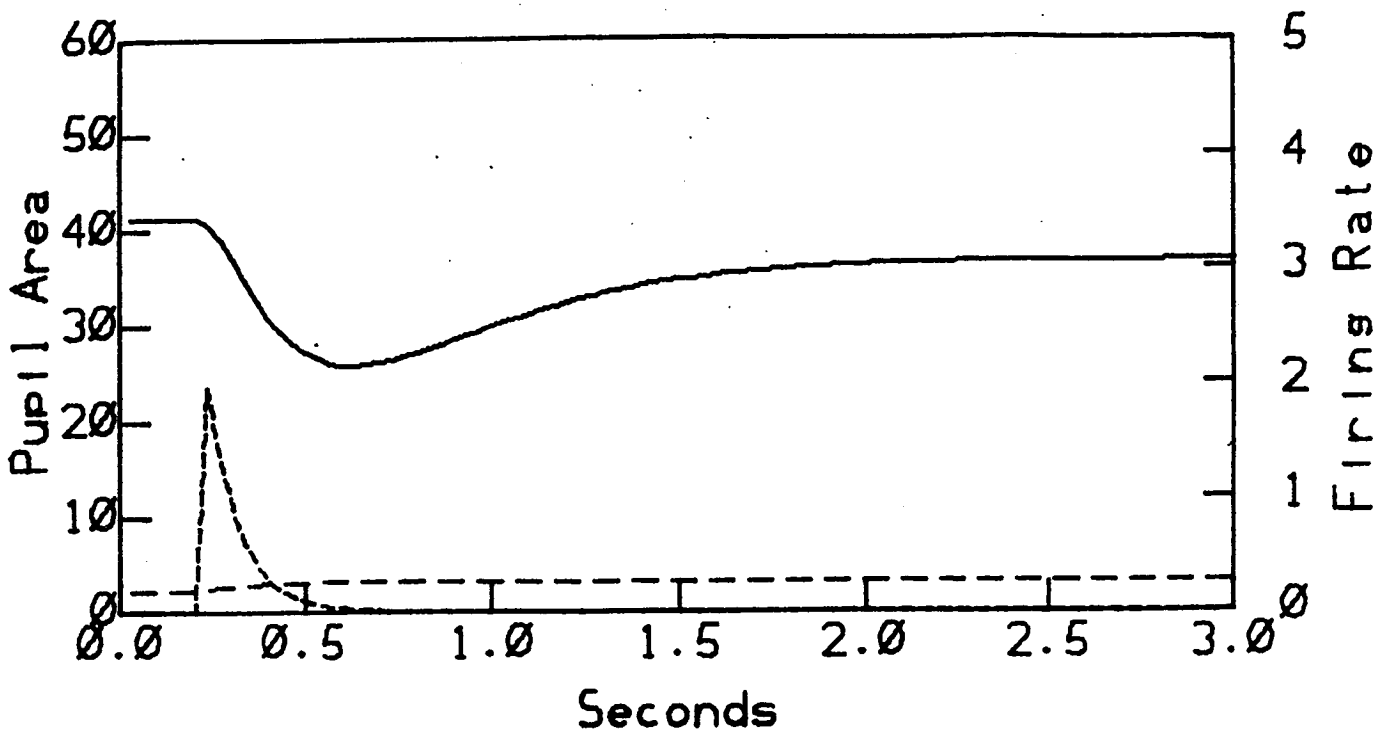
Figure 2: General Neural Unit model. Top trace shows overall structure and neural pool, while lower trace shows compostion of each unit

pathway consists of a lead-lag element in series with a half-wave rectifier. This pathway therefore yields a pulse only when stimulated with a positive-going transient (step increase in light). The DC path is simply a lag operator (leaky integrator) with a time constant similar to that of the AC path. This value controls the 'final', or steady state, pupil size. It should be pointed out that the parallel path model added the accommodative input only to the DC path, whereas this model includes it as part of the stimulus. Since the AC path does not respond to DC signals, and the accommodative input is strictly DC, this transformation is valid.

All intermediate values in the neural unit model are represented as firing rates. These values are given on a scale from 0 to 1, where a sustained value of 1 causes the pupil to reach its minimum size. When the values of each path are thus normalized, they may be compared to determine the relative contribution of each. The aggregate firing rates of the AC and DC paths of the neural pool can be separated to display their temporal characteristics (Fig. 3). In the escape condition (Fig. 3, top graph) the AC rate contributes much more greatly than does the DC. In capture, however, (Fig. 3, bottom graph), the AC produces a very narrow pulse, then the DC path dominates. The firing rates of each of the neural units are summed, then fed into the plant, where they are transformed to yield the corresponding pupil area. The pupil area is determined to have a physiological maximum of about 50 sq.mm and a minimum of about 5 sq.mm. The transformation was approximately linear, with a nonlinear saturation in the minimum ranges.

Each of the neural units in the pool have similar structure, but different gains for each pathway. Each neural unit therefore has its own frequency characteristic. Those with high AC and low DC gains respond to higher frequencies, while low AC- and high DC-gain units respond as low-pass filters. As a first approximation to the systems, four separate neurons are used. One is responsible for pupillary escape, one contributes mostly to pupillary capture, and the other two have characteristics of each, so that the intermediate response is a result of the combination of both. The neural unit responsible for pupil escape has the highest AC gain and lowest DC gain, so that it responds with a fast constriction and redilation to very near its original size. It has the phasic property that it responds more to high frequencies than low. The neural unit controlling capture, on the other hand, has a very high DC gain, and low (=0) AC gain, which causes it to respond as a first-order lag. This tonic neural unit dominates the phasic, and capture results.

Early versions of the model contained neural units which had the same AC or DC gains. It was thought that the system would be simpler to understand if fewer parameters were affected by the nonlinearities. However, using reasonable values for light intensity inputs did not yield satisfactory models. It was



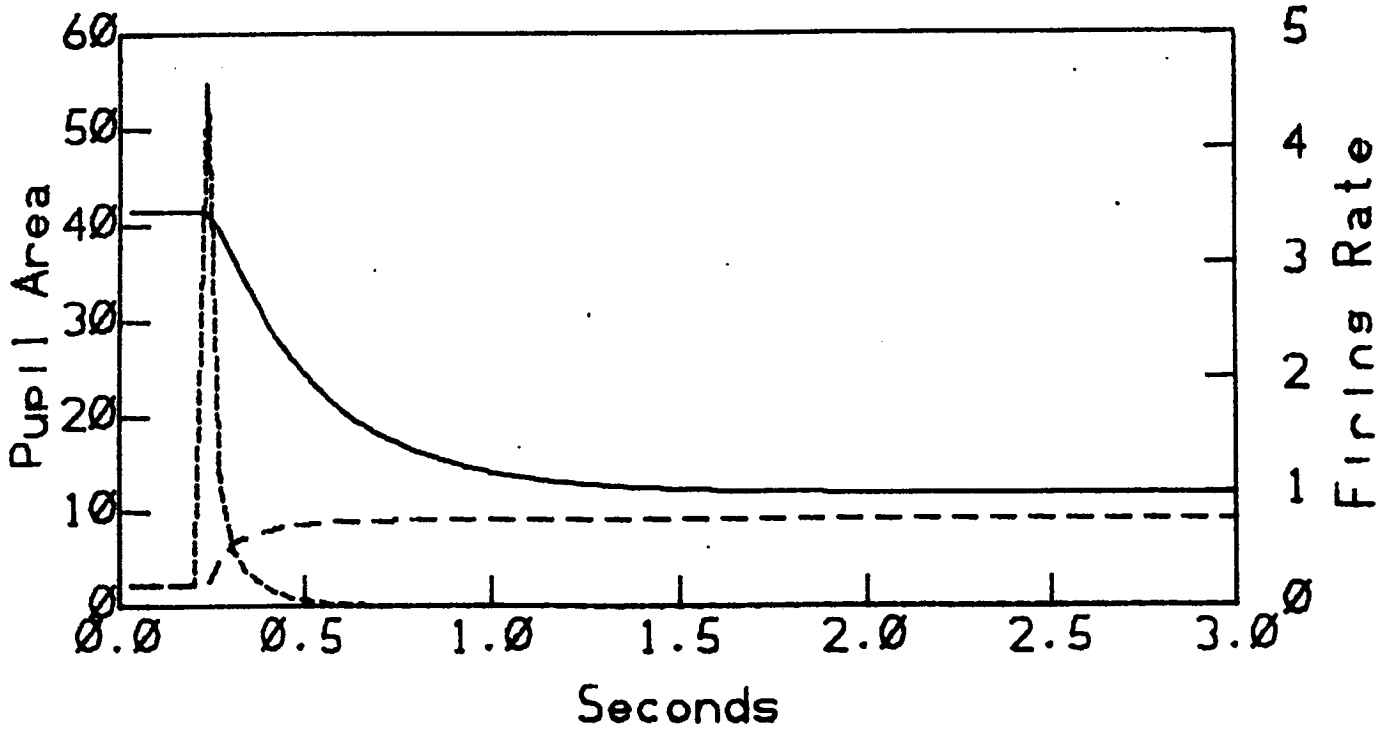


Figure 3: Shows responses and AC and DC aggregate firing rates (dashed lines). Upper graph shows escape, lower shows capture

therefore decided that each unit should have its own set of AC and DC gains.

Originally, each neural unit was allowed to respond to the entire range of the input, even if other units were active. This caused too much phasic component in the pupillary capture response. A slight modification to this total range model resulted in a much better-fitting model. Each neural unit was made to respond to only a specific range of tonic pupil size. Therefore, as one unit came 'on-line', another turned off. This is referred to as the non-coexistent neural unit model. Although the model fits the data more accurately, it is less common physiologically for neural units to turn off in this manner.

A more physiologic model, it is believed, is the condition where the units stay on after they are turned on, as in the total range model. The saturated coexistent unit model, however, limits the output of each neural unit. Thus, as the input increases, each neural unit saturates sequentially. Since the units which are activated first tend to be more phasic, and these are the units whose output saturates first, the result is a more tonic output. This is a relatively simple description of the pupil control system, yet accounts for much of its behavior.

Results

Each of the models was tested using stimuli which generate the basic responses: escape, capture, and the intermediate response. For each model, a best fit was found for the pupillary escape condition. Parameters varied included all time constants and the AC and DC gains for the N1 neural unit. These time constants were then used for all other neural units in the pool. The model responses were compared to the experimental waveforms obtained from the pupil.

Non-coexistent Neural Units

The pupillary escape conditions (large pupil, small stimuli) yielded model results which matched the data well (Fig. 4, top trace). Under these conditions, only one of the neural units was operational. The time constants for the AC and DC paths were found to be about 100ms, while the plant time constants were nearly 280 ms. The average error in each best fit was less than

0.1 sq.mm. per sample. Note that the model fits both the initial transient and the final value very well.

The non-coexistent model fit the capture response nearly as well as the escape (Fig. 4,bottom trace). The average error in this condition was less than 0.11sq.mm. per sample. It is seen that the major source of error was about 0.5 seconds after stimulus onset. It was not possible to fit the initial portion of the response (200-300ms after stimulus onset) without this error. It was felt, however, that this initial phase provided more information in the study of the system. The final value was also well-matched.

The pupil response to an intermediate change in light intensity displayed oscillations. The present model has no means by which to describe this. The fit was therefore constrained to the initial and final portions of the response, without concern about the oscillatory region. Clearly, the fit here is again satisfactory (Fig. 4, middle trace) in most regions of the response.

Coexistent Neural Unit Model

It is seen that the coexistent model provides a good fit to pupillary escape (Fig. 5, top trace). The average error of the fit was less than 0.09 sq. mm. per sample. It is also seen (Fig. 5, middle trace) that the fit to the intermediate response is slightly better than the non-coexistent model fit. However, the capture response provides the best basis for distinction of the two models. The coexistent model fit the capture response well everywhere (Fig. 5, bottom trace), with a slight error at about 300 ms. The fit to the experimental data everywhere else was so good that this early error was deemed acceptible. The average error of the response of the coexistent model was 0.1 sq.mm. per sample, slightly better than that of the non-coexistent model response.

Discussion

The models in this report accurately describe the major known phenomena of the human pupillary control system. However, the most important phenomenon described by these models must be the 'kickoff' effect noted at the onset of pupillary capture.

Figure 4 : The fit of the Non-coexistent Neural unit model to pupillary escape (top trace), capture (bottom trace) and intermediate (middle trace)

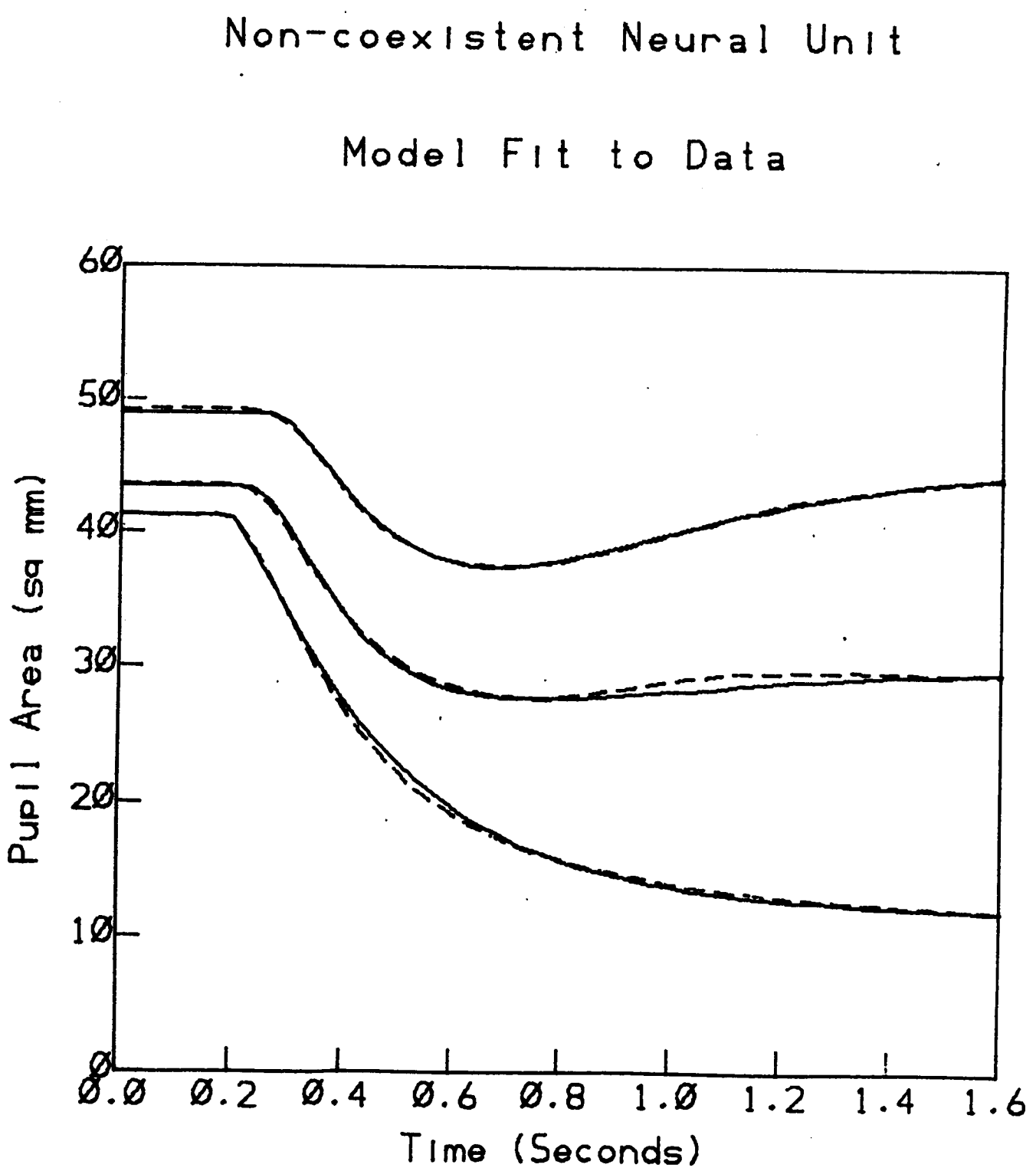
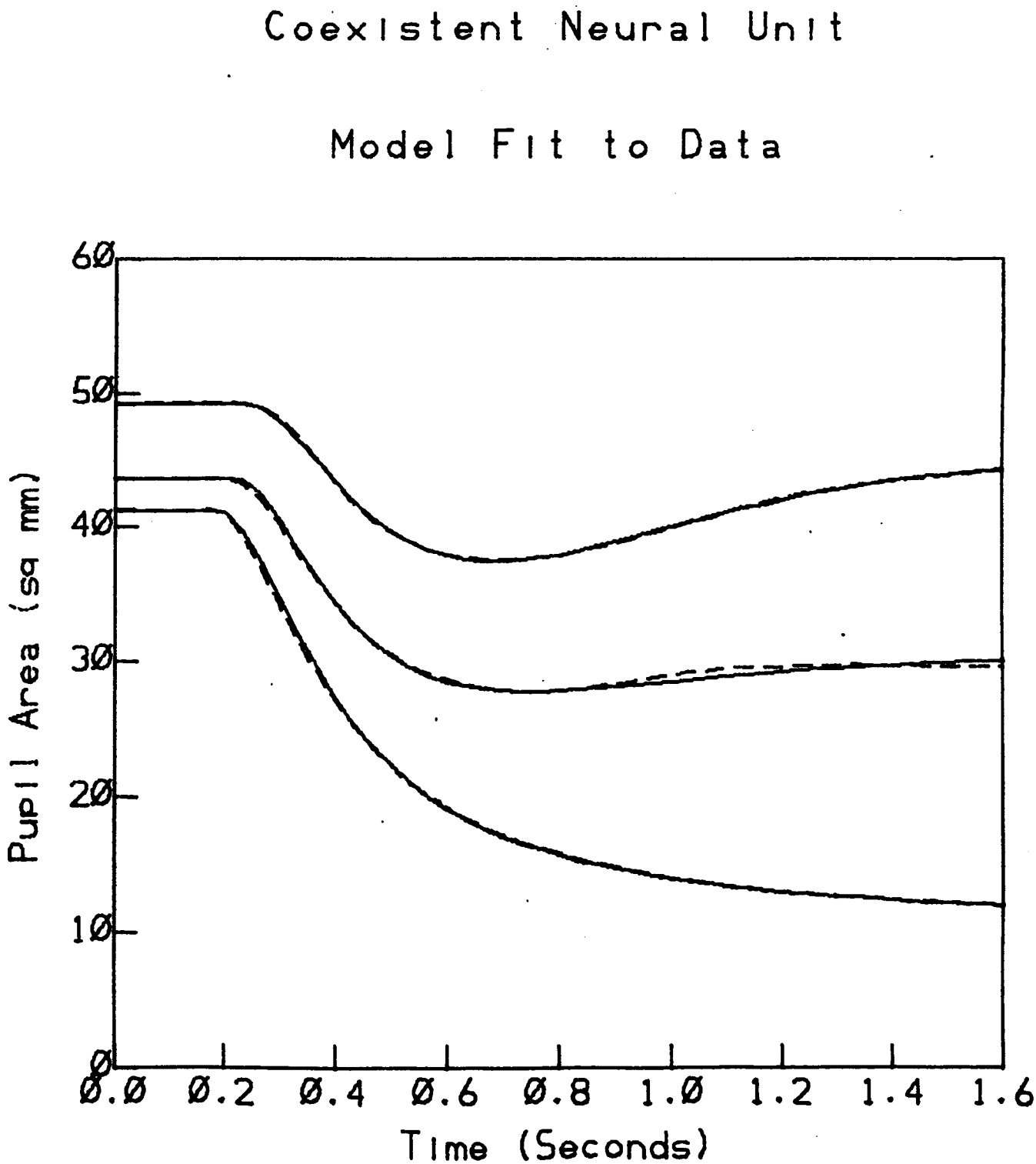


Figure 7: The fit of the coexistent Neural unit model to pupillary escape (top trace), capture (bottom trace), and intermediate (middle trace)



Although the frequency response of the pupil is similar to that of a third-order system at normal gains, its response to step changes in light in a dark background initially resembles at most a second-order system. The neural unit models predict this condtion. As the stimulus first turns on, the N1 unit, responsible for capture, is the only unit operational. This gives rise to a very quick initial response. As the other units turn on, the DC path dominates. Since this path consists of a lag, the initial second-order response tends towards a third-order response. Thus, at low frequencies, the pupil is similar to a third-order (nearly) linear system.

The fits to experimental data are generally better for the coexistent model than for the non-coexistent. It is clear that the coexistent model should describe the 'kickoff' effect mentioned above, since the phasic units will always be operational. The non-coexistent model, however, can also show this effect. All that is required is proper selection of the gains of each individual neural unit. Nevertheless, the non-coexistent model could not provide the overall good fit to pupillary capture seen in the coexistent model. The coexistent model allows for finer AC and DC gain control (continuous) than does the non-coexistent model (piecewise constant). Since it is also felt that the coexistent model is more consistent with known physiology, we conclude that the coexistent model is the best representation of the system.

Conclusion

The models presented here represent a major step towards a simple, accurate description of the pupillary control system. The models are more homeomorphic than their predecessors, and are simpler to understand. The models describe the intersting 'kickoff' effect seen in pupillary capture, which has not to date been seen in other models. We have shown that the best description of the neural pool for the pupillary system is that the units respond to a wide range of inputs, but that the output of each unit is limited. As each subsequent unit becomes active, the unit will stay active, although its output is saturated. This model coincides with those found in other neuromuscular systems.

References

1. Bremermann, H., "A Method of Unconstrained Global Optimization", Math. Biosc., Vol 9, pp. 1-15, 12970

2. Clynes, M., "Unidirectional Rate Sensitivity: A Biocybernetic Law of Reflex and Humoral Systems as Physiologic Channels of Control and Communication", Ann. NY Acad Sci., Vol 92, pp.949-969, 1968

3. Lowenstein, O., and Loewenfeld, I., "Effect of Various Light Stimuli", in THE EYE, ed. by Hugh Davson, Academic Press, p274, 1969

4. Myers, G., Anchetta, J., Hannaford, B., Peng, P., Sherman, K., Stark, L., Sun, F., and Usui, S., "Pupillometry: A Bioengineering Overview", Proc. 17th AFN. Con. on Man. Ctl., UCLA, p523-536, June 1981

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

6. Semmlow, J, and Stark, L., "Pupil Movement to Light and Accommodative Stimulation: A Comparative Study", Vision REs., Vol 13, pp 1087-1100, 1973

7. Shimizu, Y., and Stark, L., "Pupillary Escape", 10th Symp. on the Pupil, New York, June 1977

8. Stark, L., "Stability, Oscillations, and Noise in the Human Pupil Servomechanism", Proc. IRE, Vol 47, p1925-1939, 1959

9. Stark, L., and Sherman, P.M., "A Servoanalytic Study of Consensual Pupil Reflex to Light", J. Neuro-Physio., Vol 20, pp 17-26, 1957

10. Sun, F., Krenz, W., and Stark, L., "A Systems Model for the Pupil Size Effect: I. Transient Data", Biological Cybernetics, in press, 1983

11. Sun, F., Tauchi, P., and Stark, L., " Dynamic Pupillary Response Controlled by the Pupil Size Effect", Experimental Neurology, accepted, 1983

12. Sun, F., and Stark, L., "Pupillary Escape Intensified by Large Pupillary Size", Vision REs., Vol 23, No. 6, 1983

13. Troelstra, A., "Detection of Time-Varying Light Signals as Measured by Pupillary Response", J. Opt. Soc. Amer., Vol 58, pp 685-690, 1968

14. Usui, S., "Functional Organization of the Human Pupillary Light Reflex System", Ph.D. thesis, UC Barkeley, p86, 1974

15. Webster, J., "Pupillary Light Reflex: Development of Teaching Models", IEEE Trans BME-18, No. 3, pp187-194