NASA-TM-85884

NASA Technical Memorandum 85884

An Experimental Evaluation of the Sternberg Task as a Workload Metric for Helicopter Flight Handling Qualities (FHQ) Research

John C. Hemingway

March 1984

National Aeronautics and
Space Administration

ENTER:

```
  8      1      1 RN/NASA-TM-85884
           DISPLAY 08/6/1
      84N20176*#   ISSUE 10   PAGE 1553   CATEGORY 54   RPT#: NASA-TM-85884
      A-9634 NAS 1.15:85884   84/03/00  39 PAGES   UNCLASSIFIED DOCUMENT
UTTL: An experimental evaluation of the Sternberg task as a workload metric for
      helicopter Flight Handling Qualities (FHQ) research
AUTH: A/HEMINGWAY, J. C.
CORP: National Aeronautics and Space Administration. Ames Research Center,
      Moffett Field, Calif.    AVAIL.NTIS   SAP: HC A03/MF A01
MAJS: /*HELICOPTER PERFORMANCE/*RATINGS/*WORKLOADS (PSYCHOPHYSIOLOGY)
MINS: / FLIGHT CONTROL/ PATTERN RECOGNITION/ STATISTICAL ANALYSIS
ABA:  Author
```

NASA Technical Memorandum 85884

An Experimental Evaluation of the Sternberg Task as a Workload Metric for Helicopter Flight Handling Qualities (FHQ) Research

John C. Hemingway, Ames Research Center, Moffett Field, California

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

N84-20176#

TABLE OF CONTENTS

	Page
SYMBOLS	v
SUMMARY	1
INTRODUCTION	1
BACKGROUND	2
Pretest and Findings	6
Research Question	7
METHOD	7
Subjects	7
Apparatus	8
Procedure	9
Experimental Design	12
RESULTS AND DISCUSSION	13
CONCLUSIONS	22
APPENDIX A — POSTFLIGHT AND FINAL DEBRIEFING QUESTIONNAIRE	24
APPENDIX B — SUBJECTS' INSTRUCTIONS	30
REFERENCES	32

SYMBOLS

ADOCS	advanced digital/optical control system
AGL	above ground level
CHPR	Cooper-Harper pilot rating
FBW/FBO	fly-by-wire/fly-by-optic
FHQ	flight handling qualities
HUD	head-up display
ISI	interstimulus interval
NOE	nap-of-the-earth
PNVS	pilot night vision system
RCAH	rate-command/attitude-hold
RD	rate damping
RT	reaction time
SCAS	stability and control augmentation system
S-C-R	stimulus-compatible-response
VMS	vertical motion simulator
VRS	voice recognition system

SUMMARY

The objective of this research was to determine whether the Sternberg item-recognition task, employed as a secondary task measure of spare mental capacity for flight handling qualities (FHQ) simulation research, could help to differentiate between different flight-control conditions. FHQ evaluations were conducted on the Vertical Motion Simulator at Ames Research Center to investigate different primary flight-control configurations, and selected stability and control augmentation levels for helicopters engaged in low-level flight regimes. The Sternberg task was superimposed upon the primary flight-control task in a balanced experimental design. The results of parametric statistical analysis of Sternberg secondary task data failed to support the continued use of this task as a measure of pilot workload. In addition to the secondary task, subjects provided Cooper-Harper pilot ratings (CHPR) and responded to a workload questionnaire. The CHPR data also failed to provide reliable statistical discrimination between FHQ treatment conditions; some insight into the behavior of the secondary task was gained from the workload questionnaire data. A limited review of the literature on the use of the Sternberg task as a workload metric is also provided.

INTRODUCTION

The advent of sophisticated flight-control systems technologies, including digital avionics, system actuators, and fly-by-wire/fly-by-optics (FBW/FBO) systems, has lent a new dimension to the design of flight controls, and stability and control augmentation systems (SCAS). Technical innovations, including side-arm controllers, variable SCAS, and advanced multimode displays, are commonplace in modern military and commercial aircraft. Previous constraints imposed on control system designs by conventional mechanical, hydraulic, and electromechanical flight-control systems no longer apply. The contemporary design engineer must approach flight-control systems design within the total mission/systems context. He is more likely to be constrained by his own creative abilities than by available technologies. Greater operational flexibility, if not improved cost, reliability and maintainability figures-of-merit ensure the continuing application of modern technology. These advanced designs are resulting in greater demands on flight handling qualities (FHQ) research personnel to glean the full advantage afforded by the new technologies, and to determine optimal man-machine mixes.

Historically, FHQ research has relied on the abilities of highly trained and experienced test pilots to assess the adequacy of handling qualities in developmental aircraft. However, FHQ rating scales, such as the Cooper-Harper pilot rating (CHPR) scale, only generally reflect system performance and pilot workload; moreover, they are subject to pilot biases (ref. 1). Traditional FHQ rating scales alone may not be sufficiently sensitive to provide adequate discrimination between the sophisticated flight-control systems of contemporary aircraft. Thus, future FHQ research may be faced with an important problem of finding a sufficiently sensitive metric to permit system performance distinctions to be made, particularly across the central and negative portions of the FHQ scale (ref. 2).

One possible solution being investigated by several researchers is to assess the pilot's spare mental capacity while performing a control task to obtain an index of the workload associated with a specific control system and, by extension, the FHQ of the system under study. In the current investigation, a relatively straightforward secondary task — the Sternberg task — was superimposed upon the pilot's primary flight-control task in order to assess reserve mental capacity. This was an exploratory investigation in which the primary objective was to determine whether performance on the Sternberg item-recognition task can provide an objective index of pilot mental workload in an FHQ simulation study.

The author is grateful to Mr. E. W. Aiken who provided the opportunity to conduct this research under the Army's Advanced Digital/Optical Control System (ADOCS) program. Mr. C. L. Blanken and Ms. R. W. Cortilla provided valuable assistance in data collection during the actual simulation. A special note of thanks is due Dr. E. M. Huff, Dr. J. H. Stevenson, and Ms. Mary Childress for their consultation on data analysis, and to Mr. W. B. Carson and D. L. Miller who performed the major analyses. Special appreciation is also due Mr. E. J. Hartzell, Dr. J. G. Guercio, and Ms. S. G. Hart for their technical reviews and valuable recommendations.

BACKGROUND

The Sternberg task evolved from Donders' early work (ref. 3) using subtraction methodology to measure component processing times for stages believed to exist between stimulus onset and response execution. A translation of Donders' work is given in Koster (ref. 4) together with an extension of the methodology by Sternberg into an additive-factor method for detecting processing stages, assessing their attributes, and for determining their stochastic independence.

A description of the basic task along with findings from two exploratory investigations on human memory scanning was provided by Sternberg (ref. 5). A typical stimulus ensemble for the Sternberg task consists of a homogeneous set of elements (K) from which n elements are randomly drawn to compose a positive set. The remaining (K-n) elements comprise the negative set. Before each set of trials, test subjects are required to memorize the positive set, which typically varies in size from one to six elements. Stimuli are presented serially, randomly drawn from either the negative or positive subsets of K. The subject's task is to perceive a displayed stimulus, decide whether it is from the positive or negative set, and respond appropriately as rapidly as possible within the fixed interstimulus interval (ISI). Correct responses and reaction-time (RT) data from stimulus onset to response execution are collected for analysis. Response error rates for the Sternberg task are normally between 1% and 2%. Sternberg RT data versus memory set sizes are commonly presented as linear functions via regression analyses. The y-intercept value is interpreted as time for stimulus processing and response formulation, independent of set size; the slope is interpreted as the rate of search through short-term memory.

This RT measurement methodology was originally proposed by Sternberg (ref. 5) for studying the retrieval of symbolic information in short-term memory, which he later extended to research into the mechanisms underlying human information processing. In a more recent paper, Sternberg (ref. 6) reexamined the assumption underlying the basic item-recognition paradigm, and discussed the implications of findings of other researchers for the methodology and model. Four of the findings most relevant to this research are (1) mean RTs increased linearly with the size of the memory set; (2) negative and positive responses produced approximately the same slope; (3) the rate of increase is approximately 38 msec for each additional element of the positive set; and (4) the y-intercept varies about a central value of 400 msec. These basic properties have been reaffirmed by many researchers and have been interpreted to mean that memory search is "exhaustive," or sequentially completed for all elements in a given memory set.

In addition to the value of Sternberg methodology in studying basic mechanisms of information processing (see refs. 7 and 8), the relative stability of performance on the memory search task makes it an attractive candidate as a secondary task for studying workload. Knowles (ref. 9) stated that a secondary or auxiliary task can be used to discover how much additional work an operator can undertake while still performing the primary task to some specified system criteria; Knowles and Rose (ref. 10) indicated that secondary task performance is sensitive to differences in problem difficulty; that it reflects increased ease in handling the control task with practice; that it reflects differences in workload between crewmembers; and that it exposes control law difficulties during critical flight segments.

Researchers at the University of Illinois (refs. 8 and 11) have been using variations of the item-recognition paradigm in basic information processing studies to investigate structural, capacity, and resource theories of attention, including dual-task performance. Central to much of this work is a structural model of information processing based on a multiple resource theory of attention (ref. 12) which has implications for workload measurement and task integration. The model presupposes the structure of resources to include processing stages, processing modalities, and processing codes. Examples of processing stages included perceptual and central processing, response selection, and execution. Modalities included visual and auditory inputs and manual and vocal responses. Codes could be either verbal or spatial.

Wickens et al. (ref. 13) evaluated the model's ability to define resource reservoirs in a series of experiments that employed the Sternberg task among others. A primary compensatory tracking task with either rate or acceleration control dynamics was administered in the first experiment. A Sternberg task variant was administered as a concurrent task on selected trials. Dimensions that were varied on the secondary task included perceptual load (superimposing a visual grating), central processing load (memory set size), and response load (single- versus double-key press). The results of this investigation showed that stage-defined resources could be successfully differentiated by the Sternberg task. One surprising result was that RT's were not affected when the double response load was superimposed upon the secondary task. The authors suggested that the additional load was reflected in poorer performance on the primary tracking task.

In a second experiment by Wickens et al. (ref. 13), a failure-detection task was paired with the Sternberg task to evaluate the function of spatial and verbal processes in defining resource reservoirs for encoding and central processing stages. Contrary to predictions based on multiple resource models of attention, differences in slope for longer memory set sizes for both verbal and spatial stimuli failed to materialize. Instead, differences between single- and dual-task performance were reflected in the y-intercept values, which were elevated particularly for the spatial condition. The authors pointed out that higher intercept values were consistent with multiple resource theory, placing greater demands on spatial instead of on verbal resources for perceptual interactions between the central processing loads (memory set sizes) imposed on the Sternberg task.

In a follow-on study, Wickens et al. (ref. 14) conducted three experiments to examine coding effects on performance. Baseline data on verbal and spatial Sternberg tasks were collected in the first experiment. Data from both variants produced generally linear functions, although the authors reported finding a significantly greater slope for the spatial condition, as well as a weak quadratic tendency in the function. Reliable interactions were reported between memory set size, verbal versus spatial stimuli, and response hand, which the authors interpreted as providing evidence for separation and resource competition between and within hemispheres for processing these stimuli. The same Sternberg task variants and stimuli were used in the second experiment in combination with a memorization task. Results suggested that the spatial secondary task shared more common resources with the memory task than its verbal counterpart. The authors concluded that the spatial/verbal dichotomy is an important element in interpreting dual task interference patterns.

The results of the third experiment, which shared an autopilot monitoring/failure-detection task with the same two Sternberg tasks, indicated that more interference occurred with the spatial variant than with the verbal, because the failure-detection task was spatial in nature. As in the previous investigation by Wickens et al. (ref. 13), no interaction was discovered between dual-task load and memory set size. Rather than indicating that the primary task had no perceptual or central processing demands, the additivity seemed to be related more to the automaticity of certain processes. Theoretical considerations surrounding the use of the Sternberg additive factor method for assessing the demands of the primary task have been discussed within a multiple-resource modeling context in several other publications (refs. 11, 12, and 15).

The simplicity and relative stability of the Sternberg task methodology make it suitable for multimodality research as well. Vidulich and Wickens (ref. 16) reviewed recent multimodality research, and reported findings from two experiments in which Sternberg tasks were used to investigate differences between combinations of input (auditory or visual) and output (verbal or manual) modalities. Findings from these investigations are described in terms of multiple-resource theory, and help to clarify both code and modality relationships in facilitating time-sharing efficiency. Two of their findings (ref. 16) of particular relevance to the current investigation are:

1. Task priorities exerted a reliable effect on performance, and this effect was greater as the tasks shared more common resources.

2. Although clear performance differences were observed between input/output modality conditions, these were not reflected in the assessment of subjective workload ratings.

Observer ratings were not sufficiently sensitive to judge dual-task demands; however, the demands of different types of primary tasks could be assessed by specific types of secondary tasks, with greatest sensitivities obtained when task demands tap common resource pools. In other words, sharing input modalities may lead to a deterioration of the intermittent discrete reaction time task, whereas sharing outputs could lead to a deterioration of the continuous manual task (ref. 17).

Several researchers have employed Sternberg task methods to evaluate workload demands in real and simulated flight. Crawford et al., (ref. 18) using a cockpit simulator, evaluated two levels of flight control and four levels of multifunction switching, using the Sternberg task as a secondary measure of reserve information-processing capacity. Performance on the Sternberg task differed by 54% between flight-control levels, and by 20% to 31%, respectively, for simple and complex multifunction switching tasks. Corrick (ref. 19) employed the Sternberg task to evaluate four alternative display formats used to present missile launch envelope information. Although subjects failed to report any large performance-related differences between the displays, the author found large differences in secondary task performance which they attributed to the workload imposed by display formatting of the primary task.

An interesting variation of the Sternberg task was employed by Johnson (ref. 20) in studying the effects on reaction time of terrain background, downlook angle, and response-processing levels in target acquisition. The stimulus ensemble consisted of eight tank targets in place of traditional alphanumeric stimuli. Three groups of five subjects each were required to make a positive or negative set determination, recognize a target (friend or foe), or identify the target. Reaction-time data were collected and analyzed for memory set sizes of one, two, and four for each of the three acquisition tasks. Greatly inflated y-intercept (1,400 vs 400 msec), and slope (200 vs 40 msec per memory set size) values over those reported by Sternberg (ref. 6) were attributed to differences in target and task complexities. Statistically significant performance differences were found between levels of target background, downlook angles, and acquisition tasks. The results were consistent with Sternberg's findings, supporting the serial exhaustive model of memory search, and the authors concluded that the Sternberg task method served as a useful tool in understanding the observer's cognitive processes in complex target acquisition tasks.

Schiflett et al. (ref. 21) used the Sternberg task in actual flight tests on board a T-33 variable-stability research aircraft. The goal of the study was to evaluate two levels of flight control and two alternative head-up display (HUD) formats under simulated instrument meteorological conditions. The primary control task flown in the T-33 fixed-wing aircraft consisted of glide-slope and localizer intercepts, and an ILS approach to touchdown. Subjects flew four approaches for each combination of display conditions and handling-quality level. The findings of Schiflett et al. were similar to those of Corrick (ref. 19) — poor agreement was obtained between the primary flight performance measures, Cooper-Harper ratings across primary flight-control task configurations, and performance on the secondary task. It appeared that the pilots compensated for the more demanding flight-control variations, but at the expense of reserve information-processing capacity. Of particular interest were the apparent sensitivities of measures obtained on the secondary task for positive memory set sizes of one, two, and four. All data appeared to fit the Sternberg paradigm remarkably well, with both slope and y-intercept discriminants showing consistency across levels of flight-control and display type for both subjects. It should be noted, however, that no parametric statistical analyses were performed, and that analysis was restricted to exploring trends in reaction-time and response-error data. The authors concluded that the pilots had more mental reserve capacity while flying the predominantly pictorial/symbolic HUD configuration than when flying the conventional HUD format with scales and alphanumerics; however, they recommended further research be done to establish the efficacy of the Sternberg task in evaluating aircrew tasks.

Pretest and Findings

In order to determine if the Sternberg task could be effectively integrated into an FHQ simulation study, a preliminary investigation was performed in a fixed-base simulator at Ames Research Center. The Sternberg task was superimposed on a primary flight-control task for two pilot subjects during a comparative evaluation of an integrated isometric controller for nap-of-the-earth (NOE) flight (ref. 22). In addition to the secondary task, measures of pilot control activity, CHPR data, and pilot commentary were collected across the different experimental conditions. The experimental protocol for the Sternberg task was previously described; the methodology approximated that described by Schiflett (ref. 23) employing the same alpha stimulus ensemble, the same positive set sizes (one, two, and four), and the same 7.0-sec interstimulus interval (ISI). The primary flight-control task was performed in a fixed-base, rotary-wing simulator, and typified an NOE mission with maneuvering, hover and bob-up, and straight-and-level flight segments.

Linear regression fits of the reaction-time data obtained on the baseline condition (secondary task alone) generally conformed to the classical Sternberg model for serial probe recognition. The y-intercepts for the two subjects, 665 and 598 msec, respectively, were longer than the typical 400-msec values reported by Sternberg.[1] Slope values for ascending memory set sizes were 43 and 27 msec, respectively, compared with the 38-msec value reported by Sternberg (ref. 6). Numbers of reversal and time-out errors (RT > 500 msec) on the baseline task were less than 1%.

The behavior of the Sternberg function became erratic for both subjects with the addition of the primary flight-control task. Specifically, y-intercept values across all experimental conditions ranged between 850 and 1,550 msec, and the slope of the function across memory set sizes ranged from a high of 47 msec per set size to a negative 111 msec. Response error rates (reversal and time-out errors) increased from 1% on the baseline condition to over 10%. RT's for these responses were discarded from other analyses.

Pilot ratings (CHPR) regarding FHQ were generally poor, ranging from a high of 3.5 (satisfactory with unpleasant characteristics) on the cruise flight segment to 7.0 (unacceptable) on the maneuvering and hover and bob-up segments. An analysis of variance of secondary task conditions showed only one statistically significant difference ($p < 0.001$) between the baseline Sternberg condition and the other combined experimental conditions, and no memory set size effects upon RT's emerged when the primary task was added. The absence of this effect was apparent from the erratic slopes obtained on the secondary task when combined with the primary task. It is important to recall, however, that data were obtained from two pilots only, and results must be viewed accordingly. Although both subjects accepted the addition of the secondary task, poor handling qualities on the primary flight-control task appeared to have inhibited their performance on the secondary task.

Although the procedural integration of the Sternberg task into an FHQ research paradigm was accomplished, it was difficult to collect sufficient numbers of data points (reaction times) for analysis without incurring primary task overload from

---

[1]A post hoc examination of the software program revealed an implementation bias which consistently inflated RT's between 70 and 100 msec. No correction was applied to these pretest data.

the secondary task. Additionally, the subjects complained that the location of the secondary task response key on the collective control interfered with the manual, primary flight-control task, thus indicating an alternative response mode for the secondary task.

Research Question

The objective of the investigation was to determine whether the Sternberg task, used as a secondary task to measure pilots' spare mental capacity in FHQ simulation research, could differentiate between different flight-control systems. It was hypothesized that any change on the relatively stable Sternberg task would reflect variation in encoding, output, or processing loads of the primary task competing for common resources with the Sternberg task; an increase or decrease in the y-intercept would reflect a change in encoding or output components, while mental information processing would be reflected by changes in the slope. The secondary task input and output modalities were configured to achieve compatibility with the resource demands of the primary flight-control task. Thus, both primary and secondary tasks involved a sharing of the spatial input modality, whereas the output modality on the secondary task was switched to verbal (ref. 15) for this experiment.

Unanticipated performance decrements on the primary rather than the secondary task (ref. 13) and prioritization shifts (ref. 16) have already been discussed as potential problems in the application of this workload method. Consequently, all practice and at least one data run were completed for each FHQ experimental condition without the secondary task. If significant performance differences were subsequently found on the primary task, they would constitute evidence for rejecting the Sternberg task as a useful workload metric on future FHQ simulation research.

In addition to the aforementioned objectives, we also wanted to investigate the relationship between subjective workload assessment ratings, CHPR data, and performance on the Sternberg secondary task. Consequently, questionnaires and CHPR data were systematically collected throughout this investigation.

METHOD

This investigation was conducted in the Vertical Motion Simulator (VMS) at Ames Research Center. The VMS cab was configured to represent a generic single-seat helicopter, incorporating a conventional helicopter front instrument panel. The primary objective was to evaluate the FHQ of three different primary flight-control configurations and three stability and control augmentation levels in low-level flight regimes. A Sternberg item-recognition task was superimposed upon the primary flight-control task to determine whether this particular secondary task could provide an indication of pilot workload.

Subjects

Four helicopter test pilots served as subjects for this investigation; two were NASA test pilots; one was a Canadian National Aeronautics Establishment pilot, and one was a Boeing-Vertol test pilot. All subjects possessed current first-class medical certificates and had participated in similar helicopter FHQ research projects.

Apparatus

Flight-control system— The hardware for the three primary flight-control systems consisted of (1) a conventional helicopter cockpit configured with cyclic, collective, and yaw damper pedals; (2) an advanced configuration incorporating a four-axis, isometric side-arm controller with roll controlled by lateral force, pitch by longitudinal force, yaw by rotational moment, and thrust by vertical force; and (3) a mixed configuration similar to the four-axis controller described above, except that thrust was incorporated on the conventional collective. All flight-control hardware remained in place during this investigation, but the desired control configurations were actuated under software control (see fig. 1).

Helicopter control dynamics were generated by a SIGMA-8 computer programmed with a nine-degree-of-freedom generic teetering-rotor helicopter model (ref. 24) which included both stability and control augmentation. Details and rationale for the selection of the three levels of stability and control augmentation investigated in this simulation were described by Aiken (ref. 25). Basically they included (1) unaided: simulated basic UH-1 mechanical control system, without stabilization; (2) rate damping: augmented angular rate damping in pitch, roll, and yaw; and (3) rate-command/attitude-hold (RCAH): integral prefilters in pitch and roll to provide a rate-command/attitude-hold feature.

Visual-display system— The visual system consisted of raster- and stroke-written television monitors mounted perpendicularly to one another such that the imagery from each monitor was projected onto a common combining glass. Subjects viewed combined imagery, focused at optical infinity, through a lens system located above the front instrument panel. A realistic external visual scene, also presented to the subjects on the raster display, was generated by the Ames Visual Flight Attachment (VFA) apparatus — a gimbaled television camera mounted on a gantry system which traversed a terrain board. Camera motion was under computer control and responded appropriately in six degrees of freedom to pilot control inputs.

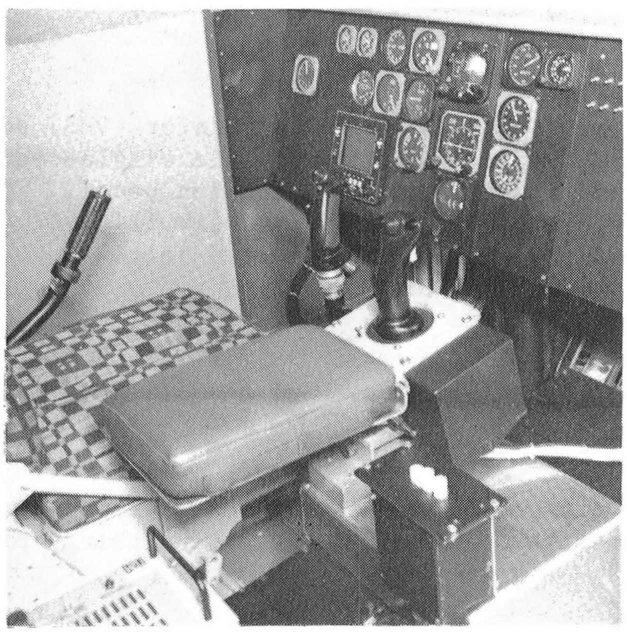


Figure 1.- Flight-control hardware.

The stroke-written television monitor contained all critical flight parameters, advisory information, and Sternberg secondary task stimuli. All Sternberg task display characters exceeded the threshold for limiting resolution and were presented just to the left of central field-of-view. Display-mode select control functions were interfaced with a PDP-11/55 computer that generated the stroke-written symbols. Primary flight display symbology consisted of modified pilot night vision system

(PNVS) display formats for various flight modes, including descent, accelerating and decelerating transitions, hover and bob-up, and cruise (ref. 26); see fig. 2.

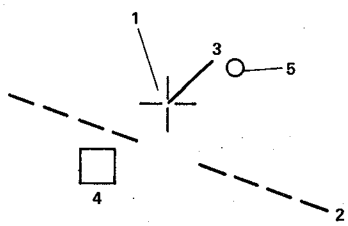
Mission scenario— A predefined mission profile comprising four distinct flight segments was used throughout this investigation. Figure 3 shows the mission segments, including a 457-m (1,500-ft) AGL, 6° descent into a nap-of-the-Earth (NOE) ground track composed of maneuvering, hover and bob-up, and cruise flight segments. The profile was defined such that each flight segment required approximately 2 min to complete.

Data recording— A Voterm voice recognition system (VRS) was used by subjects in responding to the Sternberg secondary task to avoid manual response incompatibility problems identified in the preliminary study. Reaction times were recorded by calculating elapsed time from stimulus onset to the first utterance detected by the VRS. Subjects' digitized responses (yes or no) and RT data were stored on the simulation computer following each run, together with handling qualities data. In addition to FHQ objective performance measures, Cooper-Harper ratings, pilot comments concerning each of the four mission segments, and responses to a postflight questionnaire were obtained at the end of each simulation run. Each subject also completed a final debriefing questionnaire following the simulation; questionnaires are presented in appendix A.

Procedure

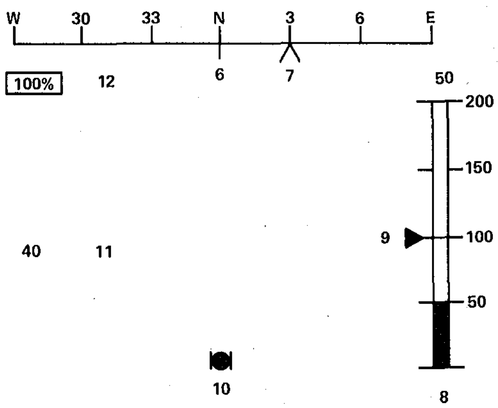
System performance and calibration checks, including picture quality assessments and flight-control response dynamics, were carried out before each test session. The Sternberg secondary task procedure was identical to that described for the pretest, except that subjects were required to respond aurally, instead of manually. The four subjects were acquainted with the primary flight-control task, introduced to the Sternberg secondary task, and given a pre-recorded set of instructions that explained the procedures and performance priorities to be employed on the primary and secondary tasks (appendix B). They were told that they would be required to fly an 8-min NOE mission composed of a 1,500-ft, 6° descent; low-altitude maneuvering; hover and bob-up; and straight-and-level flight segments. The subjects were also furnished with the appropriate altitude and airspeed requirements for each segment. The PNVS display-mode switching control functions were explained for each mission segment. The subjects were instructed to maintain a high and constant level of performance on the primary flight-control task, although their performance would be scored on the Sternberg secondary task as well. They were reminded that although the letters from the secondary task would remain on the display for 5 sec, they should respond as rapidly as possible after making their yes or no decision. Following each run, subjects were required to assign a Cooper-Harper rating to each of the four mission segments and to provide comments on selected facets of both the primary and secondary tasks.

After receiving their instructions, the subjects were brought into the simulation cab, given several practice sessions on the Sternberg task, and then tested solely on the secondary task to ensure proficiency and obtain baseline data. The core experiment, comprising the flight-control conditions depicted in table 1, consisted of testing all subjects on the conventional flight controls first, then running a pair of subjects through one of the advanced side-arm flight-control configurations. All subjects were permitted familiarization runs before commencing data collection on each primary flight-control condition, first without the secondary task, then with balanced presentations of the three memory set sizes (one, two, and



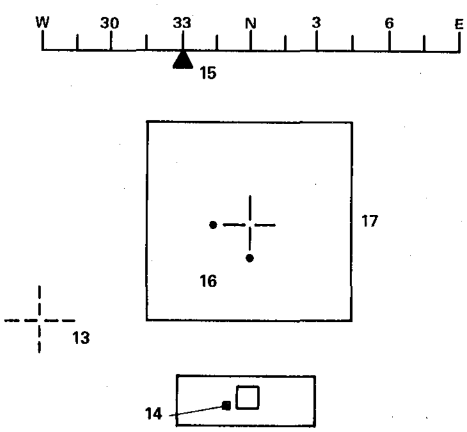
(a)

SYMBOL	INFORMATION
1. Aircraft reference	Fixed reference for horizon line, velocity vector, hover position, cyclic director, and fire control symbols
2. Horizon line (cruise mode only)	Pitch and roll attitude with respect to aircraft reference (indicating nose-up pitch and left roll)
3. Velocity vector	Horizontal Doppler velocity components (indicating forward and right drift velocities)
4. Hover position	Designated hover position with respect to aircraft reference symbol (indicating aircraft forward and to right of desired hover position)
5. Cyclic director	Cyclic stick command with respect to hover position symbol (indicating left and aft cyclic stick required to return to designated hover position)



(b)

SYMBOL	INFORMATION
6. Aircraft heading	Moving tape indication of heading (indicating North)
7. Heading error	Heading at time bob-up mode selected (indicating 030)
8. Radar altitude	Height above ground level in both analog and digital form (indicating 50 ft)
9. Rate of climb	Moving pointer with full-scale deflection of ±1,000 ft/min (indicating 0 ft/min)
10. Lateral acceleration	Inclinometer indication of side force
11. Airspeed	Digital readout in knots
12. Torque	Engine torque in percent



(c)

SYMBOL	INFORMATION
13. Cued line of sight	Overlays designated target position on background video when target is in display field of view
14. Coarse target location	Designated target position with respect to display field of view (inner rectangle) and sensor limits (outer rectangle)
15. Target bearing	Designated target bearing (indicating 330° or 30° to left of current heading)
16. Target location dots	Illumination of two adjacent dots indicates display quadrant in which designated target is located
17. Missile launch constraints	Limits with respect to aircraft reference for successful weapon lock-on to designated target

Figure 2.- Modified PNVS display symbols (from Aiken, ref. 25).

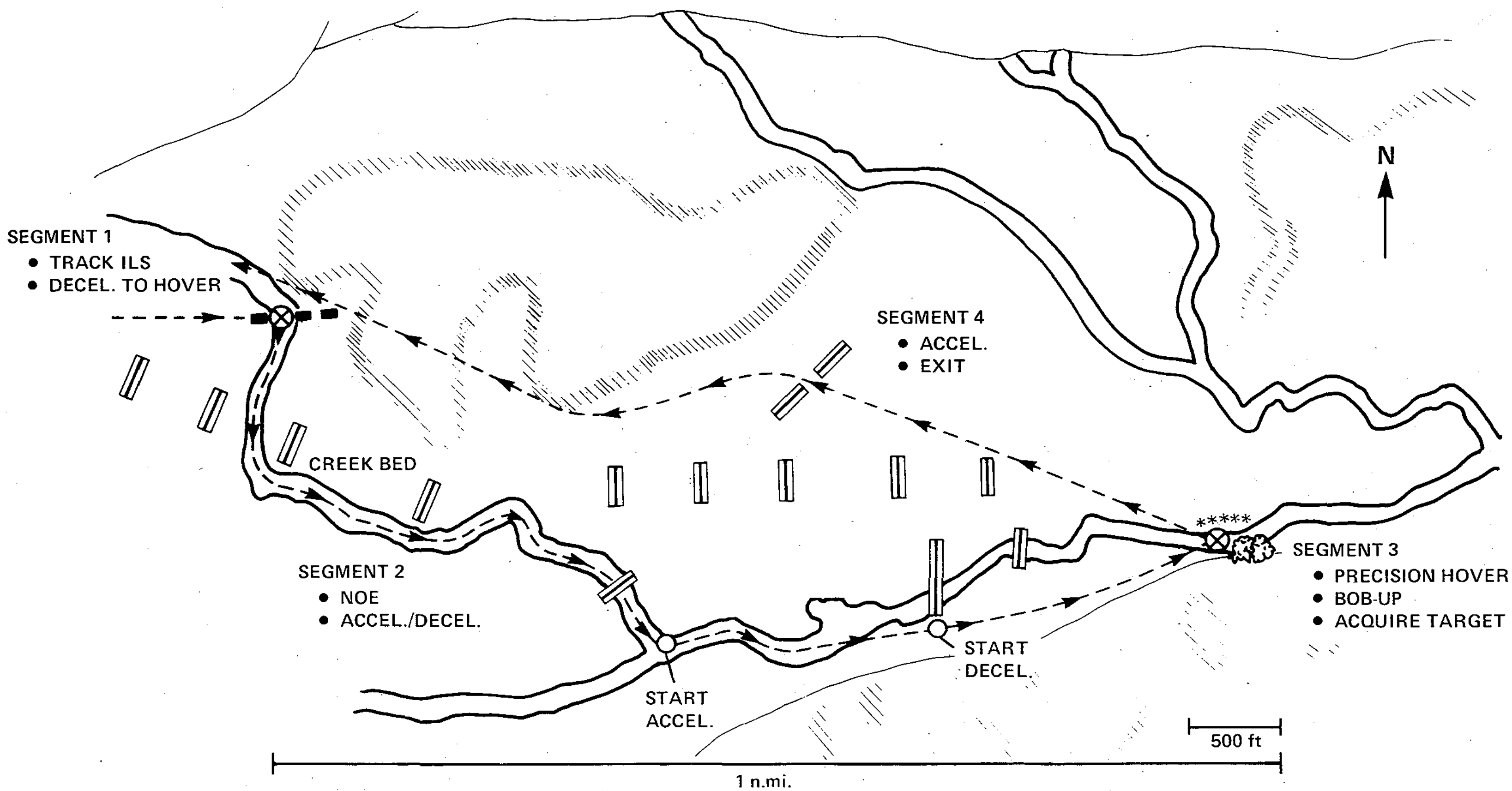


Figure 3.- Predefined mission segments for primary flight-control task.

four) of the Sternberg task while flying the mission segments. The cab controls were then reconfigured, and the subjects were tested on the other advanced control configuration. The other pair of subjects received the same treatment, except in reverse ordering of the advanced side-arm flight-control configurations. Cooper-Harper ratings and pilot comments were solicited after each run, and a 15-min rest period was permitted between each control-system reconfiguration. Training and test sessions, running two subjects sequentially, required 1.5 days to complete per pilot.

Experimental Design

In addition to collecting baseline data on the Sternberg task and on the FHQ conditions depicted in table 1, the original plan included superimposing three levels of the secondary task on each of the nine FHQ combinations given in table 1. Computer difficulties severely limited the scope of the original FHQ design; complete data were obtained for only two of the original four subjects for the rate-command/attitude-hold (RCAH) stability augmentation level across the three control configurations as illustrated in table 1. Limited data on two Sternberg memory set sizes were also obtained from one subject flying two of the control configurations under the rate-damping condition. The order of the presentations of control configurations and of Sternberg memory set sizes was initially balanced across four subjects, with the exception that the conventional flight-control configuration was always presented first. The design was to have progressed from the best stability augmentation level (RCAH) to the most difficult (unaided). The actual order of presentation for the two subjects on whom data were obtained is shown in table 2.

TABLE 1.- FLIGHT HANDLING QUALITIES EXPERIMENTAL DESIGN (THE THREE STABILITY AUGMENTATION LEVELS WERE TO HAVE BEEN EXAMINED IN SEPARATE INVESTIGATIONS

<table>
<tr><th rowspan="2">CONTROL CONFIGURATION</th><th colspan="3">STABILITY AUGMENTATION LEVEL</th></tr>
<tr><th>(A) UNAIDED</th><th>(B) RATE DAMPED</th><th>(C) RCAH</th></tr>
<tr><td>(1) CONVENTIONAL (CYCLIC, COLLECTIVE, YAW DAMPER PEDALS)</td><td></td><td colspan="2">EXPERIMENTAL</td></tr>
<tr><td>(2) THREE AXIS FORCE STICK AND COLLECTIVE (SIDE-ARM CONTROLLER)</td><td></td><td></td><td>CONDITIONS</td></tr>
<tr><td>(3) FOUR AXIS FORCE STICK (SIDE-ARM CONTROLLER)</td><td></td><td colspan="2">TESTED</td></tr>
</table>

TABLE 2.- ACTUAL PRESENTATION ORDER OF SECONDARY TASK MEMORY SET SIZES WITHIN THE FHQ EXPERIMENTAL DESIGN

SUBJECT	EXPERIMENTAL CONDITIONS	SESSION NUMBER (4 RUNS/SESSION)				
		1	2	3	4	5
3	SCAS/CONTROLLER*	C1	C2	C3	B1	B3
	MEMORY SET SIZE	0, 1, 2, 4	0, 4, 1, 2	0, 2, 4, 1	0, 4, 1	0, 1, 4
4	SCAS/CONTROLLER*	C1	C2	C3		
	MEMORY SET SIZE	0, 2, 4, 1	0, 4, 1, 2	0, 1, 2, 4		

*SEE TABLE I

RESULTS AND DISCUSSION

FHQ aspects of this research project, including the pilots' evaluations of the adequacy of the three flight-control configurations and stability augmentation levels (table 1) for helicopter terrain flight, were reported by Aiken (ref. 25, p. 5). Aiken presented the following conclusions:

1. With conventional controllers, a rate- or attitude-stabilized vehicle, and a head-up display, adequate but unsatisfactory handling qualities were achieved for the low-altitude tasks investigated.

2. Satisfactory handling qualities may be achieved with a head-up display and a properly designed two-axis displacement side-stick controller for either a rate- or attitude-stabilized vehicle. Critical side-stick design features include the force-deflection characteristics and mechanization of the trimming function.

3. Attitude stabilization is required to maintain adequate handling qualities with either the rigid three-axis (pitch, roll, and yaw) or four-axis controller configuration evaluated during this investigation.

The remainder of this report addresses workload-related findings using the Sternberg task, and the effect of introducing a secondary task on the primary flight-control task.

Linear regression fits of the Sternberg baseline data obtained from the four subjects in this experiment are shown in figure 4. These data were corrected for length of utterance and Voterm processing times to enable comparison with classical Sternberg task results. The curves and means were adjusted downward for subjects 1 and 2 by averaged yes and no utterance times of 530 and 540 msec, respectively, plus the 224-msec Voterm processing time. For subjects 3 and 4, individual utterance and processing times were recorded and subtracted on each experimental trial. There were

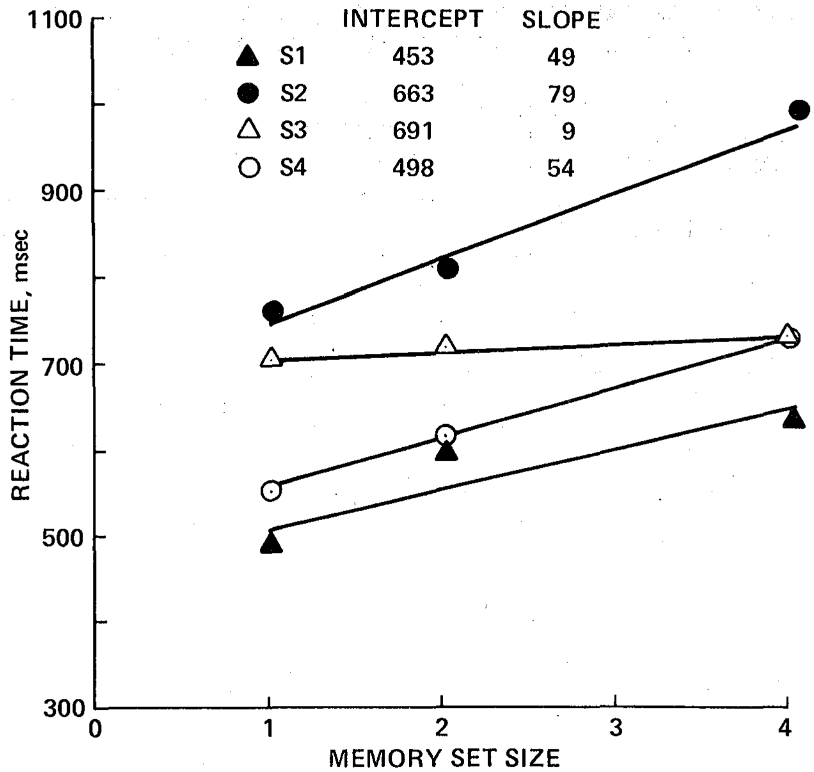


Figure 4.- Plots of subjects' baseline Sternberg data. Symbols represent means for each subject for memory set sizes one, two, and four. Error rates were insignificant over these data.

24 trials per memory set size; however, only 15 of the original 24 discrete RT's could be used per cell for baseline data analyses — a result of unrecoverable data on subject 3. Consequently, the last nine trials were deleted from each cell of the remaining three subjects to facilitate analysis. Error rates (reversal errors) for all subjects remained relatively constant, varying between 1% and 2% — results similar to those reported by Sternberg (ref. 5).

The results of an ANOVA (ref. 27) performed on the baseline data for the four subjects are given in table 3. Both main effects, subjects (S), and memory set size (M) were highly significant, ($F = 21.45$ ($df = 3,42$) $p < 0.0001$) and ($F = 19.54$ ($df = 2,28$) $p < 0.0001$), respectively, and there were no significant interactions. A strength-of-association statistic — omega-squared (ref. 28) — was calculated for both statistically significant main effects. Subjects accounted for about 27% of the RT variance, and memory set size captured about 12%.

Four ANOVAS (ref. 27) were calculated using each subject's baseline Sternberg data to separate out the effects of memory set size on RT. A summary of these analyses together with r-squared strength-of-association measures is shown in table 4. The results showed that memory set size had a significant effect on the performance of all but the third subject. The r-squared values showed that memory set size captured about 9%, 22%, 0.8%, and 20% of the RT variance of the four subjects, respectively.

An orthogonal polynomial regression (ref. 27) was run on these data to investigate the linearity assumption of the Sternberg model. The results showed that the

TABLE 3.- RESULTS OF ANOVA (ref. 27) PERFORMED ON STERNBERG BASELINE DATA; STRENGTH OF ASSOCIATION VALUES (ref. 28) INDICATE TOTAL ACCOUNTED VARIANCE

Source	df	Sum of squares	Mean squares	F	p	Strength of association
Subjects	3	2141213.75	713737.92	21.45	0.0001	0.2680
Error	42	1397803.50	38288.39			
Memory set (M)	2	922212.31	461106.16	19.54	.0001	.1154
Error	28	660737.86	23597.78			
SM	6	210458.80	35076.47	1.39	N.S.[a]	--
Error	84	2122299.70	25265.47			
Mean	1	83527381.61	83527381.61			--
Error	14	536037.48	38288.39			

[a]not significant.

TABLE 4.- RESULTS OF SEPARATE ANOVA's (ref. 27) PERFORMED ON MEMORY SET SIZE FOR ALL SUBJECTS; THE PERCENT OF VARIANCE CAPTURED BY MEMORY SET SIZE IS REFLECTED IN THE r-SQUARED STRENGTH OF ASSOCIATION MEASURES

Source	df	Sum of squares	Mean squares	F	p	Strength of association
Memory set ($\underline{S}_1$)	1	196984.75	196984.75	4.622	0.05	0.0862
Error	49	2088382.00	42620.04			
Memory set ($\underline{S}_2$)	1	510218.69	510218.69	13.961	.001	.2182
Error	50	1828596.00	36571.92			
Memory set ($\underline{S}_3$)	1	10259.06	10259.06	.643	N.S.	.0081
Error	79	1259894.00	15948.02			
Memory set ($\underline{S}_4$)	1	357086.87	357086.87	18.369	.001	.1967
Error	75	1457935.00	19439.13			

slope was significant for all but the third subject, and that no significant quadratic term existed for any of the data sets. The polynomial regression analysis is summarized in table 5.

As previously discussed, all secondary task data for subjects 1 and 2 under dual-task experimental conditions were unrecoverable because of data acquisition problems. All remaining RT data for subjects 3 and 4 on the secondary task were summarized by flight-control configuration, stability augmentation level, and mission segment (table 6). In the original design, 2 min of flight time were planned for each of the four mission segments in order to ensure near-equal data samples across segments. In actuality, pilots varied greatly in the amount of time they devoted to each segment, presumably in the interests of performing FHQ evaluations. Thus, there were unequal numbers of data points obtained for each of the three

TABLE 5.- RESULTS OF POLYNOMIAL REGRESSION ANALYSIS (ref. 27) RUN ON SUBJECTS' BASELINE RT DATA FOR THREE MEMORY SET SIZES; ANALYSIS PROGRAM AUTOMATICALLY TERMINATES AT NEXT HIGHEST VALUE

Subject	Degree	Multiple r-squared	Regression coefficient	df	F	p
1	0	–	4776.4954	2	3.04	0.0572
	1	0.08619	443.8203	1	1.51	.2250
	2	.11467	-255.1048	–	–	–
2	0	–	7741.5346	2	6.90	.0023
	1	.21815	714.2853	1	.31	.5793
	2	.22319	108.5881	–	–	–
3	0	–	6416.0000	2	.36	.6988
	1	.00807	101.2619	1	.09	.7616
	2	.00927	-38.9191	–	–	–
4	0	–	5483.3271	2	8.95	.0003
	1	.19673	597.5495	1	.02	.8970
	2	.19691	-18.3496	–	–	–

memory set sizes. The smallest sample size was six data points per set size; the largest was 44 data points per set size. Pilots also failed to respond to trials (out-of-time errors), particularly during the descent and hover and bob-up mission segments. Failure to respond during the descent was the logical consequence of requiring pilots to perform a heads-up secondary task while flying a head-down ILS approach to breakout at 100 ft AGL. Failure to respond during hover and bob-up occurred because pilots attended to external visual cues, rather than to their primary, head-up flight display. Fifty-one out-of-time errors were recorded on the core (RCAH) experiment, but were not included in the data analysis. Reversal errors (yes/no confusions) were between 1% and 2% in the baseline condition.

Reaction times, and presumably pilot attention, varied greatly in the presence of the primary flight-control task. In order to improve normality or equality of variance, several possible transformations were considered, including logarithmic transforms and Windsor transforms (ref. 29). Logarithmic transforms of these data tended to reduce the problem of outliers, not atypical of such data, accompanied by a dramatic decrease in standard deviations. Unfortunately, this transform also obscured relatively large differences in slope and intercept values, and prevented interpretation of the data within the context of the Sternberg paradigm.

In addition to these manipulations, subjects' RT data were replotted using a technique proposed by Schiflett (ref. 23) which involves a truncation of the secondary task RT's above a predetermined value (1500 msec). The application of the technique discussed by Schiflett, using an arbitrary 1500-msec RT cutoff, tended to normalize many of the extreme excursions in the data sets. The following apparent benefits were noted:

1. By truncation of data sets at 1500 msec, y-intercept values, artificially inflated or deflated by outlying RT data, were decreased or increased, respectively. This also facilitated comparisons with the Sternberg task baseline data.

TABLE 6.- MEMORY SET SIZE MEANS, STANDARD DEVIATIONS, SLOPES, AND INTERCEPTS. FROM THE STERNBERG TASK RT DATA SHOWN AS A FUNCTION OF THE FLIGHT CONTROLLER, SCAS, AND FLIGHT SEGMENT (see fig. 3)

Subject number	Experimental conditions	SEG	Memory set size X(σ) 1	2	4	Slope	Intercept
3	Baseline		697(98)	716(152)	726(114)	9.18	692
	RCAH/Conv	D	1169(619)	975(515)	1151(550)	1.25	1094
		N	873(401)	1256(1136)	1464(748)	182.71	774
		H	910(455)	1097(1127)	1051(448)	34.85	936
		S	651(167)	1138(936)	1236(692)	172.97	603
	RCAH/3-Ax	D	970(351)	1895(1350)	1425(913)	76.00	1281
		N	1246(881)	1263(692)	1217(431)	-11.98	1270
		H	1049(418)	1217(715)	1064(338)	-5.60	1144
		S	1293(478)	958(453)	978(348)	-89.86	1297
	RCAH/4Ax	D	1057(571)	1513(935)	1726(1116)	202.92	961
		N	1027(493)	1109(459)	1088(598)	13.79	1044
		H	1034(411)	1356(939)	1056(740)	-9.48	1162
		S	749(172)	1044(430)	941(197)	49.12	806
	RD/Conv	D	1179(860)	-	1388(1011)	69.70	1109
		N	1157(539)	-	907(289)	-88.33	1241
		H	1385(1059)	-	1022(192)	-120.70	1505
		S	862(415)	-	910(344)	16.13	845
	RD/3-Ax	D	1485(740)	-	1504(875)	6.32	1479
		N	1023(456)	-	1007(253)	-6.95	1030
		H	1145(441)	-	1100(488)	-14.92	1160
		S	1246(689)	-	1149(760)	-32.34	1278
4	Baseline		550(108)	610(167)	714(133)	54.27	498
	RCAH/Conv	D	738(243)	989(417)	1449(775)	236.32	507
		N	870(315)	790(149)	1209(603)	126.95	663
		H	674(110)	1010(317)	1119(215)	130.61	635
		S	778(139)	717(151)	922(367)	56.93	672
	RCAH/3-Ax	D	822(162)	1844(1122)	1862(865)	263.61	957
		N	1199(759)	1274(898)	1115(472)	-37.94	1287
		H	1160(936)	1205(447)	1379(948)	76.50	1068
		S	797(215)	1105(484)	968(324)	33.24	883
	RCAH/4-Ax	D	2232(1492)	1561(550)	1396(911)	-229.23	2253
		N	884(468)	935(249)	1127(445)	82.24	791
		H	1927(1310)	1311(890)	1599(1248)	-48.72	1668
		S	957(460)	-	879(126)	-32.90	1016
	RCAH/3-Ax	D	1104(666)	-	1556(956)	150.95	953
		N	930(267)	-	965(312)	11.48	919
		H	992(432)	-	1450(1039)	152.72	839
		S	1159(284)	-	1158(347)	-.31	1159

2. Radically high, as well as negative, slope values were decreased, or made positive by this truncation, with only three exceptions.

3. Standard deviations were dramatically and consistently decreased across all set sizes using truncated data sets.

Despite apparent advantages derived from this treatment, adequate justification for discarding outlying data was not available in this investigation because subjects were not informed of a predetermined cutoff for acceptable RT's. In fact, numerous latencies greater than 1500 msec, believed to be the result of task overload, were noted on secondary task responses during the actual simulation, despite instructions to respond as rapidly as possible. This finding is not only relevant to the feasibility of establishing empirical criteria for defining an out-of-bounds type error, as discussed by Schiflett (ref. 23), but also affects the potential utility and interpretability of secondary RT task data within the Sternberg paradigm. The possible application of Windsor transformations to these limited sets of RT data was rejected for reasons similar to those stated above, but might be considered as an alternative to truncating actual data when greater numbers of observations can be taken within selected set sizes.

The RT data obtained from this investigation do not fit within the classical Sternberg interpretation. Additionally, transformations and manipulations of the data set failed to improve interpretability within this context. It was not evident whether this was a result of (1) primary-task overload, which violates basic, a priori assumptions described by Sternberg (ref. 7), (2) a limited sample size, or (3) insufficient experimental control in progressing from relatively precise laboratory investigations to dynamic simulations.

An analysis of variance (ref. 27) was performed on the RT data from the core experiment employing the RCAH stability augmentation model (table 1) to help identify unknown sources of variation, and to determine the feasibility of pooling selected RT data across selected experimental treatments. Despite skewness in the raw data from the secondary task, ANOVA's were considered sufficiently robust to circumvent the logical problem of running test statistics on data not meeting the model assumptions. As is appropriate for a repeated measures design, F-tests for main effects and selected interactions were recomputed using the next, higher-order interaction as the error term. Strength-of-association values were also estimated for any statistically significant main effect or interaction; they are reported together with the results of the ANOVA in table 7.

From the strength-of-association values (table 7), it is evident that statistically significant findings only account for about 6% of the variation on the secondary task, compared with 39% on the baseline Sternberg condition. Although low strength-of-association values are not unexpected in complex investigations, they failed to provide support for statistically significant ANOVA findings in accounting for much of the total experimental variation. Additionally, three second-order and two higher-order significant interactions precluded pooling the data across the three latter flight segments. Of the main effects, only flight segments (S) approached a statistical level of significance ($F = 5.63$ ($df = 3.3$) $p < 0.09$). This was probably a result of difficulties in attending to the secondary task during the descent and the hover and bob-up flight segments, as discussed earlier. The difficulty level of the secondary task across flight segments was seemingly affected differentially by the choice of subjects, as well as by the particular controller being flown. The statistically significant, higher-order interactions are difficult to interpret, and no probable explanations are apparent. Thus, pilot FHQ ratings and questionnaire data were examined to identify the factors that contributed to the complicated behavior of the secondary task in this investigation.

Four CHPR's were elicited for each mission segment following each flight (table 8). An ANOVA summary is given in table 9. Of the three main effects, only

TABLE 7.- RESULTS OF ANOVA (ref. 27) PERFORMED ON THE CORE RCAH PORTION OF THIS INVESTIGATION; STRENGTH OF ASSOCIATION VALUES ARE PROVIDED AS AN INDICATION OF TOTAL ACCOUNTED VARIANCE

Source	df	Sum of squares	Mean squares	F	p	Strength of association
Subjects (S)	1	65712.30	65712.30	0.14	N.S.	-
Controller (C)	2	11048081.71	5524040.86	2.38	N.S.	-
Memory set (M)	2	6736264.78	3368132.39	4.08	N.S.	-
Segment (G)	3	27267811.48	9089270.49	5.63	N.S.	-
SC	2	4634441.05	2317220.53	4.77	<.01	0.0065
CM	4	6600664.76	1650166.19	.98	N.S.	-
SM	2	1649452.36	824726.18	1.70	N.S.	-
SG	3	4840061.00	1613353.67	3.32	<.02	.0067
CG	6	13602626.60	2267104.43	9.32	<.01	.0190
MG	6	2407021.77	401170.29	2.91	N.S.	-
SCM	4	6758336.21	1689584.05	3.48	<.01	.0094
SCG	6	1460181.65	243363.61	.50	N.S.	-
SMG	6	827935.84	137989.31	.28	N.S.	-
CMG	12	11784057.13	982004.76	.79	N.S.	-
SCMG	12	14900174.93	1241678.99	2.55	<.01	.0208
Error	1240	602715670.48	486061.02			-

TABLE 8.- COOPER-HARPER PILOT RATINGS OBTAINED ON THE CORE RCAH PORTION OF THIS INVESTIGATION

Controller configuration	Memory set size	Mission segment			
		Descent	Maneuvering	Hover/bob-up	Straight/level
Conventional	0	5/4/5/6	4.5/6/5/5	5/5/4/4.5	4/4/4/3
	1	6.5/4/5/4	5.5/5/5/5	5.5/5/4/5.5	5/4/5/3
	2	5.5/4/5/7	5.5/5/5/4	5.5/6/4/3	5/5/4/3
	4	6/4/5/7	5.5/5/5/8	5.5/5/4/4	5/4/3/3
Four-axis side stick	0	4/5/6/5	6/5/6/4.5	7/5/4/5	5/4/4/3
	1	6/5/6/5	7/5/5/4	8/5/4/5.5	5.5/4/5/3
	2	5/5.5/6/5	7/5/6/4	8/5.5/5/4.5	5.5/4.5/5/3
	4	5/5.5/6/4.5	7/5/6/4	8/5.5/5/5	5.5/4/5/4
Three-axis side stick and collective	0	5/5/5/7	5/4/5/6	6/5/4/4.5	4.5/5/4/4
	1	5.5/5/5/5	6/5/5/6	6/5/4/6	5/6/5/4
	2	5.5/5/6/5	5.5/5/5/6	6/6/4/6	5/5/4/4
	4	5.5/7/5/5	5.5/5/5/6	6/5/4/6	5/6/4/4

Note: The four numbers given in each cell correspond to the four subjects.

TABLE 9.- SUMMARY OF ANALYSIS OF VARIANCE TABLE FOR THE COOPER-HARPER PILOT RATINGS

Source	SS	df	MS	F	p
Mean	4870.2552	1	4870.2552	590.707	0.000
P/	24.7344	3	8.2448		
CONTROL	5.8932	2	2.9466	1.007	.420
CP/	17.5547	6	2.9258		
MEMSET	3.1719	3	1.0573	1.742	.228
MP/	5.4635	9	.6071		
CM	.2422	6	.0404	.122	.992
CMP/	5.9349	18	.3297		
SEGMENT	31.1927	3	10.3976	3.261	.073
SP/	28.6927	9	3.1881		
CS	4.7526	6	.7921	.622	.711
CSP/	22.9245	18	1.2736		
MS	1.6302	9	.1811	.472	.880
MSP	10.3594	27	.3837		
CMS	4.9870	18	.2771	1.017	.457
CMSP/	14.7109	54	.2724		

the flight segment approached a level of statistical significance ($F = 3.261$ ($df = 3,9$) $p < 0.075$). Neither memory set size nor control configuration was significant, and there were no higher-order interactions. The tendency for flight segment to approach significance is interesting, since it was the only dependent variable which also approached statistical significance for the Sternberg task. It is probable that a fair degree of pilot compensation was being required to perform the primary flight-control task, and that the secondary task may have been periodically ignored in favor of providing consistent FHQ ratings. The former finding tends to support the earlier contention that excessive loading on the primary task may have inhibited the classical behavior of the Sternberg function.

Each subject was required to complete a postflight questionnaire after each set of runs with a different controller/stability augmentation level combination, as well as a final debriefing questionnaire following the simulation. Examples of both questionnaires and tabulated means and standard deviations for each question from the core (RCAH) experiment are contained in appendix A. The following is based on the responses of the four pilot subjects to the postflight questionnaire:

1. Subjects considered the effects of the secondary task harmful to their primary flight-control task, especially while flying three- and four-axis control configurations. Two subjects noted that this was particularly apparent during the descent segment, with conflicting head-up/head-down visual demands. In fact, several pilots initially neglected to respond to the secondary task altogether upon entering the descent and the hover and bob-up segments. One pilot stated that the need to cross-check the secondary task had a profound effect on degrading primary task performance and increasing overall workload. A second pilot stated that the secondary task, especially with the four-axis controller, had a strikingly negative effect on

his ability to analyze handling qualities in flight, and felt that his overall flight performance was compromised.

2. Pilots reported that the need to perform display mode-select control functions had very little effect on their performance on either the primary or secondary tasks.

3. With the exception of one Ames Research Center project pilot, who considered himself to be overtrained, a tendency was noted for the remaining three subjects to consider themselves undertrained on the primary flight-control systems task. As anticipated, this tendency was greatest for the experimental three- and four-axis controllers. Two of the pilots rated themselves extremely undertrained for flying the three-axis control configuration, despite additional training runs.

4. There was a tendency for subjects to rate their performance slightly worse on the secondary task as the number of letters increased in the memory set.

5. There was a decided tendency for subjects to increase their scan pattern difficulty ratings with the addition of the secondary letter recognition task. Two of the four pilots stated that this difficulty was particularly accentuated during descent — a finding undoubtedly related to the head-up/head-down visual demands during that segment.

6. There was a slight tendency to rate the presentation location of secondary task stimuli on the head-up display as nonoptimal. This tendency was related to the head-up/head-down problem during descent, however, and only one subject suggested relocation from the nine o'clock to the twelve o'clock position. Another subject commented that the HUD is a very busy display and must be used during hover and bob-up; addition of the secondary task in this already busy display makes this segment most difficult. However, he further commented that moving secondary task symbols to another display would be unacceptable for this maneuver.

7. Subjects did not indicate that they had made many false identification errors, and for the most part, their perceptions were correct. On several of the most difficult runs, or when subjects fell behind on the primary task, their perception of the number of false identification errors appeared to be inflated. One pilot reported that he had difficulty triggering the voice response mechanism. A small number of such errors were noted, but were discarded from further analysis.

8. As noted earlier, subjects appeared to experience greatest difficulty responding to the secondary task while flying the descent, and pilot ratings on this question tended to substantiate this observation. Difficulties were also identified during hover and bob-up, followed by the NOE segment. Subjects reported only minor difficulties in attending to the secondary task during the straight-and-level flight segment.

9. The consensus regarding the relative difficulty of executing the four flight segments rated the 6° descent as most difficult and straight-and-level as the easiest. Consensus was mixed regarding the NOE and hover and bob-up mission segments, but appeared to be at least partially a function of the flight-control configuration (see appendix A, table 10). Variations in pilot ranking on this question were also related to the amount of exposure to side-arm flight controllers, with specific configurations affecting segment difficulty differentially.

Postflight questionnaires were also completed by two subjects flying the rate-damping stability augmentation model with conventional flight controls. Subject ratings on these questionnaires were compared with their previous ratings flying the same controller, but employing RCAH stability augmentation. Only minor differences were noted between ratings with one possible exception: on question 9, one pilot rated the hover and bob-up segment as most difficult to fly under RCAH, but easiest under the rate-damping stability augmentation level. The 6° descent was rated most difficult to execute under rate damping.

Tabulated results on the final debriefing questionnaire are provided in appendix A (table 11). Subjects indicated that they had received adequate training on the secondary task, and on flying the mission profile, but one pilot noted that more training would have been beneficial on the flight controllers, and two pilots indicated that more exposure was needed to the SCAS levels. For the most part, subjects appeared to remain strongly motivated throughout the simulation. Only one pilot indicated some difficulty with fatigue, and another with boredom in flying the same mission profiles. A slight tendency to consider the secondary task distracting from the primary task was again noted, as in the postflight questionnaire results.

CONCLUSIONS

The results of this investigation failed to support the continued use of the Sternberg task as a secondary measure of pilot workload in exploratory FHQ evaluations in which primary task demands may be excessive. Not only did the Sternberg function fail to materialize from data obtained under dual task conditions, but relevant main effects failed to reach levels of statistical significance. Statistical interactions were also largely uninterpretable. Strength-of-association values indicated that only a relatively small portion of the total variance was being accounted for by statistically significant effects. Some attempt was made to understand the otherwise stable and predictable behavior of the Sternberg metric which became erratic in the presence of the primary flight-control task. Possible explanations include the following:

1. Task loading was high on the primary flight-control task. This finding was supported by CHPR data, pilot responses to the questionnaires, and by direct observation of pilot failures to respond to the secondary task. It is likely that pilot responses to the secondary task were inhibited by insufficient reserve capacity owing to excessive demands of the primary task, at least during portions of the mission segments.

2. Insufficient numbers of observations were collected to ensure reliable measures of central tendency for RT's across all experimental variables — a continuing problem in the FHQ test environment in which a relatively small number of highly trained test pilots are typically used. Further decreasing the ISI on the secondary task, or expanding the length of mission segments to obtain a sufficient number of responses, would only tend to degrade performance on the primary task and would likely be unrealistic.

3. Responses to the postflight questionnaire provided some indication of conflict or overload occurring in the visual input channel when the secondary task was presented concurrently with the primary task. Vidulich and Wickens (ref. 16) reported that the Sternberg task was performed most poorly under S-C-R incompatible

visual-manual conditions, and best under compatible auditory-speech conditions for both single and dual task conditions. Experimental methods employing an auditory-speech Sternberg task may be more compatible with a heavily demanding visual-manual primary flight-control task.

4. One final factor that merits consideration is the response state used by the pilot subjects during this investigation. Damos (ref. 30) and others have pointed out that strategy is a major determinant of dual-task performance. Chiles (ref. 31) stated that the priority an operator assigns to a task is an important factor in determining the level of performance maintained on that task as other duties are added. Despite careful procedural and experimental controls exercised within the constraints of this investigation, differences in subject strategies and priorities were undoubtedly operative during this study and accounted for at least some of the variability in RT data. It is important to remember that the test pilot's primary job is the FHQ evaluation; it is difficult to modify a learned set of priorities through a limited sequence of instructions and practice trials. The Sternberg task may actually have been relegated to a tertiary position by test pilots as they concentrated on formulating responses to the CHPR scale. It would be interesting and informative to investigate the sequence of events leading to formulation of CHPR's in context of a secondary task.

It is both relevant and informative that CHPR data also failed to provide reliable statistical discrimination between the core (RCAH) treatment conditions in this investigation. Without quantitative performance or rating data, the FHQ researcher is forced to rely primarily on the pilots' subjective comments and observations regarding flight-control system evaluations. Regardless of the fact that test pilots are trained observers, it is evident that system evaluations can be affected by the number of observers, their background and experience, and other, hard-to-control, intervening factors. The negative findings obtained from this investigation do not obviate the need for continuing research to identify and develop sensitive workload metrics for FHQ investigations, but rather highlight the complex problems and difficulties surrounding this type of work. It is apparent from the current investigation that secondary tasks superimposed upon demanding primary flight-control tasks are likely to yield inconclusive results. Further research efforts are likely to be more productive and acceptable to the FHQ community at large, if an embedded task structure can be defined within the demands of the primary flight-control task. In addition to avoiding the use of additive or more demanding task structures, potential workload metrics should be as unobtrusive as possible from the pilot's viewpoint.

APPENDIX A

POSTFLIGHT AND FINAL DEBRIEFING QUESTIONNAIRE

POSTFLIGHT QUESTIONNAIRE

Following each set of simulation runs with a particular flight-control configuration, each subject was required to complete the following postflight questionnaire. Questionnaire data from the core (RCAH) experiment employing the rate-command/attitude-hold stability augmentation model were summed across the four test subjects, and means and standard deviations were computed for questions 1 through 8, based on a seven-point scale. Question 9 consisted of a 4-point priority ranking.

TABLE 10.- RESULTS OF POSTFLIGHT QUESTIONNAIRE FOR FOUR TEST SUBJECTS; MEANS AND STANDARD DEVIATIONS (SAMPLING DISTRIBUTION) ARE GIVEN FOR EACH QUESTION ON THE CORE (RCAH) EXPERIMENTAL SESSIONS

Question	Flight control configuration					
	Conventional		Three axis		Four axis	
1	4.875	0.545	5.625	0.415	5.625	0.415
2A	2.000	1.732	2.000	1.225	2.000	1.225
2B	1.250	.433	2.250	1.639	2.375	1.556
3	4.000	1.225	5.250	1.299	4.500	.500
4	3.500	.500	3.500	.866	3.250	.433
5	5.875	.217	5.625	.415	5.500	.354
6	3.625	.960	3.875	.893	3.750	1.479
7	2.500	1.500	1.750	.829	3.000	1.000
8A	5.500	.866	5.000	.707	4.250	1.090
8B	3.750	1.299	3.000	1.225	2.750	.829
8C	3.500	1.118	4.000	1.581	3.750	1.479
8D	2.500	.500	2.750	.829	2.500	.500
9A	1.500	.866	1.500	.866	1.500	.866
9B	2.000	.707	2.500	.500	2.500	.500
9C	2.750	.829	2.750	1.299	2.250	1.090
9D	3.750	.433	3.250	.829	3.750	.433

S No. ________

Date ________

Cond ________

POSTFLIGHT QUESTIONNAIRE

Please answer the following questions as accurately as you can recall. A response is made by placing a "/" through the appropriate number on the scaled line following each question.

1. The effect of the secondary memory task on your flight control task was:

Helpful — Harmful

1 2 3 4 5 6 7

2. Indicate the amount of interference that the display mode select control had upon:

a. Your performance on the primary flight control task.

None — Much

1 2 3 4 5 6 7

b. Your performance on the secondary item-recognition task.

None — Much

1 2 3 4 5 6 7

3. Were you adequately trained on the current primary flight control system?

Overtrained — Undertrained

1 2 3 4 5 6 7

The remaining questions pertain to the secondary item-recognition task.

4. Increasing the number of letters in the memory set made by performance:

Worse — Better

1 2 3 4 5 6 7

5. The display of these secondary letters on my primary flight display made my scan pattern:

Easier — Harder

1 2 3 4 5 6 7

6. The location of these letters on my primary flight display was:

Poor — Excellent

1 2 3 4 5 6 7

7. How many false identifications do you recall making while responding to the secondary memory task?

None — Many

1 2 3 4 5 6 7

8. Rate the amount of interference that the secondary task contributed to your overall flight performance on each of the following four mission segments:

   a. Six-degree Descent

   None — Extreme

   1 2 3 4 5 6 7

   b. NOE Maneuvering

   None — Extreme

   1 2 3 4 5 6 7

   c. Hover and Bob-up Maneuvers

   None — Extreme

   1 2 3 4 5 6 7

   d. Straight-and-Level Flight

   None — Extreme

   1 2 3 4 5 6 7

9. Rank order the following four mission segments in terms of difficulty of execution.

(a) Six-degree Descent (b) NOE Maneuvering
(c) Hover and Bob-up (d) Straight and Level

____ Most difficult
____
____
____ Easiest

FINAL DEBRIEFING QUESTIONNAIRE

In addition to the postflight questionnaires, subjects were required to complete a final debriefing questionnaire following the simulation. Data from this questionnaire were summed across the four subjects, and means and standard deviations were computed, based on a seven-point rating scale.

TABLE 11.- RESULTS OF FINAL DEBRIEFING QUESTIONNAIRE FOR FOUR TEST SUBJECTS; STANDARD DEVIATIONS WERE COMPUTED BASED ON THE SAMPLING DISTRIBUTION N = 4.

Question	Mean	Standard deviation
1A	4.000	0.000
1B	4.000	.000
1C	3.625	.650
1D	3.500	1.118
2	4.500	1.658
3	4.250	.433
4	3.625	.960
5	3.750	1.090

S No. ________

Date ________

Cond. ________

FINAL DEBRIEFING QUESTIONNAIRE

Please answer the following questions as accurately and honestly as you can. Your responses are needed to help in refining and modifying the secondary item-recognition task for application to other future research. Space is provided for additional comments and recommendations if you so desire.

1. Rate the overall adequacy of training you received in preparing for:

a. The Secondary Item-Recognition Task

Inadequate — Excessive

1 2 3 4 5 6 7

b. Flight Profile Familiarization for the Four Mission Segments

Insufficient — Excessive

1 2 3 4 5 6 7

c. Different Primary Flight Control Configurations

Inadequate — Excessive

1 2 3 4 5 6 7

d. Stability and Control Augmentation System Variations

Insufficient — Excessive

1 2 3 4 5 6 7

Please indicate how you feel with regard to the following statements:

2. I became bored flying the same mission profile over and over again.

Strongly Agree — Strongly Disagree

1 2 3 4 5 6 7

3. There were too many simulation runs.

Strongly Agree — Strongly Disagree

1 2 3 4 5 6 7

4. The secondary task quickly became very annoying and distracting.

Strongly Agree — Strongly Disagree

1 2 3 4 5 6 7

5. I became physically fatigued due to the excessive length of each simulation run.

Strongly Agree — Strongly Disagree

1 2 3 4 5 6 7

APPENDIX B

SUBJECTS' INSTRUCTIONS

The simulation in which you are about to participate is a joint Army/NASA investigation conducted in support of the Army's Advanced Digital/Optical Control System (ADOCS) program. The purpose of this work is to examine the relative adequacy of three different primary flight-control configurations: (1) a conventional helicopter control configuration; (2) a four-axis side-arm controller; and (3) a hybrid control system consisting of a three-axis controller plus collective. You will be flying common mission scenarios throughout this simulation study; however, the primary flight-control configurations will be varied, along with the stability and control augmentation system (SCAS) levels to help us to arrive at optimized controller/SCAS combinations. Each flight scenario will require about 8 min for you to complete and consists of four 2-min mission segments, including (1) a constant-rate descent into (2) NOE maneuvering, and (3) a hover and bob-up segment, and (4) straight-and-level flight.

In addition to Cooper-Harper rating and selected questions following each run, we will be introducing a second task during many of the simulation runs. This secondary task is quite simply a choice-reaction time task. We will be measuring the amount of time it takes you to respond orally to a letter after it appears on your primary flight display. Specifically, 15-sec before each experimental run, one, two, or four letters will be presented to you on your primary display. You must remember these letters for the next 8-min simulation run. To accomplish this, you should say the letters over and over to yourself until the run is complete. At random time intervals during the run, a letter will appear inside the box just slightly left of center on your display. If this letter is one of the letters that you memorized before the run, you must respond quickly with an oral yes. If the letter in the box is not one of the memorized letters, respond quickly with an oral no. It is important that your yes and no responses be clear and consistent from run to run since we are employing automatic voice recognition techniques.

The purpose of the secondary task is to examine your reserve capacity above and beyond the demands of the primary flight-control task. Of the two tasks, let me emphasize that the flight-control task is primary. By this I mean that you should attempt to maintain as good a level of performance on the flight-control task with the secondary task, as you did without it. You should nonetheless try to respond as rapidly and as accurately as possible to the secondary task. The usefulness of the secondary task will depend on how diligently you perform the task.

Two final points about the secondary task. First, you should respond as fast as you can after making the correct yes-no decision. This means don't guess or anticipate what the stimulus letter will be. Second, don't let your motivation lag on the secondary task during the 8-min simulation runs. Work hard to make your reaction times fast during each run; relax between runs. Slow or inconsistent response times will mask the results that we wish to investigate.

We are now ready to provide you with a few familiarization runs on the secondary task. Following familiarization, we'll move right into the investigation and collect baseline reaction-time data on the one-, two-, and four-element letter sets. Remember to respond as rapidly as possible.

(FAMILIARIZATION AND BASELINE DATA COLLECTION)

We are now ready to begin the simulation runs. Each controller/SCAS combination will remain fixed throughout any given simulation period. Following a 4-min familiarization run with each new configuration, we will move right into data collection with that control configuration, but will conduct the first simulation run without the secondary task. The next three runs will essentially duplicate your first run, but will include the secondary task with one-, two-, or four-element letter sets. One final run will be made without the secondary task to validate your baseline flight performance level. Remember to maintain the best level of performance you can on the primary flight-control task.

Following each run you will be asked for Cooper-Harper rating on each of the four flight segments which include descent, maneuvering, hover and bob-up, and straight-and-level flight. You will also be asked to complete a short questionnaire on selected aspects of the primary and secondary tasks before progressing onto each new control configuration. Good luck.

REFERENCES

1. Mudd, S. A.: The Treatment of Handling-Qualities Rating Data. Human Factors, vol. 11, no. 4, 1969, pp. 321-330.

2. McDonnell, J. D.: An Application of Measurement Methods to Improve the Quantitative Nature of Pilot Rating Scales. IEEE Transactions on Man-Machine Systems, vol. MMS-10, no. 3, Sept. 1969.

3. Donders, F. G.: On the Speed of Mental Processes. Acta Psychologica 30, Attention and Perception II, W. G. Koster, ed., North Holland Publishing Company, Amsterdam, 1969, pp. 412-4341.

4. Koster, F. G.: Attention and Performance II. North-Holland Publishing Co., Amsterdam, 1969.

5. Sternberg, Saul: High-Speed Scanning in Human Memory. Science, vol. 153, Aug. 1966, pp. 652-654.

6. Sternberg, Saul: Memory Scanning: New Findings and Current Controversies. Quarterly Journal of Experimental Psychology, vol. 27, 1975, pp. 1-31.

7. Sternberg, Saul: The Discovery of Processing Stages: Extension of Donder's Method. Acta Psychologica 30, Attention and Performance II, W. G. Koster, ed., North-Holland Publishing Company, Amsterdam, 1969, pp. 276-315.

8. Wickens, C. D.: The Structure of Attentional Resources. Acta Psychologica 30, Attention and Performance VIII, Raymond S. Nickerson, ed., Lawrence Erlbaum Associates, Publishers, Hillsdale, N.J., 1978, pp. 239-257.

9. Knowles, W. B.: Operator Loading Tasks. Human Factors, vol. 5, 1968, pp. 155-161.

10. Knowles, W. G.; and Rose, D. J.: Manned Lunar Landing Simulation. Proceedings of the IEEE National Winter Convention on Military Electronics, Los Angeles, Calif., Jan. 1963.

11. Wickens, C. D.: Processing Resources in Attention and Workload. TR-EPL-81-3/ONR-81-3, Office of Naval Research, Arlington, Va., July 1981.

12. Micalizzi, J.; and Wickens, C. D.: The Application of Additive Factor Methodology to Workload Assessment in a Dynamic System Monitoring Task. TR-EPL-80-2/ONR-80-2, Office of Naval Research, Arlington, Va., Dec. 1980.

13. Wickens, C. D.; Derrick, W.; Berringer, D.; and Micalizzi, J.: The Structure of Processing Resources: Implications for Task Configuration and Workload. Proceedings of the Human Factors Society, 24th Annual Meeting, Los Angeles, Calif., Oct. 1980.

14. Wickens, C. D.; Sandry, D.; and Micalizzi, J.: A Validation of the Spatial Variant of the Sternberg Memory Search Task: Search Rate, Response Hand, and Task Interference. TR-EPL-81-2/ONR-81-2, Office of Naval Research, Arlington, Va., Mar. 1981.

15. Wickens, C. D.; and Derrick, W.: The Processing Demands of Higher Order Manual Control: Application of Additive Factors Methodology. TR-EPL-81-1/ONR-81-1, Office of Naval Research, Arlington, Va., Mar. 1981.

16. Vidulich, M.; and Wickens, C. D.: Time-Sharing Manual Control and Memory Search: The Joint Effects of Input and Output Modality Competition, Priorities, and Control Order. TR-EPL-81-4/ONR-81-4, Office of Naval Research, Arlington, Va., July 1981.

17. Wickens, C. D.; Sandry, D.; and Vidulich, M.: Compatibility and Resource Competition between Modalities of Input, Central Processing, and Output. Human Factors, vol. 25, no. 2, 1983, pp. 227-248.

18. Crawford, B. M.; Hoffman, M. S.; and Pearson, W. H.: Multipurpose Digital Switching and Flight Control Workload. TR-78-43, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, Dec. 1978.

19. Corrick, G. E.: Missile Launch Envelope Study. TR-80-136, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, June 1981.

20. Johnson, Richard M.: Target Information Processing: The Effects on Reaction Time of Terrain, Downlook Angle, and Response Processing Level. Technical Report 480, Research Institute for the Behavioral and Social Sciences, Army Research Institute, Alexandria, Va., Oct. 1980.

21. Schiflett, S. G.; Linton, P. M.; and Spicuzza, R. J.: Evaluation of a Pilot Workload Assessment Device to Test Alternate Display Formats and Control Handling Qualities. NATO AGARD Conference Proceedings No. 312, Stuttgart, Germany, May 1981.

22. Aiken, E. W.; Blanken, C. L.; and Hemingway, J. C.: A Manned Simulator Investigation of the Effects of an Integrated Isometric Controller on Pilot Workload for Helicopter Nap-of-the-Earth Flight. Presented at the 17th Annual Conference on Manual Control, 15 Oct. 1981, JPL Publication 81-95, pp. 237-239.

23. Schiflett, S. G.: Evaluation of a Pilot Workload Assessment Device to Test Alternate Display Formats and Control Handling Qualities. Report No. SY-33R-80, Naval Air Test Center, NATC, Patuxent River, Md., July 1980.

24. Chen, R. T. N.: A Simplified Rotor System Mathematical Model for Piloted Flight Dynamics Simulation. NASA TM-78575, 1979.

25. Aiken, E. W.: Simulator Investigations of Various Side-Stick Controller/ Stability and Control Augmentation Systems for Helicopter Terrain Flight. AIAA Paper 82-1522, San Diego, Calif., 1982.

26. Aiken, E. W.; and Merrill, R. K.: Results of a Simulator Investigation of Control System and Display Variations for an Attack Helicopter Mission. Presented at the 36th Annual National Forum of the American Helicopter Society, Washington, D.C., May 1980.

27. Dixon, W. J.; et al.: Biomedical Computer Programs P-Series. U. of California Press, Los Angeles, Calif., 1981.

28. Hays, W. L.: Statistics for Psychologists. Holt, Rinehart, and Winston, Inc., New York, 1964.

29. Winer, B. J.: Statistical Principles in Experimental Design. McGraw-Hill, Inc., New York, N.Y., 1971.

30. Damos, D. L.: Development and Transfer of Timesharing Skills. TR-ARL-77-11/AFOSR-77-10, Aviation Research Laboratory, U. of Illinois, Urbana, Ill., 1977.

31. Chiles, W. D.: Objective Methods for Developing Indices of Pilot Workload. FAA-AM-77-15, FAA Civil Aeromedical Institute, Oklahoma City, Okla., July 1977.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-85884		

4. Title and Subtitle	5. Report Date
AN EXPERIMENTAL EVALUATION OF THE STERNBERG TASK AS A WORKLOAD METRIC FOR HELICOPTER FLIGHT HANDLING QUALITIES (FHQ) RESEARCH	March 1984
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
John C. Hemingway	A-9634
9. Performing Organization Name and Address	10. Work Unit No.
Ames Research Center Moffett Field, CA 94035	T-5415
	11. Contract or Grant No.
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Technical Memorandum
National Aeronautics and Space Administration Washington, DC 20546	14. Sponsoring Agency Code 505-42-11

15. Supplementary Notes

Point of Contact: John C. Hemingway, Ames Research Center, M/S 239-21, Moffett Field, CA 94035 (415) 965-5716 or FTS 448-5716

16. Abstract

The objective of this research was to determine whether the Sternberg item-recognition task, employed as a secondary task measure of spare mental capacity for flight handling qualities (FHQ) simulation research, could help to differentiate between different flight-control conditions. FHQ evaluations were conducted on the Vertical Motion Simulator at Ames Research Center to investigate different primary flight-control configurations, and selected stability and control augmentation levels for helicopters engaged in low-level flight regimes. The Sternberg task was superimposed upon the primary flight-control task in a balanced experimental design. The results of parametric statistical analysis of Sternberg secondary task data failed to support the continued use of this task as a measure of pilot workload. In addition to the secondary task, subjects provided Cooper-Harper pilot ratings (CHPR) and responded to a workload questionnaire. The CHPR data also failed to provide reliable statistical discrimination between FHQ treatment conditions; some insight into the behavior of the secondary task was gained from the workload questionnaire data. A limited review of the literature on the use of the Sternberg task as a workload metric is also provided.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Helicopter flight handling qualities Workload Sternberg task Sidearm flight controls Cooper-Harper pilot ratings	Unlimited Subject Category - 54

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	36	A03

*For sale by the National Technical Information Service, Springfield, Virginia 22161