AFHRL-TR-80-27

AIR FORCE HUMAN RESOURCES LABORATORY

AD A094317

DDC FILE COPY

PERCEPTUAL-MOTOR AND COGNITIVE PERFORMANCE TASK BATTERY FOR PILOT SELECTION

By

David L. Imhoff
Jerrold M. Levine
Advanced Research Resources Organization
4330 East-West Highway
Washington, D.C. 20014

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235

January 1981

Final Report

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

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

This final report was submitted by the Advanced Research Resources Organization, 4330 East-West Highway, Washington, D.C. 20014, under Contract F33615-79-C-0004, Project 7719, with the Manpower and Personnel Division, Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235. Mr. David R. Hunter was the Contractor Monitor for the Laboratory.

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

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

NANCY GUINN, Technical Director
Manpower and Personnel Division

RONALD W. TERRY, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFHRL-TR-80-27	2. GOVT ACCESSION NO.: AD-A094 317
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): PERCEPTUAL-MOTOR AND COGNITIVE PERFORMANCE TASK BATTERY FOR PILOT SELECTION	5. TYPE OF REPORT & PERIOD COVERED: Final
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): David L. Imhoff, Jerrold M. Levine	8. CONTRACT OR GRANT NUMBER(s): F33615-79-C-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Advanced Research Resources Organization, 1330 East-West Highway, Washington, D.C. 20011	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 02703F, 77191518
11. CONTROLLING OFFICE NAME AND ADDRESS: HQ Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235	12. REPORT DATE: January 1981
	13. NUMBER OF PAGES: 160
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Manpower and Personnel Division, Air Force Human Resources Laboratory, Brooks Air Force Base, Texas 78235	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

information processing
perceptual tests
performance battery
pilot selection
psychomotor tests
selection tests

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

A review of the literature on pilot selection and training, perceptual-motor processes, and cognitive processes was conducted. The objectives of this review were: (a) to identify perceptual-motor and cognitive tasks that demonstrated reliable individual differences in performance and (b) to identify perceptual-motor abilities and cognitive processes of demonstrated importance to successful piloting behavior. From this review, a set of tasks were identified that tapped the abilities and processes important to piloting and that showed evidence of producing reliable individual differences in performance. A conceptual framework makes explicit the link between the tasks selected and the requirements for successful pilot performance. This resulted in a large number of tasks which were candidates for inclusion in a pilot selection task battery. Psychometric and pragmatic criteria were applied to

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

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

392930

Item 20 Continued:

the tasks in the candidate pool, resulting in the identification of 15 tasks for inclusion in the final task battery. These tasks span the perceptual-motor and cognitive domains, with special emphasis on attentional and decision-making performance. The paradigms for the selected tasks are specified in detail to allow for the development and implementation of the tasks on a computer system, and a number of implications for validation of the battery are provided. The tasks included in the final battery are Perceptual Speed, Complex Coordination, Compensatory Tracking, Kinesthetic Sensitivity, Route Walking, Selective Attention, Time Sharing, Encoding Speed, Mental Rotation, Item Recognition, Immediate/Delayed Memory, Decision Making Speed, Probability Estimation, Risk Taking, and Embedded Figures.

SUMMARY

Objectives

An effort was carried out to develop a valid battery of tasks for the selection of candidates to Undergraduate Pilot Training. The battery had to satisfy several criteria, including: (a) explicit relation to the perceptual-motor and cognitive aspects of flying, (b) use of performance tasks rather than paper-and-pencil tests to avoid verbal and cultural biases, (c) evidence that the tasks reliably indicate individual differences in performance, and (d) capability of being implemented on a computerized testing station. The current effort was intended to update previously developed batteries which concentrated primarily upon motor skill performance and to broaden the conceptual basis of selection.

Approach

An extensive literature screening was conducted to identify candidate tasks fitting the above-stated requirements. Simultaneously, task-analytic data on piloting were reviewed and a conceptual framework developed to ensure the relevance of the selected tasks to important aspects of piloting. A set of psychometric and pragmatic criteria were then applied iteratively to the candidate tasks to select a final task battery of manageable size.

Specifics

A three-phase search of the literature was conducted on the following topics: (a) pilot selection and training, (b) psychomotor and perceptual processes, and (c) cognitive processes and individual differences. The products of these searches were screened by abstract and relevant documents obtained. The computer searches were augmented through reference texts and searches of relevant journals. In all, over 900 relevant articles and books were evaluated for the review.

A separate search was conducted for task analytic data on piloting. The results of this search were used to construct a conceptual framework relating various piloting tasks to their underlying perceptual-motor abilities and cognitive processes. The tasks eventually selected for the battery were required to tap those abilities and processes, and hence, the relevance of the task battery to piloting was ensured.

The goal of the review was to identify tasks which (a) had demonstrated validity as pilot selection devices or (b) reliably indicated individual differences in those abilities or processes demonstrated to be relevant to flying. A large number of tasks were selected as candidates for the task battery. Further criteria such as feasibility of implementation, sensitivity, motivational interest, independence, construct validity, and freedom from verbal bias were then used to further screen the pool of candidate tasks and to select a final battery of 15 tasks with apparent potential as pilot-selection devices. The paradigms for these tasks were then described in detail to allow implementation on a computerized testing station.

Conclusions

The final task battery represents a unique combination of perceptual-motor and attention tasks having demonstrated validity in pilot selection with tasks from the cognitive domain which tap previously untested but clearly important processes such as decision making. Because of the unique nature of the battery, extreme care must be used in the implementation, preliminary testing, and validation of the task battery. The task included in the final battery are Perceptual Speed, Complex Coordination,

Compensatory Tracking, Kinesthetic Memory, Route Walking, Selective Attention, Time Sharing, Encoding Speed, Mental Rotation, Item Recognition, Immediate/Delayed Memory, Decision Making Speed, Probability Estimation, Risk Taking, and Embedded Figures.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

The Air Force has a continuing need to improve procedures for pilot selection. While there is a long history of research on test development for pilot selection (see Fleishman, 1956; Melton, 1947; Passey & McLaurin, 1966), the increased complexity of aircraft and the changing requirements of the pilot's job emphasize the need to update these procedures. In particular, there is a need for selection test development which draws on the state-of-the-art information in abilities and processes underlying proficiency in piloting current and planned operational aircraft.

Recent efforts by Hunter (1975) and Hunter, Maurelli, and Thompson (1977) recognized, correctly, that the abilities predictive of pilot performance may tap several aptitude domains. Furthermore, there are recent developments in the measurement of individual differences that have not heretofore been adequately considered as potential predictors. Both of these lines of research clearly suggest that information processing abilities must be considered in conjunction with the more thoroughly explored perceptual and psychomotor ability domains, in order to more accurately predict pilot performance and to improve procedures for pilot selection.

This research effort reviews the literature bearing on these issues, in order to identify measures potentially valid for pilot selection. A subset of the identified measures is then selected, on the basis of specified criteria, for inclusion in a broad-based representative battery of performance tests, which may be evaluated as an instrument for the selection of pilot training candidates.

Some Working Definitions

Our interest is focused on psychomotor, perceptual, and information processing "abilities" of potential relevance to piloting. We have adopted an eclectic approach, which aims toward the identification of both abilities in the more traditional "aptitude" sense, and behavioral functions which may prove useful for the selection of pilot trainees. This outlook has guided our search of the literature and ensured its comprehensiveness.

Although there is some overlap, psychomotor abilities generally refer to coordinative, manipulative, repetitive, and/or precise body or limb movements. These abilities include multilimb coordination, wrist-finger speed, finger dexterity, manual dexterity, arm-hand steadiness, control precision, and rate control (timing) (see Fleishman, 1964, 1975). Perceptual abilities, on the other hand, involve attending to, perceiving, and evaluating sensory information in the environment. These abilities include flexibility of closure, speed of closure, perceptual speed, spatial orientation, and visualization. Reaction time can be viewed as properly belonging to either the perceptual or the psychomotor domain. The above definitions and classification of basic abilities into psychomotor and perceptual

domains are due in large measure to the work of Fleishman and his associates (see Fleishman, 1975). For the purpose of this report, psychomotor and perceptual performance domains will be combined into a single perceptual-motor category.

Information-processing abilities are more difficult to stipulate, and regardless of the definition adopted, it is clear that there will be a large overlap with perceptual abilities (e.g., Is time-sharing a perceptual or information-processing ability?). Further, we see very little difference between information processing and cognitive processes. We would argue that the latter is a broader, more inclusive label. Nevertheless, the precise definition of information processing continues to be elusive. There is no generally accepted taxonomy of either abilities or behavioral processes, which serves adequately to delimit the area, and it is beyond the scope of this research to accomplish such an organization.

Cognitive psychologists would unquestionably consider attentional, memorial, decision-making, and linguistic processes as among the most general and important for study. This list is by no means exhaustive. Furthermore, each of these areas can be broken down into "component" preocesses (e.g., short- and long-term memory comprise the memorial process). The classification problem is difficult. For the purpose of guiding this research, information processing can be defined in terms of "operations," which Carroll (1974) calls control processes that are explicitly specified or implied in task instructions and that must be performed if the task is to be successfully completed.

Encoding: the operation by which information is input to the system, including the initial set of processes that converts the physical stimulus to a form which is "appropriate" for the task. Different task demands may require different levels of analysis of the stimulus. Posner (1969) has called this dimension "abstraction"--the process by which different types of information about the stimulus are extracted; in other words, the level of stimulus analysis demanded by the task. For example, a visual search task might require only that the subject extract physical or structural information about the stimulus, a memory search task might require the extraction of name information, and a semantic search task might necessitate semantic or "meaning" information.

Construction: the operation by which new information structures are generated from information already in the system. This is what Neisser (1967) and others have called "synthesis"; in the present context, we will limit the use to situations where additional features of the stimuli must be abstracted, beyond those initially encoded.

Transforming: the operation by which a given information structure is converted into an equivalent structure necessary for task performance. In contrast to constructing, transforming does not involve any new information abstraction; rather, this operation requires the application of some stored rules to the information structure already present.

Storing: the operation by which new information is incorporated into existing information structures, while its entire content is retained.

Retrieving: the operation by which previously stored information is made available to the processing system.

Searching: the operation by which an information structure is examined for the presence or absence of one or more properties. The information structure examined may be one already in the processing system or one external to it (e.g., a visual array).

Comparing: the operation which which two information structures (again, either internal or external to the processing system) are judged to be the same or different. The information structures need not both be physical entities (as in the comparison of two objects); likewise, a physical entity can be compared to a stored representation or description in order to determine identity.

Responding: the operation by which the appropriate (motor) action is selected and executed. In many information-processing investigations, the response operation is itself the object of study.

It is important to emphasize that the adoption of this definition of information processing is not intended to be restrictive, but rather to serve as a focus of our efforts to carry out a comprehensive review of the literature.

Organization of the Report

This report is designed to serve two primary functions: (a) to provide a comprehensive review of the literature on perceptual-motor abilities and cognitive processes of potential relevance to the task of piloting, and (b) to specify and delineate a set of candidate tasks that can be justified as possessing utility as pilot selection devices when incorporated in a test battery. The report is organized in the following manner: Chapter 2 reviews the literature on perceptual-motor processes, including (a) previous research on the use of perceptual-motor tasks as pilot selection devices, and (b) perceptual-motor research that has not been applied specifically to pilot selection, but appears to be of relevance to the functions required for piloting. The emphasis in this latter portion of the literature is on recent (post-1966) perceptual-motor research. Chapter 3 reviews the literature on cognitive processes in a manner paralleling Chapter 2; that is, one segment reviews the use of cognitive information processing tasks specifically for pilot selection, while the next segment reviews research on a variety of cognitive processes that are apparently relevant to flying performance. The emphasis in this latter segment is on the identification of tasks which tap processes relevant to flying and, in addition, reveal individual differences in performance which may be reliably assessed. These include information processing and other types of cognitive tasks, including decision making.

The remainder of the report is devoted to the specification of candidate tasks which emerge from the review, selecting the best tasks based on relevance and other logistical criteria, and specifying in detail the paradigms for those tasks. Chapter 4 presents brief descriptions of a large number of candidate tasks derived from the review. Chapter 5 discusses the application of the criterion of relevance, as well as criteria such as sensitivity, feasibility, etc., to the pool of candidate tasks, resulting in the selection of a subset of the tasks with the greatest potential utility for pilot selection. Chapter 6 consists of a detailed description of the paradigms for the selected subset of candidate tasks, so that these tasks might be readily implemented in a test battery. Chapter 7 presents an overview of the report and the conclusions drawn from the effort.

Search Strategy

In order to ensure a thorough and extensive review of the literature, three separate searches were conducted. One search examined research on perceptual-motor or psychomotor processes, a second examined research showing individual differences in cognitive processes, and the third examined pilot selection and pilot training literature. In each search, several tactics were employed to ensure its comprehensiveness. The Lockheed Dialog system was used to access the computerized data bases of NTIS and/or Psychological Abstracts, and the abstracts of potentially relevant papers were obtained for further screening. Secondary sources pertaining to the search areas were examined for their own relevance, as well as to confirm the representativeness of the search. Current issues of journals in which potentially relevant material might be published were examined. In all, more than 900 articles and books were screened and evaluated for the review.

A conceptual framework describing the abilities and processes required in piloting was developed iteratively with the search process. As the search provided more and more information about piloting which could be incorporated into the framework, the framework was used to decide whether the research studies being screened were relevant to piloting. This framework is presented in its final format in Appendix A.

Although the literature search employed the three-pronged strategy described, much of the research uncovered concerning pilot selection and training involved perceptual-motor and cognitive processes and tasks. However, that portion of the search also revealed an extensive literature that, while somewhat tangential to the main concerns of this report, is nonetheless relevant to pilot selection. A brief discussion of this literature sheds some light on the difficulties involved in pilot selection.

Pilot Selection and Training

Many types of tests have been employed in an effort to improve pilot selection, especially tests of the pencil-and-paper variety.

For example, DuBois (1947) reported reasonable validity for the Aircrew Classification Battery, and a wide variety of aptitude measures have been investigated (Berkshire & Ambler, 1963), although with mixed results. In addition a wide variety of tests for motivation and personality and biographical data and social factors have been investigated as possible selection devices (Youngling, Levine, Mocharnuk, & Weston, 1977, for a summary). One prominent finding in these investigations is that many of the predictors are inadequate when considered singly, but demonstrate significant predictive power when combined into test batteries (Berkshire & Ambler, 1963).

The eclectic nature of the process of pilot selection is apparent in civilian airline selection procedures, even though the candidates are typically experienced pilots already. For example, Mensh (1970) summarized the psychological selection procedures used by some European civilian airlines for pilot selection. These methods typically include screening by general factors such as age, physical condition, and general intelligence, followed by a mixed battery of interviews, personality assessment, and specific motor and information processing tests. One European airline examines biographical data, uses a Phase I screening based on age, sport activities, general impression, command of English, arithmetic reasoning, navigation and mathematics, meteorology, instrument tests, mechanical information, technical understanding, flight techniques, physics, and simulator performance. Successful candidates are further screened in Phase II personality tests, logical thinking, short-term memory, spatial orientation, concentration and avoidance of distraction, tempo of observation (Bourdon test), medical and physical status, motivation, simple reaction time, and reaction time in two-hand coordination and rudder-control performance. This extensive screening is fairly successful in eliminating potential failures, but it is difficult or impossible to judge which portions of the screening result in this effectiveness. No attempt is made to relate the tests to specific aspects or parts of the pilot's job.

The effectiveness of individual portions of a large selection battery, such as that employed by the airline mentioned in the preceding paragraph, is a matter for investigation. Background and interest measures are usually included in the screening procedure, and Guinn, Vitola, and Leisey (1976) have examined specific background and interest measures for validity in predicting success in Undergraduate Pilot Training (UPT). They administered the Strong Vocational Interest Blank and the Officer Biographical and Attitudinal Survey to 593 undergraduate trainees and correlated the subsections of these tests with success in UPT. Then using the 17 most highly correlated professions from the Strong, and four eliminee keys from the Biographical Survey (as well as the Armed Forces Officer Qualification Test [AFOQT]), they attempted to predict success in UPT for another group. They found acceptable composite validity, but very low cross-validity when these three tests were used to predict elimination. When all three predictors were considered, 38% of eliminees were correctly identified, and 10% of those passing were

mis-predicted as eliminees. Using only the Biographical keys and the AFOQT resulted in a slightly higher identification of eliminees (45%) but much higher mis-prediction on students who passed (20%). The inclusion of the Strong professions as predictors would be warranted only if maximizing the number of pilots trained were the object, rather than minimizing the waste by minimizing the number of potential eliminees.

Baxter (1978) attempted to predict success in UPT on the basis of specific background data, in this case various records and scores obtained by students while in the Air Force Academy. Baxter used medical codes, college entrance exam scores, college grade point average, academic, leadership, and graduate order to merit, T-41 training status, and T-41 order to merit as predictors and found significant but very small relationships in a multiple correlation with UPT success. He concluded in general that these relationships were insufficient for prediction and selection, but pointed out the importance of T-41 training. Those without T-41 training were three times as likely to fail UPT as those with it, and T-41 order of merit scores averaged 53% for those who succeeded and 38% for those who failed UPT.

Other background and biographical data have been shown to be related to success in pilot training, including academic grades and officer rank (Youngling et al., 1977). One of the most important of these data is previous flying experience. For example, Bale and Ambler (1971) found a significant correlation between possession of a private pilot's license and success in UPT, and this fits well with the predictive value of T-41 training reported to Baxter (1978). In addition, background information is highly revealing concerning the motivation of potential trainees. Valentine has outlined in the Rand Symposium on pilot training (Stewart & Wainstein, 1970) a screening device based on interest in model airplanes, attitudes of family members, and individual aspirations, and other devices that test knowledge of aviation information, which reveals the potential trainee's motivation regarding aviation prior to training.

Factors other than those already mentioned may also be of importance in the piloting task, and one of the most relevant of these is resistance to air motion sickness. Ambler and Guedry (1965) developed a Brief Vestibular Disorientation Test to induce the conditions of motion sickness and tested it to see if it was predictive of success in UPT. The test involved a chair rotating at 15 rpm and required subjects to change their head position (left-upright-right-upright-forward-upright) every 30 seconds, for 5 minutes. Four judges rated the subjects in terms of pallor, sweating, facial expression, unsteadiness, rate of recovery, and overall performance. They found significant correlations between these ratings and failure for any reason ($r = .165$), failure due to tension or airsickness ($r = .272$), and failure due to airsickness ($r = .413$) for the 226 pilot trainees tested. Moreover, when the Brief Vestibular Disorientation Test was combined with other predictive criteria, predictive value of these criteria increased.

A large number of other tests for reaction to stress have been applied to the problem of pilot selection, but these have been mostly pencil-and-paper tests. Moreover, these tests have in general shown little predictive validity, and it appears that stressful situation tests like the Brief Vestibular Disorientation Test are much more appropriate and effective for pilot selection.

This introduction has been somewhat cursory, because these areas of research are not directly concerned with the perceptual-motor and cognitive tasks which are of central importance here. Nonetheless, the reports summarized here do give an indication of the complexity of pilot selection and the variety of variables with predictive potential. In the following sections, we shall consider many more variables central to perceptual-motor and cognitive performance and evaluate them with regard to their potential as pilot-selection devices.

PERCEPTUAL-MOTOR PROCESSES

The importance of the perceptual-motor skills to pilot performance has long been considered crucial, as is evident both in the extent to which pilot training has focused on perceptual-motor processes and in the predominance of perceptual-motor tests used to select candidates for pilot training. For example, pilot-training research has been concerned with the use of sensory cues, both as signaling devices (Hill, Gardiner, & Bliss, 1969) and as important carriers of information in their own right (Matheny, Lowes, Baker, & Bynum, 1971). As selection devices, perceptual-motor tests have been utilized in pencil-and-paper form (Guilford & Lacey, 1947), but the emphasis has been on performance task measures of perceptual-motor abilities (Melton, 1947; Fleishman, 1956; Fleishman & Ornstein, 1960). This and other relevant research have been summarized extensively (Passey & McLaurin, 1966; Fleishman, 1964).

While performance tests of perceptual-motor abilities added significantly to the predictive power of the basic written tests for pilot selection, they fell into disuse due to the unreliability and expense of the apparatuses. The rising cost of pilot training and the advent of reliable computerized testing stations have fueled a renewed interest in these potentially valuable selection tests (McGrevy & Valentine, 1974; Hunter, 1975; McLaurin, 1973).

This chapter reviews the role of perceptual-motor abilities in pilot training and selection research, then examines more general research on perceptual-motor abilities. Since the purpose of this examination is to identify tasks of potential value as pilot selection devices, many otherwise important areas of perceptual-motor research (e.g., physiological, animal, and sport and exercise research) are omitted on the basis that they are unlikely to yield such tasks. Moreover, the review is focused such that once sufficient evidence has been presented to justify or reject a task as a potential selection device, further research employing the same or similar task is not considered unless it concerns a different perceptual-motor ability.

Pilot Training and Selection

The current program of UPT is an intensive one, requiring nearly a year to complete and consisting of a preflight segment intended to acquaint the students with the basic flying procedures in the T-37 aircraft and to prepare them for the more intensive training to come. This preflight phase is followed by flying training in the T-37 and T-38 aircraft to train the students in basic flight maneuvers, instrument procedures, navigation procedures, and motion procedures. In addition the students receive intensive academic training acquainting them with their aircraft as well as basic flight physiology, weather, etc. Furthermore, pilot trainees undergo officer development training including physical conditioning, acquaintance with officer duties and responsibilities, and character development training. The training program uses classroom training and general

and specific flight simulators, as well as actual flying practice, to achieve its ends. (Syllabus of Instruction for Undergraduate Pilot Training, 1979).

A number of reviews of pilot training research have indicated the importance of various sensory cues in guiding pilot performances. For example, Smode, Hall, and Meyers (1966) reviewed research on pilot training, emphasizing the importance of visual aspects of flying and suggesting sensory training as a potentially fertile area of research. Matheny, Gray, and Waters (1975) have similarly emphasized the importance of sensory-perceptual factors for pilot training, showing concern for utilization of motion cues as well as visual cues. They suggested further research on visual displays and visual cue utilization, but also on motion cues and the interaction between visual and motion cues. Sinacori (1978) has pointed out the importance of human perception in simulation, comparing the capabilities of modern simulation devices to the capabilities of the human perceptual system. These overviews give strong indications that sensory cue utilization is a primary concern in pilot training research.

Research on the sensory cue utilization has been concerned both with the use of cues to provide performance feedback or warnings, and with the direct pickup and integration of information contained in the cues. An example of the first type of research was performed by Hill, Gardiner, and Bliss (1969). They were concerned with the potential of a tactile cuing system for improving flight performance. Using a flight simulator, they devised a system that provided vibration to the left or right arm of the student pilot when flying performance fell outside certain parameters. The vibration indicated not only flying error, but also the tactile cues were ignored in favor of the visual instrument display. However, in free flight, when students had to monitor an external display as well, the tactile cues improved performance. Moreover, the advantage provided by tactile cuing increased as the difficulty of the maneuver increased. They suggested that tactile cuing reduced processing load and would be especially helpful in high work load situations.

A number of devices and techniques have been developed for providing the pilot with visual cues reflecting flight performance. A recent example (Moroney, Pruitt, & Lau, 1979) involved the development of a helmet that provided the pilot with a graphic visual display of energy maneuverability--i.e., the ability to change direction, altitude, and/or airspeed indicated, in terms of energy and energy rate. Tests of the device showed that relatively inexperienced student pilots quickly performed as well on specified maneuvers as experienced pilots when they had been trained with this helmet-mounted display.

Examples of research on direct pickup and integration of information in sensory cues abound. For example, Matheny and his associates (Matheny, Lowes, Baker, & Bynum, 1971) attempted to specify important cues for piloting in the visual, auditory, motion, and

movement control domains, and to relate these cues to simulator training devices. For example, perception of depth is crucial to piloting, and both monocular and binocular visual cues contribute to depth perception. They recommended studies to determine the relative efficiency of monocular and binocular cues, on the hypothesis that monocular cues might be adequate for depth perception and therefore perspective geometry suitable as a descriptive method for identifying and quantifying depth cues for trainer application. They made similar proposals for other modality cues, as well as noting the importance of potential cue interaction. Their concern was with the adequacy and cost-effectiveness of cue implementation in ground training devices.

Young and his associates (Young, Curry, & Oman, 1977) were concerned with pilot determination of body position and tilt. While such determination is primarily dependent on vestibular and motion cues, visual cues are also important. They developed a technique of circularvection whereby a sense of tilting could be induced by a moving visual field. They noted that in actual flight simulation these visual cues regarding tilt are readily integrated with motion cues, unless cues conflict, in which case the perception of tilt seems to be dependent on the actual motion cues.

In addition to the research considered above, extensive work has been conducted on pre-training and the formation of perceptual and cognitive schemata to aid in the pickup of sensory information. This research is discussed in the next chapter.

Perceptual-motor tasks have also played an extremely widespread role as pilot selection devices. A great deal of early research on perceptual-motor apparatus tasks is compiled in a classic work by Melton (1947). He summarized research on the various apparatus tests employed by the Army Air Corps in World War II for the selection of various personnel. This work shows several of the apparatus tests to be reasonably valid in predicting success in flight-training school for pilots, including standard classification tasks such as complex coordination (CM 201), two-hand coordination (CM 101B), discrimination reaction time (CP 611D), rotary pursuit (CM 803B), rudder control (CM 120B), and other dexterity and steadiness tests. Moreover Melton discussed the experimental research being conducted on other apparatuses designed to test such functions as anticipatory timing reactions and kinesthetic discrimination.

In the post-World War II period of development of psychological tests for pilot selection, Fleishman (1956) employed a factor-analytic approach to the study of apparatus tests for the presence of various psychomotor abilities relevant to piloting. That is, the approach concentrated on using correlations between performance on different tasks to identify factors underlying variable piloting performance. Tests were then developed which were directly relevant to the abilities critical to piloting, rather than tests which were mere analogs of the pilot's job.

Fleishman and his associates demonstrated validity for a number of apparatus tests including complex coordination, two-hand coordination, rotary pursuit, rudder control, pursuit confusing, and direction control tests (1956). Moreover, Fleishman argued that the validity of these tests is not due to resemblance to the task of piloting (i.e., rotary pursuit does not resemble piloting) but that it taps an ability (control precision) which underlies piloting performance. They identified several such underlying psychomotor abilities (Fleishman, 1964) and argued that specification of the important abilities subserving pilot performance would allow construction of tests to measure those abilities more precisely, and hence to select pilots more efficiently. For example, they obtained scores made by student pilots on 24 standard maneuvers, and a factor analysis of these scores revealed six factors already identified in the laboratory; i.e., control precision, multilimb coordination, rate control, spatial relations, response orientation, and procedural integration (Fleishman, 1956). A later, more comprehensive analysis of such data also revealed a kinesthetic discrimination factor (Fleishman & Ornstein, 1960).

In view of the relatively good predictive validity evidenced by the various perceptual-motor and psychomotor apparatus tests, it seems surprising that their use as pilot selection devices was discontinued. However, as McGrevy and Valentine (1974) pointed out, the cost of operating the apparatus, on which only a single candidate could be tested at a time, and the progressive unreliability of the devices led to the discontinuation of apparatus tests. Improvements in solid-state technology have rendered many of the objections invalid, however, and the need to improve pilot-selecton efficiency and to reduce cost has stimulated a new interest in these tasks.

One result of the renewed interest in apparatus tests as selection devices has been an attempt to adapt tasks previously shown to be valid, for use in solid-state testing devices. For example, McGrevy and Valentine (1974) adapted the two-hand coordination and complex coordination tests employed by Melton (1947) and Fleishman (1956) to such a device. Performance on these tests was correlated with a variety of flight criteria, resulting in multiple correlations ranging from .18 to .20 for two-hand coordination, and from .43 to .60 for complex coordination, with the pass-fail criterion. While only some of the correlations were statistically significant (especially for complex coordination), they are considered relatively large in view of the restricted range of the sample.

Another result of the interest in perceptual-motor tasks as selection devices has been the tendency to develop and include some new tasks, especially for use in multiple-task selection batteries. For example, McLaurin (1973) included a test for kinesthetic learning in his test battery, requiring subjects to learn and perform a speeded movement pattern when blindfolded. A related kinesthetic memory task was employed by Hunter (1975), and the test proved to be an excellent predictor of success in navigation training (Hunter, Maurelli, & Thompson, 1977), as well as for success in training for enlisted men.

A test of perceptual speed employed by Hunter also appeared to be of particular value in the test battery.

Summary. Training research suggests the importance to piloting of sensitivity to various modalities of sensory input, as well as the speed and accuracy with which that input may be integrated and acted upon. Tasks such as perceptual speed, reaction time, and two-plate tapping appear relevant to the response portion, while kinesthetic memory appears relevant to the sensory sensitivity issue. Also of apparent relevance would be tests related to sensitivity to visual, auditory, and perhaps vestibular cues. Selection research has focused on the adaptation of previously employed perceptual-motor tests to current testing technology. Tests such as complex coordination, two-hand coordination, rudder control, pursuit tracking, and compensatory tracking appear relevant here. In addition, work such as that by Hunter et al. (1977) indicates the potential value of searching out and validating new perceptual-motor tasks for pilot selection.

Perceptual-Motor Research

The primary concerns of research in the perceptual-motor field have been the speed and accuracy with which movements may be executed, and the use of sensory information in controlling and modifying movement behavior. This segment of the review will consider first basic movement speed and accuracy, then perceptually guided movement control, and finally other perceptual-motor research. Within each subsection, a brief review of research prior to 1966 will be followed by a more in-depth consideration of recent research in the area.

Basic movement speed and accuracy. Early research on basic movement processes was largely concerned with the speed with which a movement could be completed, producing such common-sense findings as movement time increasing as a (logarithmic) function of distance, or the size of the movement (Brown & Slater-Hammel, 1949). However, it is easy to think of instances where this relationship does not hold. A word can be written on paper or on a blackboard with equal speed, although the movement distance is much greater on the blackboard. Fitts (1954) demonstrated that movement time depended on the required precision of the movement, as well as the distance traversed. Fitts and Peterson (1964) required subjects to move a stylus from a starting position to a target 3, 6, or 12 inches away, and varied the width of the target from .125, .25, .50, to 1.0 inches. They showed that distance and precision tend to complement each other; that is, if the distance is doubled and the target width also doubled, movement time remains the same.

The utilization of individual differences in basic movement processes has been evident in the work of Fleishman and his associates, who used factor analytic techniuqes to infer psychomotor abilities subserving performance on various tasks. For example, they find that performance on apparatus tasks such as the rotary pursuit task and complex coordination task are highly correlated; people who

do well on one tend to do well on the other, and vice versa (Fleishman, 1958; Fleishman & Hempel, 1956; Parker & Fleishman, 1960). They infer the existence of a general ability which they call control precision; the ability to make highly controlled and precise, but not over-controlled movements. Performance on tasks, such as the rotary pursuit task, then becomes an indication of the presence or absence of this control precision ability in a given individual. Moreover, control precision might be considered to underlie performance on other tasks, such as those envolved in piloting an airplane, and if so, performance on one control precision task like rotary pursuit should predict performance on the other task, piloting.

Using this same procedure, Fleishman and his associates also identified a basic movement ability they called speed of arm movement, and referred to the speed with which a gross, discrete arm movement could be made when accuracy is not a requirement (Fleishman, 1958; Fleishman & Hempel, 1954, 1955; Parker & Fleishman, 1960). Since accuracy is not required, the effects of movement control via sensory feedback are negligible. The ability has been shown to load heavily on a two-plate tapping task, which is a repetitive, patterned movement task. Such patterning of movement may be under open-loop control via a motor program. If so, speed of arm movement might represent a motor programming function, as well as a simple muscular response function. This suggestion is supported by the fact that the factor "speed of arm movement" contributes variance in complex tasks at high levels of proficiency and after extensive practice (Fleishman & Hempel, 1954, 1956).

A third factor that might be thought of as a component of basic movement processes has been identified as reaction time (RT), the speed with which an individual is able to respond to the onset of a stimulus, independent of the modality of the stimulus or the response required (Fleishman, 1954, 1958; Fleishman & Helpel, 1955; Parker & Fleishman, 1960). While this ability is crucial to the speed of overall response, it does not involve movement per se, and Fitts and Peterson (1964) have shown RT to be unaffected by the variables of distance and precision which control movement time. RT appears to be a measure of perceptual speed, of the time taken to detect a stimulus and initiate the neural commands for the response. RT abilities are best detected using a simple reaction time task, in which subjects keep their finger on the response button and press at the onset of the stimulus, thus avoiding confounding RT with movement time.

Some evidence suggests that RT also involves central processes in basic movement. For example, Fleishman and Hempel (1955) again show that RT makes its major contribution to individual differences on complex tasks at high levels of proficiency. Also, work on psychological refractory periods (Welford, 1952) shows that when a second stimulus follows a first too closely (less than 250 msec), RT to the first stimulus is lengthened. This is explained by assuming a single decision channel whch is occupied by the first stimulus. If the second occurs too quickly, the decision regarding initiation of response is disrupted and RT lengthened.

A great deal of more recent research in basic movement processes has concerned the source of control for those movements when peripheral feedback is not crucially involved. Among others, Keele (1968) has suggested that movements may be pre-programmed, by developing a sequence of centrally stored commands--a motor program--that are "structured before the movement begins and allows the entire sequence to be carried out uninfluenced by peripheral feedback" (Keele, 1968). Others have added to the notion of a motor program, suggesting that detection of movement error may take place by comparing the issued motor commands--efference--to a centrally stored standard of the movement to be executed (Laszlo & Manning, 1970). Both the notion of motor programming and that of closed-loop feedback at a central level strongly suggest the involvement of central processing in basic movement control.

Laszlo (1967) has been concerned to show that motor control and motor learning can take place in the absence of peripheral feedback. He had subjects trained to tap a light Morse key during nerve compression block, which eliminated kinesthetic feedback. After eight 40-second training sessions, subjects performed the key-tapping test to 90% of their original proficiency, with evident differences in the ability to perform the task without kinesthetic feedback. A single subject was trained under conditions that reduced kinesthetic, tactile, visual, and auditory feedback, and reached 79% of normal tapping performance after seven training sessions. Results such as these might lead to speculation on the existence of a function involving differences in ability to use central movement control in the absence of peripheral feedback, and that such a function would be different from what is generally thought of as motor learning.

Another approach to the study of movement control has been taken by Beggs, Sakstein, and Howarth (1974). They investigated whether measures of angular accuracy in different basic movement tasks might be related to each other, revealing a "basic movement ability" characteristic of individual subjects. The following tasks were included in their study: beam walking, in which the subject was required to traverse a 24-foot beam 1 inch off the ground as rapidly as possible without falling. Difficulty was varied by using beams of different width, 1.0, 1.5, 2.0, 2.5, 3.0, and 4.0 inches wide. The second task was line-drawing, requiring the subject to keep a stylus within a track 1.0, 2.0, 3.0, 4.0, 5.0, and 6.0 cm. wide for a distance of 250 cm., moving as rapidly as possible. An aiming task required touching a target 20, 30, 40, or 50 cm. away and returning to a home button, in time to a 120-beat/minute metronome, thus requiring one circuit from home button to target and back per second. They find a linear increase in time to move in beam walking and line-drawing, and decrease in accuracy on the aiming task, as difficulty increases, for all individual subjects. However, correlations between scores on the different tasks are very low for individual subjects, revealing no general "movement control ability." While the authors suggest that the task parameters vary too widely for the correlations to emerge, nonetheless no evidence for a general movement control ability is provided.

Other researchers in basic movement processes have been concerned with the amount of information transmitted by different types of movements. For example, Kvalseth (1976a) compared movement times for rotary hand movements (elbow and wrist held secure) to those for movements initiated at the elbow, with wrist and finger held steady. His task used a 2.5-inch stylus to tap in target circles 30^0, 60^0, or 120^0 apart and with a width of 2.5^0, 5^0, 10^0, and 20^0. Subjects were told to tap alternate circles as quickly as possible while making less than 5% errors. He finds that rotary hand movements transmit more information--about 30% more than when the movement is initiated from the elbow. Moreover, of the five subjects tested, three were twice as fast with rotary hand movements, while the other two were equally fast with the hand or arm movements. He also indicated (Kvalseth, 1976b) that movement distance and precision affect response variability as well as mean response time, such that variability also increased with difficulty. This task involved only rotary hand movements, employing a linear potentiometer with a control knob affecting movement of a pointer. Turning the control knob caused the pointer to be moved along a scale, and targets were placed along the scale using the same width and separation of targets as in the previous study. The individual differences in rotary hand movement are adequately revealed by either of these tasks.

While a consideration of choice reaction time research has been reserved for later in this review, a number of relevant studies have examined the effect of choice on basic movement. For example, Guiard and Requin (1973) looked at the effect of a preparatory period on movement time and RT. Subjects initiated a preparatory period of .5, 1.0, 1.5, or 2.0 seconds by placing a stylus on a pressure plate. A stimulus then occurred to right or left or center, and subjects released the plate and moved the stylus to the target on the side which the stimulus occurred. They found that reaction time decreased as preparatory period increased for 9 of 12 subjects, whereas movement time was unaffected for all subjects. This suggests a difference in the ability to prepare a movement decision but no effect of preparation on movement per se. Likewise, in a two-choice reaction time study, Semjen and Requin (1976) showed that, while movement time fluctuated with distance and precision required, reaction time increased only with longer distances.

These types of results seem to suggest that while basic movement is clearly possible in the absence of central and peripheral feedback, different abilities may be called upon in the different situations. As an example, Flowers (1975) used a typical Fitts tapping aiming task and compared performance to a task requiring duplication of a rhythmic pattern in the tapping, but no aiming and hence no guidance by peripheral feedback. He found that subjects with a strong hand preference--either hand--performed much better on the task requiring aiming than did ambilaterals, whereas both groups performed the same on the rhythm reproduction task (the role of handedness and laterality is discussed in greater detail later). Also, in an interesting study Marteniuk and Roy (1972) showed that subjects seemed to be able to code and utilize cues to position or location, but were unable to use

information regarding actual distance of the movement. Christina (1973) examined the effect of enforced sensory set (concentrate on the stimulus) or motor set (concentrate on the movement) on RT and movement time. He found that the enforced motor set lengthened RT, but neither set affected movement time. These and other studies suggest the relatively central nature of reaction time. Even though movement may be at least partially under central control, it seems to require little processing capacity.

Perceptual-motor movement control. Research on movement guided by sensory and perceptual feedback from initial responses has also revealed a number of individual differences. For example, Fleishman and his associates identified an ability, which they term multilimb coordination, that is separate from gross body coordination and typically involves coordination of two hands, two feet, or a hand and a foot (Fleishman, 1958; Fleishman & Hempel, 1956; Parker & Fleishman, 1960). This ability is similar to control precision, except that it involves multiple limbs, and the coordination of limbs is dependent on the use of visual and kinesthetic feedback. Because control precision is also involved in situations requiring multilimb coordination, a pure measure is difficult to establish. However, multilimb coordination has been shown to be revealed by apparatus tests such as two-hand coordination.

Performance on discrimination or choice reaction time tasks, in contrast to simple reaction time performance, is not dependent on the movement distance and precision required, but rather is a function of the information processing required in the situation. For example, Hick (1952) demonstrated that choice reaction time increased as a function of the number of aternative possible responses. That is, the greater the number of alternatives, the greater the uncertainty concerning the response. Since the information necessary to reduce this uncertainty was assumed to be processed at a constant rate, CRT increased as the number of alternatives increased. It has been shown, however (Mowbray & Rhoades, 1959), that CRT decreased with practice, so that reaction times with four alternatives became as fast as those with two alternatives. It might be suggested that the rate of processing changes with practice or learning, but most researchers (i.e., Broadbent & Gregory, 1962) suggest that the compatibility of the stimulus and the response is the key variable. That is, the extent to which a subject learns to associate a particular stimulus with its appropriate response determines choice reaction time with practice. If the responses are highly compatible with the stimulus, i.e., in the same order or location, the association is quicker and more complete, and therefore choice reaction time will tend to decrease with practice.

Fleishman and his associates have shown that when alternatives are added in a reaction time task, the relevant psychomotor ability shifts from reaction time to response orientation. Response orientation may be described as the abiliy to select the correct movement in relation to the correct stimulus under speeded conditions (Fleishman, 1957, 1958; Fleishman & Hempel, 1956; Parker & Fleishman,

1960). They have discovered that this ability is common to tasks involving visual and directional discrimination, and rapid orientation of movement patterns. It seems likely that response orientation involves the utilization of available stimulus-response compatibility, and may also involve the ability to associate responses to stimuli and initiate the responses when the stimulus and response are less compatible.

A great deal of perceptual-motor research has involved continuous tracking tasks, e.g., requiring subjects to move a control to try to keep a pointer in contact with a continuously moving target. Fleishman and his associates (Fleishman, 1958; Fleishman & Hempel, 1955, 1956) have shown that tracking performance is subserved by an ability they call rate control, which involves making continuous anticipatory motor adjustments relative to changes in speed or direction of a continuously moving target. This complex ability requires the use of visual and kinesthetic feedback to predict the movement of the target and attempt to time the motor movement so that the pointer will coincide with the target. The problem of response timing is discussed in more detail in a later segment of the review.

Recent perceptual-motor research has examined in detail the effect of different types of sensory feedback on performance in a wide variety of tasks. For example: the importance of kinesthetic cues in a two-hand coordination task is suggested by Fleishman (1972) and Fleishman and Rich (1963). Subjects were classified as high or low in kinesthetic sensitivity based on a simple psychophysical weight comparison test; then, they performed 10 trials on the two-hand coordination test. Kinesthetic sensitivity was unrelated to performance on early trials, but increasingly related to performance on the later trials.

This change in resource demand with practice on a task is not the only complexity evident in this research. Different types of cues may be substituted to allow movement control by a different-sense cue than the one normally used, and different cues may be differentially useful according to the perceptual style or strategy employed by the operator. For example, Lewis and Griffin (1976) noted that pursuit tracking performance is normally attributed to visual guidance, with complete vibration breakthrough accounting for drastic errors. They had subjects track using a free-moving or a spring-centered control device, under four levels of vibration. They found that subjects maintained control under vibration much better with the spring-centered control, suggesting that the kinesthetic feedback aids in control and that vibration may act to disrupt such feedback when the free-moving control is provided.

Ware and Barnhill (1975) investigated perception of intermediate spatial locations, rather than the usual perception of the upright. They used a 4-foot-diameter black wooden circle with pins marking degrees around its circumference, and a 3.75-foot luminous pointer in the center. Subjects were required to adjust the pointer to each hour position, excluding 3, 6, 9, and 12, while blindfolded. They were

given either no feedback, auditory feedback ("you're 4° too high"), visual feedback ("remove blindfold after movement"), or kinesthetic feedback ("keep hands on adjuster while the experimenter positions it correctly"). After two trials with the appropriate feedback, subjects were tested with no feedback. Kinesthetic feedback resulted in learning for the field-dependent subjects (see also Witkin, 1949).

In a study discussed briefly earlier, Marteniuk and Roy (1972) demonstrated the importance of kinesthetic cues in the reproduction of angular movement. They used a hardboard half-circle, with a lever control and pointer for moving a specified number of degrees. They guided subjects through a 70° standard movement, then asked them to reproduce it under the following conditions: (a) start and end points changed so that only the distance itself served as a cue; (b) start point changed but same end point, allowing some location cue; and (c) random passive movement during introduction of the standard, but start and end location cues available. If distance per se is codable, this condition should disturb distance coding. They found poorest performance when both start and end cues (i.e., kinesthetic feedback) were eliminated and no disturbance of performance by passive movement, suggesting that distance cues are not useful but that kinesthetic cues are crucial to performance.

The importance and usefulness of kinesthetic feedback is also evident in a study performed by Lackner (1974). He noted that the task of pointing at targets is performed very accurately when visually guided; that is, when the subjects can see both the targets and their hands. When the subject cannot see their hands, however, pointing performance is depressed. He suggested that the substitution of other types of feedback might improve non-visually guided pointing. Lackner used an apparatus in which six pegs were fixed 10, 20, and 30 degrees left and right of the median visual plane. Subjects could see the pegs, but could not see their hands. The task required the subjects to point at the pegs with a metal stylus, 20 times each in sequence, with the location of the point recorded where the stylus contacted a recording strip. A pre- and post-exposure trial involved no feedback. On the exposure trial, the experimental group subjects could contact the bottom of the pegs with the stylus, providing kinesthetic feedback regarding location, while the control group again received no feedback. While only 2 of 10 control subjects improved performance on the post-exposure test, 9 of 10 experimental subjects showed improved performance. Thus, exposure to tactile and kinesthetic feedback improved performance for most subjects even on a following no-feedback exposure. Analysis of movements during kinesthetic feedback exposure suggest that these subjects did not acquire stereotyped movement responses, but rather were able to use what they learned during feedback to guide movement in the absence of feedback. This could suggest an individual sensitivity to kinesthetic cues that might play a strong role in movement behavior.

Karlovich and Graham (1968) looked at control of movement in a key-tapping task with synchronous, delayed, or decreased visual feedback, when various levels of auditory placing stimuli were always

present. They found very few tapping errors even in the presence of delayed visual feedback, suggesting that auditory feedback is adequate for motor control in this case. They did find that tape durations were longer during delayed and decreased visual feedback, and suggested that subjects used an increase in tactile and kinesthetic cues to compensate for the distorted visual feedback.

In the studies mentioned previously (Lewis & Griffin, 1976; Lackner, 1974; Karlovich & Graham, 1968), it has been pointed out that visual feedback normally plays a very large role in the guidance of movement in tasks such as tracking and perception of location. A large number of recent studies have been concerned with the effect of alteration of visual feedback on performance. Wallach and Smith (1972) used prisms to alter the accommodation and convergence of the eye. When the adaptation procedure was such that kinesthesis provided the only true cues for distance, subjects showed a strong kinesthetic adaptation in a later pointing task. When accommodation and convergence were increased by the prisms, the adaptation procedure provided veridical visual depth cues as well as kinesthetic cues. In this case, a visual adaptation takes place in judging size and distance of objects, but a large kinesthetic adaptation is evident in the pointing task. Wilkinson (1971) used prisms that caused a change in judged visual direction of a target and in the subjective perception of the location of the hands. After adaptation, performance on a pointing task appeared to be a function of both prism-induced changes, but performance with the untrained arm was a function of visual direction adaptation only. Yachzel and Lackner (1977) showed that multiple exposures to visual rearrangement led to long-lasting adaptation effects. After adaptation, pointing without seeing the hand showed diminished variability with practice, but no diminishment of constant error. The aftereffects of the visual rearrangement persisted for at least 2 weeks. Moreover, the adaptation evident in the pointing task transferred to different movement tasks when the hand could not be seen.

Using a slightly different procedure, Meisel and Wapner (1969) examined the influence of different types of visual cues on spatial localization. They used a lighted box with three luminous bars, placed in the median visual plane, or to the left or right of the median, and a prism to increase covergence of a single eye, placed on either the left or right eye, or on neither. They presented visual stimuli every 3^{0} from 36^{0} left, to the median plane, to 36^{0} right, for 25 total stimuli. Subjects responded with a button press indicating whether the stimuli were left, center, or right on the median visual plane. They found tht the extended lighted box shifted perception in the direction of extension, whereas the convergence prism shifted perception in the direction opposite greatest convergence.

Fleishman (1972; Fleishman & Rich, 1963) demonstrated that greater sensitivity to visual-spatial cues (a spatial orientation ability) was related to successful performance on a two-hand coordination test on early trials, but as learning proceeded, a

kinesthetic sensitivity ability became more important. The pattern of results discussed above also seems to suggest that utilization of kinesthetic and other sensory cues can be substituted for visual feedback in a large number of tasks, and that the importance of kinesthetic cues increases over time. Moreover the studies of visual adaptation to prisms point out the extent to which the mechanisms are interconnected; that is, when a movement occurs guided by vision that is adapted to prisms, a kinesthetic adaptation to the visually distorting lenses also occurs.

While spatial abilities have long been of interest to psychometricians, there has recently been a renewed interest in the ability to perceive holistic patterns, orient oneself within those patterns, and guide oneself through them (McGee, 1979; Egan, 1979; Koslowski & Bryant, 1977). For example, Koslowski and Bryant (1977) asked subjects to rate how good their sense of direction was. They then had the subjects imagine they were standing in front of a familiar building or campus, indicated by a circle on a sheet of paper. The task was to indicate the direction of several other buildings on campus, relative to their imagined position. Subjects were also taken through a novel human maze for several trials, and asked to indicate the direction of the starting point and to draw a representation of the route. Subjects who rated their sense of direction highly made smaller errors in indicating directions of familiar buildings, and learned to represent the novel maze over trials; however, those with poor sense of direction made large errors and showed little learning of the maze.

Weitzman (1979) classified subjects as good or poor in sense of direction based on accuracy of pointing to landmarks in familiar and novel environments. Subjects with a good sense of direction performed better on a movement reproduction task, and moreover appeared to employ more appropriate spatial strategies. Allen, Siegel, and Rosinski (1978) presented subjects with a sequence of slides representing the visual experience along a walk, and found that subjects incorrectly recognized unpresented slides from the same walk, but easily distinguished slides from other walks. Moreover, thme distance estimations of subjects were related to actual distances, even when the slides for the walk were presented in random order rather than sequentially. In appears, therefore, that the crucial ability of spatial orientation revals individual differences in these performance tasks, and that they might be used to reliably assess the spatial ability.

Other perceptual-motor processes. One of the more intriguing propositions in recent perceptual-motor research is that proficiency with spatial tasks may be tied to the degree of cerebral lateralization. Levy (1973) has argued that degree of specialization of the brain affects both verbal and spatial performance. That is, a less lateralized person would suffer interference, particularly with spatial materials, while a more specialized brain would handle the tasks with less interference. One difficulty, however, has been to provide an adequate operational definition of degree of lateralization.

Gilbert (1977) attempted to see whether lateralization was related to a visual-spatial deficit, and to see whether left-handers are less lateralized. He defined left-handedness in terms of stated preference, family members being left-handed, and manual dexterity tests. He then gave subjects standardized spatial abilities tests (i.e., the Weschler Adult Intelligence Scale Design and Object Assembly tests) and looked at laterality differences with visual half-field reaction time for face recognition (a right hemisphere ability) and letter discrimination (a left himisphere ability).

Gilbert found that RT for face recognition was fastest in the left visual field (right hemisphere), but unrelated to handedness. Of those subjects with a right visual field bias for face recognition, an equal number were left- and right-handed, but this group showed poorer performance. Thus there is some suggestion that a right-hemisphere bias for visual materials results in superior performance, but no evidence that handedness is related to that superiority. Fennel, Satz, Abell, Bowers, and Thomas (1978) used high school and college students, assuming that the latter would have superior verbal abilities, and gave them dichotic listening and visual half-field tasks to categorize them according to cerebral dominance. They also determined preferred handedness for each subject, and administered the Wechsler Block Design and the Primary Mental Abilities visual-spatial subtest. There were no differences in performance on the spatial tests between those categorized as right or left hemisphere dominant. Their handedness classification shows the same pattern as classification by the dichotic and visual half-field tests. Overall, this suggests that left-handed individuals may be lateralized differently, but it does not lead to a spatial performance deficit.

Handedness has been implicated in other perceptual-motor tasks as well. Flowers (1975) compared performance in a visually guided target-aiming task to rhythmic tapping performance not requiring visual feedback. He found that subjects with no strong hand preference performed more poorly on the target-aiming task than those with a preference, and among those with a preference, there was no advantage to being right- or left-handed. No differences as a result of handedness occurred on the rhythmic tapping task. He suggested that sensory--in this case visual--feedback is better utilized by more strongly lateralized individuals (when laterality is indicated by strong hand preference). Berlucci, Crea, diStefano, and Tassineri (1977) looked at simple reaction time as a function of whether the stimulus was presented in the same visual field as the response side. They found faster RTs with the hand ipsilateral to the visual stimulus, even when the hands were crossed. This advantage therefore could not be attributed to spatial compatibility of the stimulus and response. Similar effects were found in a choice reaction time test. They suggested that anatomical relationships between the visual fields, cerebral hemispheres, and the hands may account for this advantage and that these anatomical relationships may interact with stimulus-response compatibility. Rigal (1974) found no differences on a right-left discrimination test or a preferred-hand efficiency test between right-handed and left-handed subjects. Overall, this

promising area of research appears at present to be resulting in data too confusing and conflicting to allow safe inferences about lateralization and its relation to human processing to be drawn.

We noted earlier that timing or rate control can play an important part in tasks such as compensatory tracking, and by analogy it is important to any steering-type task (Fleishman, 1953). The problem of timing a response to coincide with the occurrence of an external event is a fairly general and complex one, and has been approached from a number of points of view. For example, Adams and Creamer (1962) trained subjects to time simple responses such as pressing a button or speaking a word, to co-occur with changes in a stimulus (changing direction of a sine wave). They found that such training transferred positively to performance in a pursuit tracking task, and hypothesized that this ability to time responses was dependent on perception of kinesthetic feedback. It is clear, however, that this ability also depends on the presence of redundancy in the stimulus pattern. The ability to perceive this redundancy allows a subject to predict target movement prior to timing the response. It is less clear, however, whether the ability to perceive this redundancy in stimulus patterns is a separate individual ability, or merely a function of memory storage capacity.

Recent work on timing behavior has concerned itself with the effects of stress or difficulty variables on timing performance. Dorfman and Goldstein (1975) utilized a complex timing task with multiple inputs and examined the effects of speed/stress, task coherency, and preview on timing performance. Three dials were employed, each with a central target line and a moving point. The subjects' task was to reverse the direction of the pointer by pressing an appropriate response button, at exactly the time the pointer crossed the target line. Speed stress was varied by having the pointer cross the target line 15, 30, 45 times per 30-second trial. Coherency involved the velocity of the pointers, which were either all three equal (total coherency), two of three equal (partial coherency), or all different (no coherency). Preview involved the percentage of the dials visible to the subject, i.e., completely visible, small mask, or large mask. As might be expected, correct responses decreased and latency increased as the difficulty of all three variables increased. The partial and no coherency groups, and small and large masking (preview) groups, do not differ, indicating these variables interact with timing performance, while speed/stress has a consistent effect and does not interact. While the effect of preview is small overall, it appeared to be especially helpful at high stress, whereas coherence was minimally helpful in that situation. Thus, it appears that differences in timing performance are influenced by a variety of variables which may have simple or complex effects.

Wrisberg and Herbert (1976) used a coincident timing task in which subjects had to sweep their arm from right to left to attempt to knock over a 12 cm^2 hardboard target at the exact moment a sweep hand arrived at a target point on a dial. Subjects received 4 days of practice at 50 trials per day, so that the asymptotic performance of

well-practiced subjects could be established. They then examined the effect of fatigue on timing performance, either by inducing local fatigue requiring maximal contractions of the right shoulder muscles or by inducing general fatigue using a treadmill. Subjects then again attempted to perform the timing task. On the first postfatigue trial, performance was depressed relative to day 4 performance. Moreover, the two fatigue groups differed, in that general fatigue induced a speed-up, so the target was knocked over too early, whereas local fatigue produced a slowdown and the target was knocked over too late. These differences diminished by the second postexposure trial and disappeared thereafter, suggesting that fatigue, at least at the level induced here, does not have long-term effects on timing performance.

While we have discussed at length the effects of sensorimotor feedback on perceptual-motor performance, more cognitive feedback in the nature of knowledge of results may also affect performance. For example, McCaughan (1978) looked at performance in a mirror star-tracing task, in which subjects were given fictitious feedback indicating success, failure, or random performance, or given no feedback. He found in general that performance was best with success feedback, and subjects tended to perceive the task as a function of ability, whereas with random feedback people assigned a role to luck in the performance of the task. Sterner and Carpp (1977) also showed a role for cognitive functions in perceptual-motor performance. They trained subjects in a pursuit rotor task, requiring one group to perform seven 10-minute sessions of rehearsal of tracking. These subjects were told to recall and rehearse the visual and physical movements needed to track. The other group alternated tracking sessions with naming of objects presented on slides. Sterner and Carpp found significantly enhanced performance in tracking for the group that mentally rehearsed. They suggested that this might be attributed to the formation of a visual image to help guide the task, but they are more inclined to believe that the rehearsal allows additional cues for tracking to be transferred to long-term memory when those cues could not be stored during actual tracking performance.

Summary. The discussion of research on basic movement control does not deny the possibility of central processing involvement in these tasks, but merely suggests that in most cases the tasks are less dependent on central processing. A large number of potentially reliable tasks are available: simple reaction time, successive reaction times (the psychological refractory period), Fitts' tapping/aiming task, rotary pursuit, two-plate tapping, and rotary hand movements. It seems very likely, however, that the underlying abilities and operations for these tasks have considerable overlap, and only the most useful should be included in a test battery. Other tasks, such as beam-walking, line drawing, etc., were less promising in terms of experimental results, and appear to be less adaptable to a test battery.

Those tasks considered in the discussion of research on perceptual-motor movement control, on the other hand, depend on the utilization of sensory feedback cues to correct and guide movement.

The evidence on such cue utilization is complex; visual and kinesthetic cues appear to be of greatest importance, but different sensory cues can substitute for the normally preferred one. Moreover, people not only have differing sensitivities to the cues but also differ in the cue they prefer when multiple cues are available. All of these variables affect performance. Tasks of potential relevance revealed in this literature include apparatus tasks such as complex coordination, two-hand coordination, compensatory tracking, and choice reaction time. Other performance tasks including directional pointing, spatial orientation, distance estimation, and angular movement reproduction also appear to have potential value. Again, however, these tasks are unlikely to tap totally independent underlying abiltiies, and selection to reduce overlap and employ the most efficient tasks is still required.

The results of research on hemispheric lateralization are complex, and further complicated by the inadequate definitons of lateralization in terms of handedness, or performance on dichotic listening and half-visual field tests. While there is some evidence of a right hemisphere specialization for spatial ability (Gilbert, 1977), this evidence is sometimes contradicted (Fennel et al., 1978). It has been suggested that degree of lateralization per se and not right hemisphere specialization underlies visual performance, and perhaps even that the lateralization underlies a general sensory feedback utilization ability (Flowers, 1975). Moreover, the complicated anatomical relationships postulated by Berlucci et al. (1977) are not amenable to easy testing. There appears to be some justification for inclusion of a test for lateralization, e.g., a dichotic listening task, if it is not costly in terms of implementation, and if results of that test are viewed with caution.

Similarly, a direct test for response timing such as that employed by Adams and Creamer (1962) might be included, or the response timing function might be considered to be revealed in the rate control ability subserving compensatory tracking, or related to the kinesthetic sensitivity measure. The influence of cognitive factors on perceptual-motor performance is clear from this review, but it seems likely that tasks suggested by the review of the cognitive literature should tap the abilities underlying that influence.

COGNITIVE PROCESSES

Contemporary aircraft have become increasingly sophisticated with technological advances, automating many functions previously perfomed by the pilot, but also increasing the load placed on the pilot's cognitive functioning. The speed and accuracy with which information is perceived, encoded, stored, transformed, and compared, the speed with which memory is searched and accessed, and the speed with which appropriate decisions may be made are all crucial to pilot performanmce. Recognition of this fact is reflected in the literature on pilot training and selection. For example, current pilot training research places a great deal of emphasis on cognitive pre-training techniques to improve comprehension and integration of necessary aviation information (Gerlack, 1974; Crosby, 1977), and programs for judgment training have been formulated (Jensen & Benel, 1977). Selection tests have focused on selective attention (Gopher & Kahneman, 1971) and time-sharing tasks (North & Gopher, 1976) as predictors of pilot performance, but recent test batteries have included a number of tasks involving memory, spatial visualization, comprehension, and other cognitive functions (McLaurin, 1973; Pew & Adams, 1975; Hunter, 1975). Indications that these test batteries may be effective suggest the examination of the cognitive literature for other potential tasks.

This chapter provides a review of the role of cognitive tasks in pilot training and selection research, followed by a more general review of the literature. This latter review considers research on cognitive operations thought to be relevant to piloting, and is particularly concerned with demonstrating or inferring that a task reliably indicates individual differences in the relevant cognitive process. The goal of this chapter is to identify a number of tasks with potential utility as pilot selection devices.

Pilot Training and Selection

The importance of cognitive factors in pilot training is often pointed out. For example, in a review of pilot training literature, Smode, Hall, and Meyers (1966) emphasized components of the pilot's task such as time-sharing and performance under stress, and they urged further research on a variety of instructional variables. Similarly, Matheny, Gray, and Waters (1975) suggested research or instructional variables such as cognitive pretraining, task sequencing, contextual training, and individualized instruction. In general, efforts to improve pilot training have focused on (a) the use of simulators as general and specific training devices, for actual transfer to actual flying performance, (b) the identification of techniques and adjunct aids to facilitate comprehension and integration of the material, and (c) the identification of information processing components relevant to piloting, for specific training.

Clearly the variety of simulation devices employed in pilot training constitute an instructional aiding device, with varying

fidelity to the actual task of piloting. The analog nature of simulation tasks renders them tangential to this review, however, since they are difficult to relate to a specific component of flying. A variety of other instructional aids are being employed, though. Members of the Rand Symposium on pilot training (Stewart & Wainstein, 1970) included among current training innovations study carrels to aid in accessing information for review and study, films and slides for procedure reviews, single concept films to aid in learning specific maneuvers, automated briefings, and video trainers. All of these devices are employed to make requisite UPT information readily available for study, practice, and review.

In addition to this concern with instructional devices, significant effort is being made to improve instructional techniques by optimizing the presentation of information and practice to facilitate learning and integration. One such technique is cognitive pre-training, which involves formation of appropriate cognitive structures prior to training, so that information is readily integrated rather than forcing the trainee to infer the structure while learning the specific information. This technique has been applied to a number of problems in pilot training. For example, Gerlach and his associates (Brecke, Gerlach, & Shipley, 1974; Gerlach, 1974, 1975) chose a representative flight maneuver and analyzed the instruction manual and briefings for cues crucial to performance of that maneuver. They then developed an algorithm for the systematic presentation of these cues. Using a cognitive pre-training procedure, they compared the effect of pre-training with no cues to pre-training containing cues arranged either systematically according to their algorithm, or concurrently with the order in which they appear in the flight training manual. The pre-training included both a written presentation and a taped briefing on the flight maneuver, and both contained either no cues or cues either concurrently or systematically arranged. Gerlack looked at acquisition of a complex skill--the flight maneuver--as a result of the training conditions, and found the following results: (a) pre-training which includes instructional cues results in better performance than pre-training with no cues, (b) systematically developed cues are better than concurrently arranged cues, and (c) systematically developed cues result in lower within-group variability of performance, suggesting that individual strategies play a smaller role. These are encouraging results suggesting that properly structured pre-training including instructional cues can greatly facilitate training of complex perceptual-motor skills.

The technique of cognitive pre-training has proved useful in other areas of training as well. Crosby (1977) suggested that the formation of appropriate visual schemata through pre-training would ease the transition from instrument flying to composite flying depending largely on external information sources. His pre-training procedure improved performance markedly on a laboratory discrimination task designed to be an analog to the external information sources available to the pilot. Other versions of the pre-training technique have been employed to improve training of complex skill learning as well as preparation for simulator training and formation flying skills.

A much different approach was taken by Leshowitz, Parkinson, and Waag (1974) who conducted a series of auditory and visual information processing studies with the aim of defining the specific processing demands of the piloting task. They selected different tasks for their experiments to use different input modalities and to tap different processing stages. In Phase I of these programs, four experiments were conducted. The first was an array search experiment, requiring subjects to scan and recall a 3x4 array of letters. They feel that the potential for training efficient scanning strategies would be useful to the pilot for acquiring and retaining information from an array of instruments and dials. A second experiment presented a visual pattern followed at some interval by a visual masking stimulus. On the average, if 250 msec elapsed before presentation of the mask, it had no effect on memory for the visual pattern, but large individual differences in speed of visual encoding were evident. A third study required processing of auditory patterns. Two sequences of tones were presented, and subjects required to say if the second was the same or different from the first. They found that processing of the auditory patterns could be accomplished when duration was about 200 msec--a considerably longer time than is required for visual processing. The fourth experiment in Phase I required subjects to remember a visual figure and mentally rotate it to decide if it is the same or different from a second stimulus. They found no need for language mediators to perform this task and suggested it may be related to important piloting abilities of visualization and decision-making. The second phase of their research program examined the relationship of digit span to processing in a series of dichotic listening studies. They showed that recall in these studies was an increasing function of digit span, and interference with recall of a digit sequence due to a stimulus suffix was an inverse function of digit span. When digit recall was used as a primary task and reaction to a visual probe as a secondary task, processing capability (as measured by the probe reaction time) showed wide individual differences but no relation to digit span. This suggests that digit span measures a passive memory component useful in avoiding interference, but unrelated to active processing capacity. They suggested that this program indicates a number of areas where training might improve the processing of visual and auditory information crucial to pilot performance.

One element of pilot performance for which there is significant agreement that the element is crucial is the ability to time-share. Gabriel and Burrows (1968) were concerned with the ability of pilots to allocate attentional resources appropriately between instrument and within-cockpit sources and visual information sources external to the cockpit. They devised an 8-week training program to enhance display reading and hazard detection performance in simple adaptive simulation devices. Following this training, experimental subjects showed significantly better hazard detection performance than a control group in a highly complex flying task. Moreover, this improvement in no way compromised performance on other flying tasks, suggesting that the experimental subjects were able to time-share more effectively in the complex task.

Another skill frequently mentioned as crucial to piloting is the exercise of judgment. Jensen and Benet (1977) have looked at the potential for training judgment in civilian airline pilots. They defined judgment as the ability to search for and establish the relevance of all available information regarding a situation, to specify alternative actions and determine the expected outcome of each alternative, and the motivation to choose and to authoritatively execute a suitable course of action within the time frame permitted. Thus their definition includes an intellective information-seeking and decision-making part, and a motivative propensity for risk-taking factor. They pointed out that some judgment training already goes on, when potential emergency situations are analyzed and discussed. This procedure teaches potential pilots to discriminate alternative action courses, and specifies correct response selection with the student pilot generally required to memorize these emergency procedures. They favor a situational approach to judgment training, suggesting that development of a judgment training curriculum must be preceded by specifying the desired behaviors and developing the principles for judgment training as well as training instructors to implement these procedures. They recommended training on four levels: ground school, computer-assisted instruction, simulator instruction, and actual aircraft instruction (although the situational approach is less practical in the aircraft, where a poor judgment could be disasterous). The basic technique involves the presentation of a situation, demonstration of information seeking and response selection and of the effects of stress and work load on the decision process. Evaluation of this training through transfer to real-world flight decisions will be difficult at best. Nonetheless, the several processes lumped together as judgment are crucial to flying and deserve attention in the selection and training process.

Information processing and cognitive functions have been recognized as a potential basis for selection as well as training, although the use of information processing tests for pilot selection has focused almost exclusively on attentional phenomena. Gopher and Kagneman (1971) demonstrated individual differences in selective attention and showed that these differences were predictive of various flight criteria. They employed a two-part dichotic listening task in which subjects monitored one ear for the occurrence of digits in a stream of words, and reported the digits when they occurred. In Part 2, subjects monitored either the same ear or the other one, and received pairs of digits simultaneously (one to each ear) and were to report the digits from the relevant ear. They found most errors in Part 1 to be omissions due to verbalizing the previous digit while the next one is occurring. Also intrusion errors occurred when the left ear was relevant, suggesting the familiar left-hemisphere bias for dichotic listening. In Part 2 instruction errors were most common, but perhaps not because of confusion due to switching the relevant ear. Most errors occurred in Part 2 when the left ear was relevant for both parts. Both omission errors in Part 1 and intrusion errors in Part 2 correlated significantly with various flight criteria including the pass-fail criterion.

The ability to divide attention between two or more inputs, often referred to as a time-sharing ability, has long been considered crucial to the task of piloting. Melton (1947) noted that rotary pursuit in a divided attention task not only provided moderate predictive validity for early flight training success, but maintained its validity in the later stages of flight training. North and Gopher (1976) required subjects to perform a digit processing reaction-time task and a compensatory tracking task, each task alone and both tasks concurrently. They looked at the ability of the subjects to adjust performance in the dual-task situation to different proportions of their single task performance. They were able to demonstrate that flight students with high potential were reliably discriminated by performance in the dual-task situation.

A different approach to divided attention as a predictor of flight performance was taken by Amos (1978). She reasoned that the ability to time-share was dependent on the fixed processing capacity of the subjects and, therefore, the ability of a subject to time-share should be a function of the amount of processing capacity remaining when performing two tasks concurrently. That is, "because an individual's residual attention acts as an upper-limit on his time-sharing performance, a realiable measure of residual attention may be an effective predictor of pilot performance" (Damos, 1978, p. 436). She employed a primary compensatory tracking ta k, and for a secondary task she used a choice reaction time test in which one of two, four, or eight digits was presented and responded to with a corresponding pushbutton. A cross-adaptive system was used to keep tracking performance in the dual-task situation relatively constant. Subjects were also rated on a pilot performance scale, after 10, 20, and 30 hours of training in an introductory flight course. Mean reaction time as a measure of residual attention was correlated with performance on the 10-, 20-, and 30-hour flight checks, yielding correlations of .59, .64, and .68, respectively. Only the last correlation was statistically significant, but the increase in correlations with flight training suggests that residual attention might be an important predictor for success at higher levels of flight training, where general intelligence and motor tests begin to lose their predictive validity.

Walker and Walker (1979) also investigated dual-task performance among pilots using a primary tracking task with different levels of difficulty as an analog of piloting, and a secondary auditory discrimination reaction task to stress performance of the primary task. They trained and tested groups of pilots, including Air National Guard pilots, Air Force pilots, and World War II "aces." They found that performance on a tracking task of moderate difficulty (requiring leading or anticipating marker movement) was predictive of success with various cockpit simulation devices, as well as, to a lesser extent, actual combat performance. Fournier and Stager (1976) also validated a dual-task paradigm for operators of communication equipment aboard military anti-submarine aircraft. Performance on a visual decision task combined with an auditory short-term memory task was related to a number of indices of operator performances.

While research on single-test predictors within the field of information-processing has been largely confined to attentional phenomena, other information-processing tasks have appeared in more recent batteries of tests designed for pilot selection. For example, McLaurin (1973) developed a battery of performance tests to assess aerospace ground equipment crews. This battery included the following tests:

1. Arithmetic computation - subject adds or subtracts 1 or 2 digit numbers; scores the number correct and the mean and standard deviation of latency.

2. Short-term memory continuously changing - after a series of random digits presented, subject must press the button to indicate the third number previous; the score is the number correct.

3. Paired associate learning - six items of one letter paired with two digits are presented three times; on fourth trial, letter only is presented and subjects recall paired digits; score is the number correct.

4. Switch activation - digits 1, 2, 3, or 4 are presented on a green or black background; subject must leave home key and activate appropriate switch for digit up for green, down for black background - score is the mean and standard deviation perceptual speed and motor speed and the number correct.

5. Complex short-term memory - three, four, five, six, or seven digits are presented; subjects add the last three and punch in the total - score is the mean and standard deviation response latency and the number correct.

6. Pattern memory - bar graph configuration is presented, followed by a second pattern; subjects respond whether second pattern is the same or different from first - score is the number correct.

7. Average estimation - five digits presented for 3 seconds; subjects estimate the mean - score is the mean and standard deviation response latency and the number correct.

8. Complex counting - subjects count and remember the number of times digits 1, 2, and 3 occur in a random sequence; when one occurs three times, subject pushes its button and starts again for that number; score is the number correct.

9. Kinesthetic learning - for five trials, at auditory tone subject activates buttons 1, 4, 2, 3, in that order, on sixth trial is blindfolded - score is the mean auditory reaction time, mean movement time, and number correct.

Using this test battery, McLaurin (1973) found significant correlations with performance grades in different segments of training, with the correlations ranging from .30 to .61. Moreover, when combined with the Electronic Aptitude Index (EAI), this battery added significantly to the predictive validity of the EAI. The contribution of the individual tests to the overall validity of the battery was not evaluated.

One of the more extensive attempts to construct and validate an information processing performance test battery has been undertaken by Hunter, 1975 (see also Hunter, Maurelli, & Thompson, 1977). This battery was constructed to be implementable on a small computer test bed, and with the requirement that it be as free from verbal biases as possible. A number of these tests are very similar to those employed by McLaurin (1973). The tests included:

1. Kinesthetic memory - same as McLaurin's kinesthetic learning task, but with 12 blind trials and with some geometric figures rather than digits as stimuli.

2. Perceptual speed - same as the kinesthetic memory task, but not blindfolded, and includes a measure of study time prior to response initiation.

3. Performance under stress - same as perceptual speed, except some figures are shaded and require a different response.

4. Associative learning - line drawings or novel geometric figures are paired with the geometric responses at presentation; each pair is presented three times for study, followed by a paired-associate recall trial.

5. Memory - immediate/delayed - random presentation of a sequence of geometric figures at 2 sec (immediate) or 5 sec (delayed) intervals; subject presses a button corresponding to the item which occurred two items previously.

6. Concept identification - presentation of four geometric figures with an element in common (e.g., right angles); subject presses a button indicating whether or not a fifth figure belongs in the set.

7. Divided attention - requires simultaneous performance of a visual and an auditory compensatory tracking task.

In an initial test of the development of this battery, Hunter (1975) demonstrated the reliability of all the measures except the associative learning test. He went on to perform a factor analysis of the data from over 300 subjects on the many dependent measures in these tests, uncovering seven distinct factors. These included Visual Tracking, Auditory Tracking, Figural Memory, Positional or Kinesthetic Memory, Percpetual Speed, Motor Speed, and a somewhat vague factor he referred to as Associative Speed.

A later study went on to validate this test battery with Air Force Officer trainees and with enlisted personnel (Hunter, Maurelli, and Thompson, 1977). They found that several of their tests were individually predictive--for success in navigator training (the officers) the kinesthetic memory response time was the best predictor, followed by number correct for kinesthetic memory and number correct for immediate memory, the next best predictors. The test battery did not achieve predictive significance as a whole for navigator training, but this may have been due to the relatively small sample size (n = 77).

For the enlisted personnel, performance was validated against three training criteria. For the criterion of training disposition-pass/fail, the best predictors were kinesthetic memory-response time, and perceptual speed-number correct. For the Air Training Command standard score criterion, the best predictor from this battery was kinesthetic memory-response time, while for the number of washbacks criterion, the best predictor was perceptual speed-response time. Overall the test battery showed small but significant predictive value for the first two of these criteria, but little or none for the washback criterion. They concluded that several measures within this test battery are of potential use for improved selection of Air Force personnel, and that potential is especially evident with regard to the navigational training of officers.

Hunter and Thompson (1978) performed extensive studies to examine the validity of a battery of tests including several pencil and paper tests, the two psychomotor tests developed by McGrevy and Valentine, and the Automated Pilot Aptitude Measurement System (APAMS) which essentially assessed performance on a flight simulation device. While this study was beset by problems of changing samples and complex scores for the psychomotor and APAMS tests, they were able to demonstrate validity for a battery consisting of these two psychomotor tests and seven reference pencil and paper tests. They recommended use of the pencil and paper tests for initial screening, followed by use of the psychomotor and APAMS tests. In addition, they pointed to the potential utility of information-processing tasks, and tasks concerned with higher level integrative ability, for pilot screening and selection.

A wide variety of other test batteries have been developed, which employ a number of information processing tasks. Pew and Adams (1975) developed a test battery to select and assign subjects in Advanced Simulator for Undergraduate Pilot Training (ASUPT) experiments. The battery included a digit span test to estimate active memory capacity, a rotated letters test for spatial abilities, dichotic listening and the Stroop test to estimate facets of selective attention, a sentence verification test of comprehension, and time-sharing tasks. The battery also included a critical tracking task and job sample tests. They concluded (Pew, Rollins, Adams, & Gray, 1977) that all of the tasks but one of the time-sharing tests warranted inclusion in the test battery.

Test batteries incorporating information processing tasks for evaluation in a variety of job environments continue to be developed. Emurian (1978) described a mini-computer controlled battery which incorporates probability monitoring, arithmetic operations, target identification, warning light monitoring, and blinking light monitoring. This battery is similar to that employed by Alluisi and his associates (Alluisi & Chiles, 1967) in a variety of research settings. Rose has developed a test battery based totally on information-processing tasks (Rose, 1974; Rose & Fernandes, 1977). Kennedy and Bittner (1978) have relied heavily on information processing tasks in the development of their Performance Evaluation Test for Environmental Research (PETER). It seems likely that the role of cognitive tasks as selection devices will continue to grow and be emphasized.

The role of simulators in flight training was discussed briefly, but performance on simple ground-training devices has also been used to predict success in UPT. In a sense, simulator performance in a job sample task could be viewed as reflecting differences in the composite of perceptual-motor and information processing functions. LeMaster and Gray (1974) looked at objective measures of ground-training device performance for naive UPT candidates, and found the job-sample approach to be highly predictive of performance in the T-37 phase of training. It did not predict attrition for reasons other than lack of flying skill, however.

Summary. Our overview of pilot training and selection research makes clear the importance of cognitive processes for flying. A portion of the training literature has focused on the employment of comprehension aids to improve the program of training per se. This research has examined the efficacy of various devices (e.g., information retrieval carrels, slides, and movies) for training pilots. In addition, attempts have been made to specify the important cues in flying, and to develop a pre-training regimen which facilitates acquisition and use of those cues (Gerlach, 1975).

A related training approach would require the specification of information-processing components that correlate with actual flight performance. Leshowitz, et al. (1974) conducted exploratory studies attempting to relate performance on visual and auditory information-processing tasks, and digit span tasks, to training of pilots. Others have agreed that two elements crucial to piloting are judgment and the ability to time-share. Training programs aimed at improving those abilities have been undertaken, with more obvious success in regard to time-sharing, which is easier to define and to measure.

Where pilot selection research has focused on a single information-processing skill, that skill has tended to be related to attention. A wide variety of studies have supported selective attention tasks (Gopher & Kahneman, 1971) and divided attention tasks (Damos, 1978; North & Gopher, 1976) as reliable predictors of pilot performance. In general, however, recent selection research has

focused on the development of multi-test batteries in which a variety of cognitive skills must be employed. Similarly, job sample approaches to pilot selection could be viewed as multi-component in nature, requiring a variety of cognitive and other processes for successful performance. Both from the point of view of updating tests to the state of contemporary research, and in view of the predictive validity evidenced by these tests, a search for new and superior cognitive tasks with which to formulate selection tests appears to be mandated.

Individual Differences in Cognitive Processes

Although psychologists concerned with cognition and information-processing have generally attempted to discover and account for regularities in processing across individuals, several forces have coalesced recently to focus attention on individual differences in processing. Basic researchers in cognitive phychology have become discouraged with the inability of numerous models to consider and account for individual variability. Psychometricians have become disenchanted with factor-analytic approaches to the description of intelligence, and have begun to construct a theory-based view of intelligence. One such theory-based view deals with intelligence in terms of individual differences in various information-processing components. The third force at work here involved applied psychological research, especially with regard to personnel selection. Advances in technology have changed the demands of many jobs, decreasing physical requirements but emphasizing the role of information-processing and decision-making skills.

While applied psychologists have always recognized the value of tests of cognitive abilities as screening and selection devices, they have typically relied on the criterion-validated tests produced by psychmetricians. Recently, however, they also have perceived the potential value of relating such tests to the theory and data of cognitive psychology. Moreover, a concern for the potential verbal bias inherent in pencil and paper tests has encouraged the development of performance tests capable of revealing differences in cognitive ability without being subject to that bias. Tasks from the cognitive laboratories appear to be a tempting source for such performance tests.

A number of test batteries made up of cognitive tasks have been constructed for specific selection purposes, including those discussed previously by McLaurin (1973) and Hunter (1975). One extensive effort in the development of an information processing based test battery was that of Rose and his colleagues (Allen, Rose, & Kramer, 1978; Fernandes & Rose, 1978; Rose, 1974; Rose & Fernandes, 1977). They began with a set of cognitive operations modeled on those described by Carroll (1976), including encoding, transformation, memory storage and retrieval, and responding. They then selected sets of tasks from the cognitive research literature which had been extensively replicated and involved the different cognitive operations, and tested them for reliability and success in revealing individual differences.

Rose and Fernandes (1977) employed a test battery including a letter classification task (Posner & Mitchell, 1967), lexical decision task (Meyer, Schaneveldt, & Ruddy, 1974), graphemic and phonemic analysis task (Baron, 1973), short-term memory scanning (Sternberg, 1967), scanning for words and categories (Juola & Atkinson, 1971), linguistic verification (Clark & Chase, 1972), semantic memory retrieval (Collins & Quillian, 1969), and recognition memory (Shepard & Teghtsoonian, 1961). A second effort (Fernandes & Rose, 1978) focused primarily on a variety of memory tasks, and Allen, Rose, and Kramer (1978) added a letter rotation task (Shepard & Metzler, 1971), visual scanning (Neisser, 1967), mental addition (Hitch, 1978), and sentence recognition and sentence recall tasks, (Bransford & Franks, 1971). While not all of these tasks proved to be efficient and reliable indicators of individual differences, many showed substantial reliability and validity, and the effort is very encouraging regarding the potential usefulness of information-processing tasks in test batteries.

Current research in psychometrics, basic cognitive psychology, and applied psychology points to the importance of understanding individual differences in cognitive performance. All three approaches have contributed to the recent spurt of research on individual differences in cognitive processes; moreover, they provide a starting point from which inferences may be made about other research not specifically designed to examine individual differences. In the following sections of this review, we consider research on a variety of cognitive processes relevant to flying, conducted from several theoretical points of view, in order to identify tasks that are reliable indicators of individual differences in those processes.

Perception and attention. This segment of the review considers research on the speed and individual structure of perception, and research on how attention limits the speed and amount of information perceived. While this research ranges from discussions of foveal visual asymmetries to exploration of the ability to allocate resouces to disparate perceptual inputs, they are related in that they address the nature of the information perceived.

While it is clear that individual differences do occur as a result of differing sensitivities of the various sensory receptors, more complex differences are also apparent in a number of perceptual phenomena. For example, Schaller and Dziadosz (1975) noted differing asymmetries in foveal visual perception. They presented subjects with a 5x7 array of 35 circles, one of which was crossed by a vertical or horizontal bar. Subjects were required to indicate the direction of the bar by pressing the appropriately coded response key. While performance in general was better for the top half of the array and decreased with increasing distance from the center, subjects displayed biases for stimuli as a function of the direction of the stimulus from the center. About 2/3 of the subjects were superior on items to the left of center, while 1/3 were superior on right-side stimuli. They rejected the hypothesis that reading-scanning habits account for these asymmetries, since none of these subjects should have a reading bais

right-to-left. Rather they tentatively attributed the superiority of the top half to an attentional bias, and suggested that greater visual acuity in one eye leads to a strategy favoring one visual field over another, and resulting in a right- or left-field bias.

Another perceptual phenomenon in which individual differences are evident is the figural aftereffect, which occurs when prolonged inspection of figures modifies the apparent size, location, or tilt of contours later viewed in the same part of the visual field. Differences in the rate of recovery from the aftereffect have traditionally been explained by reference to satiation and recovery of cortical activity. Over (1970) pointed out that many of the differences reflect response differences; that is, formation of a response set, criterion adopted for judgment, and the speed of judgment. Therefore figural aftereffect differences cannot be taken as a pure measure of individual differences in cortical activity.

A number of researchers have looked at differences in the way sensory patterns are perceived and responded to. Mavrides (1970), for example, employed star patterns varying in the radius of the smallest *enclosing circle*, *radius of the largest enclosed circle*, *number of* points, external interpoint angle, and interior point angle. Sixty-six pairs of such stimuli were rated for similarity. Results indicated that the subjects did not use all features of the stimuli in making the judgments, but instead tended to rely on the number and "sharpness" of the points. Isaac (1970) has indicated that such dissimilarity judgments can be used in conjunction with multidimensional scaling techniques to describe individual differences in perceptual structure. He collected judgments of dissimilarity on pairs of photos which were used to construct multidimensional scale configurations for each subject. The subjects then performed an encoding task (given a target photo, select three others so the target will be "odd" or different in the set), and in a second session they "decoded" these sets of photos, selecting the odd photo. In addition to the 12 sets encoded by the subject, he decoded 36 sets encoded by other subjects and 12 sets constructed on the basis of predictions from the scaled perceptual configuration. Results indicated that subjects behaved in the same manner to their own and predicted sets, and were much less successful with other target sets. Isaac (1970) feels that the fact that predicted sets are decoded more accurately than other sets supports the hypothesis of individual differences in perception.

Using a similar technique, Forsyth and Shor (1974) asked subjects to rate similarity of schematic faces differing in 1, 4, 7, or 10 dimensions. The similarity ratings concerned expression of interest in one set of faces, and expression of fear in another. They found three groups of subjects differing in their mode of response. Group 1 responded primarily to direction of mouth curvature for both judgments, up reflecting interest and down fear. Equal weight was given to any attribute in both judgments. Group 2, on the other hand, weighted dimensions differently for the two judgments, responding to pupil dilation or interest but not for fear, and mouth direction for

fear but not interest. For Group 3, pupil dilation and mouth direction interacted to influence the interest judgments, but not the fear judgments. Scores of the Shor Cognitive Elements Test revealed a lack of flexibility in response strategy for Group 1 that was not evident for Groups 2 and 3. A similar study by Forsyth and Huber (1976) also revealed homogeneous subgroups of respondents to ambiguous figures. This research suggests that individual differences in perceptual structure may be specified using similarity judgments and multidimensional scaling techniques. The only qualitative statements to emerge with regard to these perceptual structures tend to be negative statements, like the indication of lack of flexibility for Forsyth and Shor's (1974) Group 1. The type of perceptual structure that leads to superior performance is not evident.

An alternative approach to the study of qualitative differences in perception is possible through the examination of the speed with which visual arrays can be searched for a target (see Neisser, 1967). In displays of lines of letters, the speed of search tended to be a linear increasing function of the number of lines to be searched, with a slower search rate when the array consisted of letters with low discriminability (Neisser, 1964). Atkinson, Holmgren, and Juola (1969) confirmed that the slope of the search rate was a linear increasing function of the size of the search set. Chiang and Atkinson (1976) investigated individual differences in two parameters of the visual search task. The slope represents time to encode the search item and compare it to the target, while the intercept represents the duration of the binary decision and response production. Their results were differentiated on the basis of sex (although the variance of group means on these parameters indicate large individual differences), and showed a significantly faster visual intercept for men. Moreover, correlations with Scholastic Aptitude Test (SAT) scores suggested that high SAT scores were associated with fast search rates for men, but with slower search rates for women. Results such as these validate the proposed processes represented by the parameters (i.e., encoding, comparison, decision, and response), and suggest that individual differences in the two parameters may reveal important variations in these underlying processes as well as the effect of these variations on overall performance.

Researchers have attempted to develop direct methods for measuring perceptual speed. For example, Hunter (1975) employed a task in which he measured the time from the onset of a stimulus to the initiation of a motor response. The task demonstrated reasonable predictive validity, and warrants inclusion in a test battery. However, it seems likely that even this simple task includes a decision speed factor, as well as a perceptual speed component, contributing to the latency.

In addition to these perceptual differences, the limitations imposed on perception by attentional processes must also be considered. The concept of attention is a complex one, which must take into account the ability to process one signal without

interference from others (selective attention), the ability to process two or more signals simultaneously (divided attention or time-sharing), and even the ability to maintain selectivity over time (vigilance). Each of these aspects of attention is a potentially important source of individual differences which might prove to be effective as part of a selection instrument.

The study of vigilance typically involves measuring the ability to detect signals in noise, as when monitoring a radar scope for the occurrence of a signal. Broadbent and Gregory (1963) have presented evidence that in the monitoring situation the behavior of subjects corresponds to a statistical decision regarding the probability that a signal has in fact occurred. Specifically they employ the model of signal detection theory which suggests that partially overlapping distributions of noise and signal plus noise exist, and that the likelihood of detecting a signal for any subject is determined by the degree of overlap of the two distributions (how discriminable is the signal from the noise) and the subject's response criterion (is it more important to never miss a signal, at the risk of responding when no signal occurs, or to avoid responding to no signal at the risk of missing some signals). This model assumes that the variance of the noise distribution is equal to that of the signal plus noise distribution. It has typically been assumed that individual subjects perform according to this model, although individual data tends to be noisy and unreliable, and the support for the model tends to be based on averaged group data.

Craig (1977) has reanalyzed a number of vigilance studies in terms of the performance of individual subjects. He finds that while 50% of the subjects respond to the manner predicted by the signal detection model, a substantial portion (30%) perform consistently in a manner not predictable from signal detection theory. The performance of these variant subjects reflects a lower likelihood of detecting a low probability signal. It seems clear that differences in the pattern and criterion for responding to signals in noise do exist, in spite of the general unreliability of the task for detecting such differences.

A primary technique used in the study of selective attention is shadowing, in which subjects are presented with two simultaneous messages and required to repeat one message aloud as closely as possible while ignoring the other message. In a previously described study, Gopher and Kahneman (1971) adapted the shadowing task by presenting streams of digits dichotically and requiring a report of digits from the ear designated as relevant. They found significant individual differences in the number of omitted digits, as well as the frequency with which digits from the irrelevant ear were reported, and successfully related these differences to various flight criteria.

While selective attention is conceptualized as an ability to focus on relevant messages, it has a complementary aspect which involves the ability to resist interference from irrelevant dimensions. Stroop (1935) developed a test employing three cards; a

list of color names printed in black (w), a set of color patches (c), and a list of color names printed in incongruent colors (cw). it typically requires a longer time to name the color patches in the cw list than in the c list, indicating that the word interferes with color naming. Rose (1974) has shown that the cw-c comparison is a reliable measure of a factor of interference, and that individual differences in susceptibility to interference can be obtained (Jensen, 1965). Pew and Adams (1975) included this cw-c subset of the Stroop test in their test battery, and although it showed rather low correlations on its own with flight criteria, they felt it contributed significantly to the power of the test battery (Pew, Rollins, Adams, & Gray, 1977).

Whereas the selective attention paradigm usually requires the individual to process only one of two or more simultaneous stimuli, divided attention or time-sharing requires the subject to process all of the inputs. In the dichotic listening task employed by Gopher and Kahneman, if subjects had been required to report the digits from both ears they would have been dividing attention. Another common procedure for evaluating time-sharing ability is the subsidiary task technique. Subjects might be required to perform two tasks simultaneously to a specified criterion level for both tasks, in which case the ability to allocate processing capacity would be critical. This was the approach taken by North and Gopher (1976) as described in another part of this review. Alternatively, subjects might be told to perform the primary task as well as possible and only perform the subsidiary task when it will not affect primary task performance. This approach was utilized by Damos (1978), as also described elsewhere in this report, and constitutes an assessment of the overall processing capacity of the individual in terms of secondary task performance level. Both of these tasks have been shown to predict performance on various flight criteria.

An alternative approach to considering the amount or allocatability of processing capacity is to consider the possibility of different types of capacity. Das, Kirby, and Jarman (1975) have presented a version of Luria's information-processing theory which assumes two modes of processing--simultaneous and successive processing. They have shown that these modes of processing cut across various abilities, such that both types of processing may be involved in memory, and they have shown that spatial abilities are particularly dependent on simultaneous processing (Kirby & Das, 1978). An implication of this point of view is that measures of time-sharing ability or processing capacity may be misguided, since more than one "capacity" might be involved. Hawkins, Rodriguez, and Reicher (1979) conducted dual task reaction time tests which suggested that the ability to time-share was task-specific--that is, tasks with more components in common yielded higher correlations between performance levels. Similarly, Navon and Gopher (1979) have argued convincingly for the existence of multiple resources, and claim that the ability to time-share depends on which resouces are required for performance of the tasks. For the purposes of a test battery for pilot selection, however, tasks implicating a time-sharing ability have already been

shown to be predictive of flight criteria (Damos, 1978; North & Gopher, 1976; Pew, Rollins, Adams, & Gray, 1977). The importance of the multiple resources argument is to underline the need for face validity of the dual tasks with regard to piloting, so that time-sharing is being measured using the appropriate resources.

The literature on cognitive perception reveals a number of tasks which yield individual differences in performance. Multi-dimensional scaling techniques reveal individual differences, but do not in general suggest that a particular perceptual structure would lead to superior performance on piloting tasks. Tests of perceptual speed and visual search tasks do appear to be appropriate as candidates for a pilot selection battery. Attention research has yielded a rich variety of tasks which have apparent relevance to the function of piloting. The dichotic listening task used by Gopher and Kahneman (1971) has proved valuable as an index of selective attention, and a variety of dual-task measures are useful as indices of time-sharing (Damos, 1978; North & Gopher, 1976). In addition, vigilance monitoring tasks have resulted in individual differences in performance, as has the Stroop test for susceptibility to interference. Both of these processes are apparently relevant to piloting, and the tasks are potential candidates for inclusion in a test battery.

Encoding. The initial step in the processing of sensory input is to convert the raw input to a format useful for the processing system. This conversion or encoding of the input is often conceived as an interative process, in which sensory data are successively recoded and elaborated into more complex and meaningful forms (Craik & Lockhart, 1972). The concern of this portion of the review is to examine differences in the performance aspects of encoding, i.e., the speed of encoding, and flexibility in employing alternative coding formats.

Studies of sensory storage indicate that physical stimuli are held very briefly beyond the time of their occurrence, in an unencoded form. This brief prolongation of the stimulus allows the system opportunity to encode more information, and also allows for an assessment of the speed with which information is being encoded. For example, Sperling (1960) presented a 3x4 matrix of letters to subjects for only 50 msec, and found that subjects typically recalled about four letters. However, when cued with a tone to recall one row of letters, subjects recalled the entire row, indicating that the entire matrix was present immediately after presentation. Recall of the cued row decreased as the lag between matrix offset and cue onset increased, but wide individual differences were observed in the delay function. Of the four subjects in the study, the best showed 85% recall after 150 msec while the poorest showed only 60% recall. Similar results have been obtained with auditory echoic memory research (Moore & Massaro, 1973).

Posner and Mitchell (1967) refer to the successive recoding of the stimulus as the process of abstraction, and they developed a

technique to measure directly the amount of time taken for each successive recoding. They presented subjects with pairs of letters to which they made a timed same-different response based on physical identity (AA vs. Aa), name identity (Aa vs. Ab), or rule identity (both vowels--Ae vs. Ab). Posner and Keele (1967) report that reaction-times increased with the increasing complexity of classification, with an average of 50 msec more required for a name match than for a physical match. Moreover, individual differences appear quite large; Hunt and Lansman (1976) calculate that 85% of the variance in the name-access effect was not associated with the average trend and hence was due either to measurement error or individual differences.

Hunt, Lunneborg, and Lewis (1975) have related measures of encoding speed to more general cognitive abilities. They find that students who score in the upper quartile on verbal intelligence tests are 35% faster in the Posner matching task than are those who score in the lower quartile. High verbal students also read out more letters to immediate report in the Sperling iconic task, and are better able to report dichotic stimuli by category than are low verbals. They suggested that "high verbal" subjects may simply be faster encoders. (See, however, Hogaboam & Pellegrino, 1978, discussed below.) Rose and Fernandes (1977) utilized a modified version of the Posner task, and replicated the previous findings as well as establishing reliable parameters for individual differences investigations.

Hogaboam and Pellegrino (1978) have utilized the classification task for category verification of lexical and pictorial stimuli. They presented subjects with single stimuli every 5 seconds and required a response indicating whether or not it fit in a specified category. The stimuli varied in format (word or picture), taxonomic and printed frequently, and type of decision (yes-no). They replicated previous results in this type of research showing faster reaction times to pictorial stimuli, with positive responses, and for high frequency category examples. However, all of their measures of processing speed showed no relationship to verbal ability as assessed by SAT verbal scores. This is in direct conflict with the results of Hunt et al. (1975). Hogaboam and Pellegrino suggested that their task was more representative of normal processing activities than the name-matching task, and that the differences obtained by Hunt et al. may reflect greater flexibility of processing, allowing high verbal subjects to adapt more readily to the unusual processing demands of the name-matching task.

A somewhat different approach to studying differences in encoding processes has been employed by Baron (1973). He argued that in identifying a printed word (i.e., abstracting from the printed stimulus to its semantic representation) two strategies are available. One strategy requires first translating the stimulus to a phonemic representation, then using that representation to retrieve the meaning of the word. The other strategy involves the use of the visual representation itself to retrieve meaning. Since this second strategy requires fewer steps in processing, it results in faster

identification of the word and should be the preferred strategy. Baron and McKillop (1975) employ a task in which subjects were required to classify sentences as meaningful or not. They used three types of sentences: sensible (S), e.g., "It's not so," homophone (H), e.g., "It's knot so," and nonsense (N), e.g., "It's boat so." They arranged the sentences into three lists containing two types of sentences each--SH lists, HN lists, and SN lists. They reasoned that a phonemic strategy would be most effective in selecting the homophone sentence as meaningful in HN lists, while the visual strategy should be most effective for SH lists. Since either strategy should be effective with SN lists, they predicted that those employing a visual strategy would be faster since that strategy should be intrinsically faster. They compared the five most "visual" subjects (fastest on SH lists) to the five most phonemic (fastest on HN lists) and found that the subjects using a visual strategy were much faster on the SN lists. While the visual strategy is faster in this case, it is possible that when the material is more difficult a phonemic strategy would be superior, since it should result in a code that is retained more easily in memory while the word is accessed.

Rose and Fernandes (1977) employed a modified version of the task used by Baron and McKillop, and although the results are not directly comparable (Baron and McKillop report only the extreme subject's scores) they appear to replicate the primary results. Moreover, the very high reliability found for the parameters in this task suggest that it would be a good task for revealing individual differences in processing strategy that affect speed.

Encoding research has been concerned both with the speed with which encoding and recoding occur, and the flexibility with which subjects employ different codes. Speed of initial encoding may be inferred from the delay functions in a Sperling-type sensory memory test, and the letter-matching task of Posner and his associates assesses recoding speed with demonstrated reliability. The Baron task assesses the flexibility of encoding in subjects with high reliability, and also suggests the importance of imaginal coding for speeded performance. The use of imagery has, in fact, resulted in a number of individual differences in performance.

Imagery. An important aspect of the encoding process involves the fact that encoding may take place in more than one format, and the format utilized may have consequences for performance. This fact has been alluded to indirectly, discussing the superiority of visual encoding in the work of Baron and McKillop (1975) and in that of Hogaboam and Pellegrino (1978). Research on imaginal coding presents some confusing results, but facility with such coding may be crucial to pilot performance.

One of the difficulties involved in assessing research on imagery is the variety of ways in which the construct of imagery is operationalized. In general, differences in the ability to form and use an image code are assessed in one of two ways: (a) self-report ratings of the vividness of images and (b) performance on spatial

abilities tests such as the Minnesota Form Board Test. Although these two approaches distinguish subjects who show performance differences as a function of imagery, they do not appear to measure the identical process (DiVesta, Ingersoll, & Sunshine, 1971). The problem is further complicated by the definition of imagery in terms of instructional set, or in terms of the type of material (concrete or abstract) used in the task. Nonetheless, performance differences are evident under a variety of definitions of imagery.

Much of the research interest in imagery has been stimulated by Paivio's (1971) dual coding hypothesis; i.e., he suggested that information may be coded verbally and imaginally, but that imaginal coding is superior. Ernest and Paivio (1971) presented lists of mixed concrete and abstract words to subjects, instructing them to (a) form an image, push an RT button, and describe the image, or (b) form a verbal associate, push the RT button, and give the associate. Subjects were rated as high or low imagers based on a battery including both vividness ratings and sptatial abilities tests. They found shorter RT overall for high imagers, as well as for concrete words and for the verbal association instructions. The imaging instructions led to faster RT with concrete words, while the word types did not differ under verbal instructions. The overall superiority of high imagers interacted with word type such that the superiority of high imagers was even greater for abstract words. They suggested that both verbal and imaginal codes are formed with either instruction, but high imagers used both codes and were therefore superior in overall performance.

In order to determine the source of this advantage for high imagers, Griffith and Johnston (1973) had subjects perform a paired-associate learning task in which the responses were concrete or abstract nouns. Subjects received either standard rote learning instructions or imagery instructions, and expended processing was measured during study and test using a subsidiary reaction time test. RT declined over the interval allotted for both study and test, but imagery instructions lowered RT only during study, while concrete response resulted in lower RT only during the recall interval. Apparently the instructions to code imaginally facilitated learning of the word list without affecting retrieval, whereas concrete items were no easier to learn, but noticeably easier to retrieve.

Klee and Eysenck (1973) required subjects to decide whether sentences were meaningful or not, while varying the meaningfulness and concreteness of the sentences and the type of interpolated activity (verbal or visual interference). There were no imagery instructions, and the assumption was that if imagery were used to comprehend a sentence, then visual interference should be more disruptive, while verbal interference should disrupt verbal coding. Concrete sentences were comprehended more rapidly than abstract, but this effect interacted with type of interference such that the advantage for concreteness was much less than visual interference. Similarly, verbal interference appeared to affect comprehension of abstract sentences. When subjects were rated as high or low imagers, high

imagers were significantly faster to comprehend, with the advantage being greater for abstract sentences. They suggested that visual imagery is a superior form of processing resulting in superior comprehension (although Baron and McKillop's caveat regarding the lack of difficulty in the managerial should be kept in mind).

In a somewhat different approach to studying the role of imagery in information processing, Seamon (1972) asked subjects to perform a Sternberg type item-recognition task in which the search set consisted of one, two, or three words. Subjects were instructed either to covertly rehearse the items, to form separate images for each item, or to form a single image in which the items interacted. He found evidence for serial exhaustive search with covert rehearsal and separate images, but the group forming an interactive image employed a parallel search procedure--RT did not increase as a function of set size. The indication is that search and retrieval strategies depend on the type of encoding, and encoding in terms of interactive images is exceedingly effective. Sheehan (1966a, 1966b) examined differences in the effective use of literal pictorial codes (i.e., an image to group and relate non-visual information). He used a pattern construction task, presenting subjects with matrices of 9, 12, or 16 simple geometric forms, which they reproduced using wooden blocks. They were then required to do an incidental task, followed by an attempt to reconstruct the pattern. Subjects were rated as high or low imagers based on their mean vividness ratings on the Betts Questionnaire upon Mental Imagery, and they rated the vividness of the pattern image after reconstruction of each pattern. He found accuracy of reconstruction correlated with vividness ratings within but not between subjects, and found that subjects rated as high imagers increased their vividness ratings with experience while low imagers did not. It is surprising that accuracy and vividness are related only within subjects, but this may be because vividness ratings were collected after construction. Subjects might tend to rate the imagery as more vivid when they feel their reconstruction has been more successful.

Marks (1972) modified Sheehan's procedure by using a more interesting picture-memory task, by collecting vividness ratings prior to performance, and rating subjects as high or low imagers based only on a visual imagery inventory (the Betts scale employs images for all senses). He found superior performance for high imagers as well as an effect favoring female subjects, and no evidence of a within-subject correlation between vividness ratings and performance.

McKellar, Marks, and Barron (in Marks, 1972) investigated the relation between imagery ability and memory, hypothesizing that high imagers would make more effective use of an image mnemonic (method of loci). They required subjects to take a walk, noting landmarks, until the walk was well-learned. They then instructed subjects who were learning a serial word list to form an image associating each word with a successive location on the walk. The use of this technique significantly improved memory, but low imagers used it just as effectively as high imagers. The authors warned that imagery may not

be a stable trait, that all people may have the potential for high imagery and be able to use it when required. It seems likely, however, that a consistent preference for image-coding could be facilitative in a number of situations.

Further evidence that the use of image-coding is flexible is provided by Tverskdy (1973). She presented subjects with simple line drawings, under instructions to prepare for a recall by grouping the items, or to prepare for a recognition test by presumably encoding visual information for discrimination. She then gave both tasks to both groups, and found recall performance enhanced by appropriate instructions, as well as enhanced recognition performance with recognition instructions. She argued that the low correlation between correct items on the two tests suggests that subjects used both grouping and pictorial encoding strategies and used them with differential effectiveness according to instructions.

An evaluation of the literature on individual differences in imagery is clouded by the variety of definitions of imagery and by the variability of results as a function of task demands and instructional set. Nonetheless the use of imagery appears to facilitate comprehension (Klee & Eysenck, 1973) and recall (Ernest & Paivio, 1971) and would appear to be important when any kind of visual comparison and discrimination were required. One possible technique for assessing imagery is to employ a comprehension or recall task varying materials and/or instructional set, but typically responses to an imagery questionnaire are used to predict performance in those tasks. Similarly, the reconstruction task employed by Sheehan uses imagery ratings to partition subjects. The variation of the Sternberg task employed by Seamon is a plausible alternative, since the slope of the reaction time function can be used to infer imaginal coding. However, the evidence for individual differences in this paradigm is not as strong as might be desired. The use of such a task within a test battery might be justified by the apparent desirability of imagery ability in a variety of performance situations.

Comparison and transformation. One possibility is that the advantage of imaginal coding lies not just in its passive representational advantage, but depends as well on the ability of the subject to transform the image and compare it to some standard, either internal or external. The ability to mentally manipulate and compare information is crucial to performance of any task requiring comparison of incoming information to some standard, in order to evaluate performance in real time. Shepard and his colleagues have developed a technique for examining information processing in the manipulation and comparison of short-term memory (STM) (Shepard, 1978; Shepard & Metzler, 1971). In this procedure, subjects are shown a visual display, usually a letter or a complex geometric line drawing. After a brief inter-stimulus interval, a second stimulus is displayed, usually rotated with respect to the first, and the subject is asked to compare the second to a retained image of the first and then decide if they are the same or not. They typically find that the time required to answer this question is a linear increasing function of the degree

or rotation of the second stimulus. As might be expected, however, there are large individual differences in speed of rotation, and Cooper and Shepard (1973) have noted that the subjects appear to fall into two groups; faster responders and slow responders.

The slope of the reaction-time function in the Cooper and Shepard study does seem to represent the speed with which an image may be rotated, but rotation speed is not the only source of potential variation in this task. Encoding the stimulus and deciding to respond both require time, and their times are reflected in the intercept parameter of the RT function. Egan (1979) presented evidence to suggest that speed differences may be apparent in the intercept parameter. For example, when the RT functions for two-dimensional and three-dimensional stimuli are compared, both slope and intercept are different, suggesting a longer encoding time for three-dimensional stimuli.

Just and Carpenter (1976) recorded eye movements during performance of a three-dimensional rotation task. They were able to decompose performance into subprocesses, including an initial search for corresponding elements, rotation, and comparison processes which included rotation of part of the figure followed by comparison, and a confirmation process applied to remaining parts of the figures. Not only would the speed of each process contribute to variability, but the sequence of application could also be a contribution.

Speed is not the only variable affecting performance in the rotation task. There appear to be styles of comparison analogous to the preferred strategies for encoding that affect performance in the rotation task. For example, Cooper (1976) has examined the differences in visual comparison strategies in detail. She utilized random geometric shapes as targets, with the "diffferent" or foil stimuli being perturbations ranging from very similar to quite different from the original. The second stimuli were presented at one of six rotations at 60^0 intervals, and the interval between the first and second stimuli was .1 sec or 3.0 sec. At the .1 sec inter-stimulus interval (ISI), she found that subjects fell in one of two patterns of response. The first type, whom she calls holistic processors, were faster on same responses and their different responses were not a function of pattern similarity. The second group, called analytic processors, were slower on same responses, and reaction times to different stimuli decreased as similarity decreased. This pattern held at the 3.0 ISI in a second study. In addition, holistic processors were faster overall and error rate was unrelated to RT, while for analytic processors error rate and RT were positively correlated. She suggested that holistic processors make a single, holistic comparison--if it results in a match, fine, but if not, they simply respond "different," and all of the different stimuli are treated alike and have equal RTs. Analytic processors may do the same, except that simultaneously a point by point comparison is undertaken. If no immediate match occurs, RT is a function of similarity--the time taken to locate a feature which discriminates the foil from the original.

Allen, Rose, and Kramer (1978) investigated an adaptation of this comparison paradigm for the presence of parameters sufficiently reliable to test for individual differences in comparison speed. They found the slope of the RT function, as well as the standard deviation and range, to be sufficiently reliable to warrant inclusion in a test battery.

Contemporary research on the process of transforming and comparing stimuli has employed almost exclusively variations of the paradigm developed by Shepard and his colleagues. These variations include the use of two-dimensional vs. three-dimensional stimuli, letters vs. line drawings vs. geometric figures, and comparison of two stimuli when both are present vs. comparing a stimulus to a mental image. The rotation task has been extensively replicated and shown to yield reliable individual difference parameters, and is therefore a likely candidate for inclusion in the test battery. Moreover the evidence suggests that variants to the task, such as comparing performance on two-dimensional vs. three-dimensional stimuli, may yield additional valuable information.

Memory search and capacity. The speed and accuracy with which information in memory can be searched is a major source of potential variation in performance. Sternberg (1966, 1975) has introduced the item-recognition paradigm to allow assessment of the speed of search, along with the search strategy employed. This task usually involves presentation of a set of items to be held in STM, followed by a presentation of a probe item to which the subject must respond by indicating whether or not the probe is a member of the set held in STM. When RT is plotted as a function of set size, a linear increasing function occurs analogous to that found in the visual scanning paradigm. Again the available parameters are slope, indicating the speed of the search or comparison process, and intercept, indicating the duration of the encoding, decision, and response processes. Typically individual variations are so great that Sternberg (1966) reports individual subject data. Rose and Fernandes (1977) replicated the Sternberg work using target sets of 1 to 4 digits and, in addition, found a moderate level of reliability in the slope and intercept parameters.

The existence of parallel functions for "yes" and "no" responses in the item-recognition task has usually been taken as evidence that STM search was serial and exhaustive in nature. Further research has questioned the probability of this interpretation, however. Clifton and Birenbaum (1970) examined serial position effects in this paradigm by varying the number of digits that preceded and followed the probed item in the target set, as well as varying the delay of probe onset. They found that, at very short delays (.8 sec) for probe onset, probe items near the end of the target set were responded to more quickly, given an equal target set size. This recency effect indicates a failure to search exhaustively under some conditions. Moreover when individual data were examined, three subjects showed the slope for "no" responses to be twice that for "yes" responses, also indicating self-terminating search. The other nine subjects had the usual

parallel functions indicating exhaustive search. Sternberg (1975) has suggested that exhaustive search is efficient because the search is so rapid that it is faster to examine all items before deciding on a response than to decide after each item. However, as the search rate becomes slower, the advantage of exhaustive search disappears and self-terminating search becomes more likely.

Juola and Atkinson (1971) also extended the item-recognition paradigm to include the cases in which the items were not digits, but words and categories. They also found RT to be a linear function of set size, but noted that the slope was much greater when the set was category labels and the correct probe a category exemplar. While this variation offers the potential for examining further individual differences, Rose and Fernandes (1977) found very low reliability estimates for the slope and intercept parameters in this task.

Sternberg also noted (1966) that when the required response in this paradigm was the location of the probe in the test list, search times were much slower and self-terminating. Anders (1971) examined the process of search and retrieval in this paradigm, using a search list of 12 items, and requiring a timed location response as well as a verbal description of the search and a confidence estimate. In this paradigm, probability correct was only .69, with a slow, self-terminating search process. Interestingly, organizaton at the time of storage is crucial to understanding the search process. Of the 12 subjects, five organized the list into four groups of three items, and five organized it as three groups of four items. They then accessed the first item of a group (item 1, 4, 7, or 10 for the first group, and 1, 5, or 9 for the second) and searched serially through that group. Variations from self-terminating search reflected a tendency to exhaustively search the accessed group. Moreover, since 12 items exceed most estimates of STM capacity, a separate analysis was presented for the first six times (LTM) and the last six (STM), which suggested that STM is searched about twice as fast as LTM under these conditions.

In addition to studies focusing on the variations in parameters in the item-recognition test, attempts have been made to show that scanning rate is related to other measures of STM performance, particularly the span of immediate recall. In the Chiang and Atkinson (1976) study mentioned above, subjects not only performed a visual search task but also a memory search and digit span task. They found the slopes in the two search tasks to be highly correlated (.83) and the intercepts as well (.97), with the slope of the visual task slightly higher and the intercept of the memory task slightly higher (reflecting the fact that the encoding time shows up in the slope for the visual task and the intercept for the memory task). While the correlations suggest that the same processes are involved in the visual search and memory search tasks, the digit span results were not signficiatnly correlated with any of the slopes or intercepts. Their results do show a correlation between memory span and the slope and intercept parameters for male subjects, however, while the non-significant correlations for female subjects tend to be in the

opposite direction. Cavanaugh (1972) has made the relation between scanning and memory more explicit. He assumed a fixed capacity for the more complex the item, the shorter the span. By comparing span size and scanning rate for different types of items, he showed that time per item to scan is proportional to space per item for recall. That is, smaller memory span and longer scanning times may both be a function of the fixed capacity of STM.

One of the most frequently used techniques for assessing memory capacity has been the memory span task. Research has focused on the digit-span task, both for its potential to reveal important psychological processes and as a result of its pragmatic role in intelligence testing. The importance of digit-span for revealing individual differences has been taken for granted since the inclusion of the task in the first intelligence tests by Binet. Nonetheless the source of these differences has remained a topic for investigation, with results such as Cavanaugh's (1972) suggesting ties between memory search and capacity, while Chiang and Atkinson (1976) show no relationship between search parameters and digit span.

Whimbey and Leiblum (1967) investigated the effects of intervening activity on digit span performances. They presented subjects taped sequences of five to nine digits at two digits per second, and required either no intervening activity, simple activity (write the list suffix 0 prior to the digit list or complex activity (examine six word-letter pairs, and when a word is spoken following the digits, write the appropriate letter response before recalling the digits). As might be expected, the intervening activities depressed recall, but intercorrelations on the three tasks were very high, suggesting that the same process or processes were involved in each case.

The existence of a modality effect in immediate serial recall (auditory presentation superior to visual in immediate recall, visual presentation superior at delays) has often been attributed to better rehearsal for visually presented materials since the visual presentation does not interfere with acoustic rehearsal, while at no delay, auditory inputs are easier to recover. Jensen (1971) investigated the possibility that span performance was affected by modality, and perhaps indirectly that span was related to effectiveness of rehearsal. He presented control groups with either auditory span tasks or visual span tasks only, on consecutive days, and used them to estimate the reliability of the tasks. The experimental groups received one modality task the first day and the other modality the second. Noting Kay's (1958) finding that, in a paired-associative task, 12% of subjects were significantly better with visual presentation and 4% significantly better with auditory presentation, Jensen reasoned that differences in span might rely on differences in the efficiency of usng a single modality. If so, one should expect correlations to be significantly less than perfect between auditory and visual presentation. His results replicate the modality effect, but the correlation between auditory and visual

performance was almost 100%, suggesting that span performance relies on the same processes in either modality.

Searching for the processes underlying individual differences in immediate serial recall, Lyon (1977) hypothesized that effective use of various mnemonic devices might cause the variation. He reasoned that if rehearsal were the crucial underlying process, presenting the stimuli too fast to allow rehearsal should either eliminate or reorder the differences in performance. Similarly, if chunking or organizational activities underlay the differences, forced grouping of the digits in groups of three should eliminate or reorder them. His tests revealed neither of these hypotheses to be supported, and he suggested that differences in span might be related to speed of stimulus identification (Huttenlocher & Burke, 1976). That is, an item which is identified more quickly should have more time in which to consolidate and hence a higher probability of recall.

Martin (1978) has pointed out that papers related digit-span to other psychological processes (Cavanaugh, 1972) have done so across materials rather than across subjects. She investigated whether digit-span was related to other processing measures, including immediate and delayed free recall, estimates of short-term memory and long-term memory capacity derived from the recall tasks, and an order recall task in which 12 digits were presented and recall required by location. She found that digit-span performance did not correlate with free recall performance or with the theoretical estimates of STM and LTM capacity, suggesting that differences in digit-span did not result from differences in item memory or memory and processing capacity. The significant relationship between digit-span and order recall suggested that ability to retain order information might underlie differences in digit span.

Immediate memory span has also been shown to be related to performance in more complex tasks. Martin (1968) trained subjects in a simple concept learnng task in which pairs of letters were presented varying on four dimensions (letter name, case, number of letters, and position relative to center), of which only one dimension was relevant. When subjects overlearned the initial concept, those who made few errors in initial concept learning performed more poorly on a shift task than those who made many errors on initial learning. Those who made more errors on initial learning showed enhanced shift performance, and a larger memory span for relevant attributes. This difference in span was not evident when measured after initial learning, and suggested a strategy difference in which slower initial learners lengthened their span for relevant attributes in shift learning and hence performed better than those who learned the initial task quickly.

In summary, the relationship between rate of memory search and memory capacity is unclear, but the investigation of each has resulted in tasks which reliably index individual differences and appear to tap important cognitive processes. The memory search paradigm as

developed by Sternberg clearly suggests itself as an important candidate for inclusion in a test battery. While variants of the Sternberg task have not always produced reliable parameters (Juola & Atkinson, 1971), some variants which do not produce error-free performance might also be considered. Tasks which probe for location information, or which use larger target sets and access longer-term memory might prove to be additionally informative.

The suggestion that search rate and digit-span performance are both determined by memory capacity has received mixed support. While correlations between performance levels on a variety of tasks indicate a tie between the task and capacity (Cavanaugh, 1972), Chiang and Atkinson (1976) show no such relation. Span does not appear to be a function of rehearsal or organizational strategies (Lyon, 1977), but may be related to speed of stimulus identification (Huttenlocher & Burke, 1976) or to retention of order information (Martin, 1978). It is unclear whether span tasks are indicative of processing capacity, but they are reliable and related to processes of apparent importance, so as to warrant inclusion in a test battery.

Memory organization and retrieval. Whenever people are presented with information which they need to use, but which is too much to hold in consciousness, they must attempt to store that information in LTM. Mandler (1967) has argued that storage in LTM does not merely involve consolidating a memory trace, but rather involves the active organization of the material into a format that will facilitate retrieval. While the form of this organization is often inferred from the pairwise consistencies of recall output, Mandler, Worden, and Graesser (1974) presented evidence that the organization is more complex than simple pairwise associations. Specifically, they argued that the organization is hierarchical in form, with a limit of about five subordinate nodes accessible from each superordinate node (this limit seems likely to be related to STM capacity, either in its role as a coding device or as a response device). While organization may well take place at the time of storage, the form of organization is inferred from retrieval patterns, and the function of organization is presumed to be the facilitation of retrieval.

Earhard (1970, 1974) has extensively investigated individual differences in the ability to organize, and the role those differences play in performance on various memory tasks. In her studies, she used a measure of subjective organization to pre-classify subjects as good or bad organizers. Earhard and Endicott (1969) compared serial anticipation performance, on which both inter-item associations and order information were available, to paried-associate performance dependent only on inter-item associations. They found superior performance for good organizers with a preferred list order in serial anticipation learning, but no advantage with the non-preferred order (preferred orders were frequent output orders by other subjects while non-preferred orders never occurred in other subject outputs). In the paired-associate task, however, the good organizers were superior with both list orders, learning the preferred pairings more rapidly, and increasing their advantage over low organizers over trials. They

suggested that the advantage in recall for high organizers rested on their ability to rapidly form and utilize available inter-item associations. The only time this superiority faltered was with the non-preferred list order in serial anticipation, which was probably learned on the basis of order information.

Earhard (1970) investigated the possibility that this advantage for good organizers might be due simply to superior STM functioning allowing subjects to maintain the stimulus array more successfully, rather than to the ability to form inter-item associations. She presented subjects with a 16 trial free recall of 22 unrelated words, and scored the outputs for subjective organization to divide the subjects into good and poor organizers. Both groups then performed a short-term retention task (Peterson & Peterson, 1959) in which recall of three consonants was measured after brief filled delays of 0-18 seconds. She found the usual decline in STM performance as delay increased, but there were no differences between high and low organizers on STM performance. This finding suggested that organizing ability did not depend on differences in STM performance.

Earhard (1974) also investigated the relationship of organizing ability to performance on transfer tasks. Tulving and Osler (1967) had shown that when subjects practiced on a part of the list prior to learning the complete list (part-to-whole transfer), or practiced on a whole list before learning a subset of the list (whole-to-part transfer), learning took longer than for a control group which did not practice (i.e., negative transfer obtained). Tulving and Osler concluded that subjects form an appropriate organization during practice which interferes with learning. Earhard reasoned that if that were so, good organizers should suffer more from the inappropriate organization formed at practice. On a whole-to-part transfer task, good organizers showed more organization and better overall recall on the practice trials. On the transfer task, those subjects showed initial difficulties, but by the fifth trial, their performance was significantly better than that of poor organizers. With part-to-whole transfer, no differences were evident on the practice list, probably due to its shortness. In this case also, however, good organizers showed better transfer, with the difference emerging over trials. While negative transfer was obtained when compared with control groups, skill at organizing did not distingush different levels of the effect. Poor organizers were equal to good ones on early trials, but the better organizers showed a higher final level of performance. These results suggested that the subjective organization measurement relates to a more general memorial ability than simply reflecting a tendency to remember words in the same sequence from trial to trial.

Another approach to the study of differences in long-term memory processes was developed by Collins and Quillian (1969). They assumed that information was organized and stored in LTM hierarchically, in a fashion analogous to that described by Mandler (1967), and that the speed of any search and retrieval process would be a function of the location of the information in the hierarchy. They required subjects

to verify sentences such as "a canary is a bird" (subset relation) or "a canary has wings" (property relation), and found that reaction times increased as the subset relation became more remote. Moreover, retrieving a property from any node required an extra amount of time. While this task potentially provides an opportunity to measure individual parameters, both in terms of speed of movement between levels of the hierarchy and speed of retrieval of properties at any given level, there is a drawback. We must assume that the experimenter-constructed hierarchy is congruent with that possessed by the subject. Moreover to simplify measurement, Collins and Quillian made the assumption that properties are stored only once, at the highest appropriate level (so the property "has wings" is stored at the level "bird" rather than "canary"). This assumption does not appear to be met consistently by individual subjects. Rose and Fernandes (1977) replicated the Collins and Quillian results and examined the parameters of slope and intercept of both subset and property sentences for reliability in revealing individual differences. While the intercept parameters were reliable the slope parameters showed very low reliability, although slope parameters should be most crucial for assessing search and retrieval speed.

Underwood, Boruch, and Malmi (1977) investigated a number of memory tasks, looking for individual differences on a variety of attributes. Fernandes and Rose (1978) replicated many of these tasks, looking for reliable parameters of individual differences in specific memorial functions. Fernandes and Rose examined the following memory tasks: (a) free recall of concrete and abstract words, (b) continuous recognition of words as a function of lag or delay since prior presentation, (c) interference susceptibility in which word-number pairs were presented in lists of five and six lists were presented with the same items re-paired for each list, (d) situational frequency in which a word list was shown, (e) list differentiation in wnich subjects were shown three lists, then asked to judge on which list each item occurred, and (f) a memory span test. Of these, the continuous recognition test (Shepard & Teghtsoonian, 1961) replicated the previous work by Rose and Fernandes (1977). While they obtained reliable individual difference parameters with all these tasks except the interference susceptibility paradigm, it should be noted that the intertask correlations were quite high, suggesting that they may all relate to some general memory function rather than specific functions, such as frequency and temporal coding for the situational frequency and list differentiation tasks respectively.

Hunter (1975) employed a recall task which presented subjects with a sequence of geometric figures at 5-second intervals and required them to indicate the figure that was presented two items (15 sec) prior to the presented one, thus potentially tapping a long-term retrieval process. This paradigm demonstrated very high reliability and loaded heavily on a figural memory factor.

The literature on memory organization and retrieval suggests a variety of tasks in which individual differences in performance are evident. Earhard has amply demonstrated that organizing ability is

not dependent on pure STM functioning, but is related to performance on a variety of memory tasks. Collins and Quillian's work presents an important technique for assessing retrieval, but the crucial slope parameter has proven unreliable for assessing individual differences. Work which has evaluated tasks implicating several aspects of retrieval performance has tended to show a high inter-task correlation to performance. It seems likely that a few tasks which provide several measures would be most appropriate for a test battery. For example, a multi-trial free recall task would allow assessment of initial retrieval, learning, and organiation, while a continuous recognition task provides more sensitive retrieval measures as well as the possibility for evaluating response criteria through signal detection measures.

Learning and comprehension. Fleishman (1953) has suggested that the rate of acquisition of motor skills might be an important individual difference variable for personnel selection. It seems likely that the rate at which cognitive information is acquired might also be important. Moreover, learning new information presupposes comprehension of the input, such that learning rate may well depend on the speed and accuracy of comprehension. A variety of research approaches have been employed in an attempt to understand these processes.

Educational psychologists have been concerned with learning rate, from the point of view of specifying techniques to increase the learning rate of slower school children. They typically employ analogs to classroom materials, such as a subsection on multiplication, and measure the time required for the children to achieve mastery. To check the efficacy of an alternative teaching strategy for increasing the learning rate, Anderson (1976) administered three successive programmed units on matrix arithmetic to eighth grade students. Some students received the units under a normal learning procedure, while others received mastery training (in which tutors pointed out errors and gave directed reviews and retests until an 85% criterion was met). She found that the mastery group took significantly longer to reach criterion on the first unit, and on the second unit although the difference was smaller. By the third unit, elapsed time for the groups was not significantly different, suggesting that learning rate is an alterable characteristic. Yen (1978) tested a large group of students, grades 5 to 10, on a paired-associate task using noun-CVC pairs and on a definition-learning task. She then attempted to fit individual subject data to the parameters of a specific information processing model, and found reliable differences in the acquisition rate parameter--how much information was encoded into LTM per trial. While it seems likely that learning rate parameters could be estimated from any number of memory tasks, there is some question of the construct validity of acquisition rate since it appears to be a readily modifiable function.

A different approach to the study of learning is concerned with the comprehension of verbal materials. One technique for examining

comprehension is the linguistic verification task employed by Clark and Chase (1972). They presented subjects with a simple sentence, followed by a simple picture, and required subjects to indicate as quickly as possible whether or not the sentence contained the same information as the picture. These sentences varied in linguistic complexity, with false and negative sentences being more complex than true and positive ones.

Clark and Chase argued that the sentences were represented as base forms plus markers; i.e., the sentence "star below plus" was represented as the base form "star above plus" and the marker "no." Transformations of this type were based on Clark's theory of semantic features. Rose and Fernandes (1977) replicated the Clark and Chase study, and found that the four indicated parameters accounted for nearly all of the variance in performance. Unfortunately, only two of the parameters showed adequate reliability. Pew and Adams (1975) also used a simplified version of the linguistic verification task as part of their test battery, with results sufficiently encouraging for them to conclude (Pew et al., 1977) that the task warranted inclusion in the test battery.

The base processes involved in this task are representation of the sentence, representation of the picture, and a comparison of the two. Clark and Chase (1972) assumed a specific type of representation involving a deep structure plus markers indicating additional information. Carpenter and Just (1975) have assumed a propositional format of representation, and a component by component comparison of the sentence and picture representations. In this case, the single parameter representing the speed with which components are compared would become crucial, since both sentences and pictures would be represented in the same format. McLeod, Hunt, and Mathews (1978) gave subjects a modification of the verification task and attempted to fit individual subject verification times to the comparison rate parameter. They found that subjects fell into two groups: one group appeared to fit the model perfectly, apparently using a propositional format of representation, while the other group did not fit the model and appeared to use a visual-spatial format of representation. This suggestion is strengthened by noting that the second group performed much better on psychometric tests of spatial ability. The point is that parameters from a model of the linguistic verification task may be useful in revealing variations in comprehension, but they might also be misleading if it is assumed that all subjects use the same format in comprehension. The McLeod et al. data suggest that either propositional or visual-spatial strategies might be used to comprehend sentences.

A different approach to comprehension has been suggested by Bransford, Barclay, and Franks (1972) and Bransford and Franks (1971). They suggested that subjects do not represent individual sentences when comprehending discourse. Rather, they use the information in successive sentences to construct a scheme representing all of the sentences, but not retaining information about individual sentences. Bransford and Franks demonstrated that when a series of

related sentences were presented--sentences with overlapping information content--then in a later recognition task subjects were unable to distinguish sentences they had heard from those which fit the scheme but had not been presented, suggesting no memory representation for individual sentences. Moreover, subjects displayed increasing confidence that they had heard sentences before when they contained more information, suggesting that sentences are recognized by comparing them to the whole scheme constructed for them and judging the degree of overlap. Allen, Rose, and Kramer (1978) utilized the sentence recognition task of Bransford and Franks in an attempt to estimate reliability for the various measures in revealing individual differences, but the reliabilities were quite low, suggesting that this task might be less appropriate for assessing variations in verbal comprehension. Allen et al. also created a new task to measure the tendency to cluster and integrate information in a sentence recall task. They presented related sentences similar to those of Bransford and Franks and asked subjects to recall the information. They measured the extent of clustering by noting the number of related ideas recalled together. This task also demonstrated inadequate reliability, although its construct validity was quite good.

In summary, research indicates that learning rate may be an important variable, though indications that it is readily modified might suggest that it is more appropriate as a training variable than a selection variable. The techniques used to examine learning rate in educational psychology do not appear to be readily adaptable to a test battery, but Yen has shown that a learning rate parameter may be estimated from simple memory tasks. Such an estimate from a multi-trial free recall task would appear to be warranted.

The comprehension task developed by Clark and Chase has been widely replicated, and reliability for some of the parameters has been demonstrated. The task appears to be an excellent candidate for the test battery, with the potential to reveal preferences for visual or propositional coding as well. Other tasks such as that developed by Bransford and Franks appear to tap important aspects of comprehension, but have failed to exhibit sufficient reliability to justify inclusion in the test battery.

Problem-solving. It is very difficult to conceive of problem-solving proficiency as a single ability or cognitive function. On the contrary, problem-solving research has focused on a wide variety of problem types, and related successful performance to several psychological functions. Moreover, those functions are not always the seemingly obvious ones. Karlins, Lee, and Schroder (1967) gave subjects a set of problems representative of Peace Corps tasks. For example, a community development. problem required subjects to learn about a native culture in order to secure native cooperation in building a hospital. Karlins et al. measured the number of questions asked and number of categories sampled, and gave each subject both an intelligence test and a creativity test. They found that the information-seeking behavior of the subjects was unrelated to intelligence, but strongly related to creativity. They suggested that

problem-solving proficiency may rely on mental fluency or flexibility, especially when the problems are of an unusual nature.

Mandelsohn and Lindholm (1972) suggested a slightly different role for creativity in problem-solving. They had subjects memorize a list of 25 words, while a tape played 25 others, then gave them an anagram solution task using 10 words from the printed list, 10 from the taped list, and 10 new words. They varied the knowledge of relevance of the word list to the anagrams, and varied instructions to attend to both lists (broad attention) or attend only to the printed list (narrow attention). All groups were given Mednick's Remote Associations test as an index of creativity, and the Minnesota Multiphasic Personality Inventory (MMPI)-R scale as an index of repressive defense. They found that a high creativity index and low repressive defense index correlated highly with anagram solution, indicating to them that attending to and storing a wide range of stimulus information was crucial in problem-solving. So it appears that indices of creativity may be indicative of problem-solving facility, both because of the flexibility in encoding the original problem and the fluency of response generation in seeking a solution.

Creativity is not, however, the only function complicated in effective problem-solving. It seems obvious that a variety of information processing functions should be important in problem-solving. Whitely (1977) looked at differences in the ability to solve verbal analogies (prototypes of intelligence test item) as a function of individual differences in various information-processing components. She indicated that response latency in analogy solution could be accounted for by differences in memory accessibility, decision-making, and response implementation, suggesting that even when all the answers are correct, different mental processes might be implicated. There is also a suggestion that, while total response time does not predict accuracy, better analogy-solvers encode very quickly and spend more time on the decision component. Moreover, response accuracy seemed to be a function only of the judgment or decision-making component of the subject.

The importance of a variety of processing functions in problem-solving has been noted before. Hunt and Lansman (1976) summarized a study by Bloom and Broder (1950), which compared the behavior of good and poor problem-solvers on syllogistic reasoning problems. Some of the differences noted were as follows: (a) better problem-solvers isolate key elements of the problem and encode the problem to highlight those elements, (b) good solvers investigated the implications of each statement more systematically, (c) poor solvers did not read or comprehend instructions carefully, and (d) poor solvers sometimes reacted emotionally to the problem content. Thus it would appear that differences in encoding, comprehension, and LTM retrieval would all contribute to variations in problem-solving ability.

Perhaps the most comprehensive analysis of reasoning and problem-solving in the recent literature is presented by Sternberg

(1977), who analyzed several inductive and deductive reasoning tasks in terms of the information-processing components required for solution. The tasks he studied were similar to general intelligence test items, including analogical reasoning, classification, and series completion (inductive tasks), and linear, categorical, and conditional syllogisms (deductive tasks). He suggested a set of processing components as well as meta-components (control processes and strategies which direct the application of the processing components to the problem). The correlations between performance on these tasks, then, are a function of the degree to which the components required for performance overlap, and differences in performance level should reflect the facility with which the processing components are applied.

As an example of Sternberg's analysis, he has suggested that inductive reasoning involves six processing components: encoding, inference, mapping, application, justification, and response. For the analogies test, the subject must encode the problem, infer the relationship between the first pair of words, map the inferred relation onto the third item, apply the relation to potential responses, justify the relation if the application is not perfect for either response, and finally respond with the chosen answer. In the classification task, however, the inferred relation is among all three items and hence a mapping of a higher order relation is required. Sternberg (1977) suggested that individual differences in performance on these types of problems result from differences in the duration, difficulty, and probability of application of these operations. For example, subjects who perform well on the problems spend a relatively longer time encoding the problems, and older children tend to perform the operations exhaustively rather than in a self-termination fashion, resulting in fewer errors.

One of the most important analyses of human problem-solving behavior has been performed by Newell and Simon (1972), using the technique of computer simulation of individual problem-solving behavior. Two important principles emerged from this analysis. First, the external problem must be represented mentally, and the formal encoded description of the problem determines what operations are brought to bear when solving the problem. The second important notion is that of a production system, which is essentially a program for a changing or recoding the problem until a solution is reached. While their studies clearly indicated individual differences in problem-solving procedure, these differences were specified in terms of the production systems or programs necessary to model individual subject behavior, and therefore may be too specific to use to circumscribe a limited set of processes crucial to problem-solving. While the technique of computer simulation continues to be actively applied to problem-solving research, it is unlikely to yield tasks which can be related to those common processes. Therefore, this interesting area of research is not reviewed further at this time.

The review of problem-solving research has indicated that problem-solving is related to creativity, especially in unusual problem situations. The importance of creativity appears to be the

fluency or flexibility it provides, for encoding and recoding the problem and for generating a variety of responses. A variety of studies by Bloom and Broder (1950), and Sternberg (1977) have suggested the importance of other information processing components. While many of these components are discussed elsewhere in the review, it seems likely that some problem-solving tasks capture relatively unique components (i.e., the mapping component for verbal analogies) as well as the superordinate aspect of coordinating and applying the components successfully. Therefore a verbal analogy test is a candidate for the test battery. An Ususual Uses Test does not specifically measure a perceptual-motor or information-processing function, but its ability to index the flexibility of various cognitive processes makes it a potential candidate for the selection battery as well.

Decision making. The term "decision making" has been used to describe a wide range of cognitive behavior. In fact, several different classification schemes have been offered to conceptualize the collection of problem-solving and decision processes which compose decision making. Based on a simplification of the taxonomy provided by Nickerson and Feehrer (1975), the most basic processes involved in decision making include the following: problem structuring, information acquisition, data evaluation, hypothesis evaluation, preference specification, action selection, and decision evaluation. Each of these general subtasks may play an active role in decision making activities required of a pilot during the flight of a mission, which, like for other military decision situations, usually involves "fairly well-defined objectives, significant action alternatives, relatively high stakes, inconclusive information and limited time for decision" (Schrenk, 1969). More specifically, a comprehensive survey of the behavioral science literature dealing with the subject of flying skills, in particular the long-term retention of such skills has led Prophet (1976) to conclude that one of the most critical aspects of flying performance is decision making.

Through decomposition into subprocesses, decision-making performance in a variety of tasks has been subjected to considerable empirical investigation. Comprehensive reviews of the experimental literature are available, for example, in Lee (1971), Nickerson and Feehrer (1975), Rapoport and Wallsten (1972), Slovic, Fischhoff, and Lichtenstein (1977), and Slovic and Lichtenstein (1971). Many of the experiments performed have had the goal of describing human decision behavior and gaining a better understanding of the cognitive processes humans employ to make choices and to solve decision-related problems. The specified tasks studied have often been well-structured to permit optimal or prescriptive models (e.g., Bayesian, Regression, mathematical expectation) to be applied to the same task parameters or input data which are presented to the experimental subjects. By comparing the output provided by subjects with the output of a prescriptive model, the investigator has often been able to determine how reliably, and to what degree, human judgments match or depart from normative judgments. Furthermore, by tracing the processes apparently

used by subjects to arrive at a problem solution, a variety of intuitive or heuristic strategies have been identified.

The most general finding from behavioral experiments that have used the person versus model paradigm is that the intuitive responses of the unaided human mind usually differ from those generated by an optimal model, but they do so in a systematic fashion, depending on personal, situational, and task factors. Thus, borrowing engineering terms to describe human behavior, people have been referred to as suboptimal or inefficient information processors and decision makers. For example, when acquiring information, a person may gather a significantly greater or lesser amount of information--depending on the specific circumstance--than warranted by the optimal model. Or, when evaluating hypotheses by making diagnostic inferences from observed data, people may extract either a significantly greater or lesser amount of diagnosticity--again, depending on the specific circumstances--than warranted by the optimal model. Or, when evaluating hypotheses by making diagnostic inferences from observed data, people may extract either a significantly greater or lesser amount of diagnosticity--again, depending on the specific circumstance--than warranted. Many such deficiencies and/or limitations exhibited by decision makers seem to be due to a variety of biases in the way they process information for decisions (e.g., Slovic, 1976; Slovic & Lichtenstein, 1971; Tversky & Kahneman, 1974). As a result, people are unable to integrate/aggregate/combine various dimensions of information to arrive at subjective assessments or decisions that match the consistency and accuracy of those produced by a normative model.

Although substantial individual differences in the accuracy of decision responses have long been evidenced in the experimental literature, recent research trends have indicated that people also vary systematically with respect to the particular strategies that they use to solve decision problems. Apparently, because of limited memory, attention, reasoning, and computational capabilities, people (without sophisticated mathematical knowledge) cannot appropriately apply optimal models to data parameters; instead, they resort to simplified cognitive-strain-reducing processing strategies. For example, when evaluating the truth of competing hypotheses on the basis of unreliable data, some people may incorrectly treat the data as though they were perfectly reliable (e.g., Johnson, Cavanaugh, Spooner, & Samet, 1973). By applying a variety of methodological techniques to trace and describe processing strategies (see Svenson, 1979), investigators have observed reliable differences in the way individuals approach complex tasks (e.g., Johnson et al., 1973; Payne, 1976). Certainly, the degree of strategy complexity adopted by an individual, and the level of consistency with which it is used, may reflect upon characteristic differences in the way individuals treat information in making real-world decisions. However, such strategy selection preferences can depend on the decision task situation as well as on the decision maker (Beach & Mitchell, 1978; Crolotte & Saleh, 1979; Miller, Rice, & Metcalfe, 1979).

Because of the importance of subjective judgments and values in decision making, consistency--sometimes referred to as internal consistency--has been widely studied as a measure associated with the quality of decision making performance. Irrespective of whether a decision is correct or incorrect on the basis of some external criterion or environmental outcome, the decision may or may not be consistent with the relevant information or values possessed by the decision maker, or for that matter, with other decisions made by the same decision maker. Overall, the empirical findings have suggested that individuals vary markedly in their level of consistency. One common line of research, for example, has examined the degree with which the preference structure of an individual adheres to the axioms of decision theory. A case in point is transitivity (represented in its simplest form by--A preferred to B and B preferred to C implies A preferred to C), where the frequency of intransitive or inconsistent responses has been observed to vary considerably across individuals for simple (e.g., Tversky, 1969) as well as more complex task formats (e.g., Fischer, 1976, Wallsten, 1972). Another typical research paradigm has involved reversals or inconsistencies in the use of two different response modes for eliciting subjective evaluations for the same entity, as for example between bids and choices for bets (Lichtenstein & Slovic, 1971) or between greater-than/less-than judgments and point estimates for the truth likelihood of intelligence reports (Samet, 1975). Here again, differences are found to be large with respect to each individual's proportion of inconsistent responses (e.g., preferring A to B but saying A is worth less in the consistency with which people combine subjective probabilities (Peterson, Hiehla, Miller, & Bourne, 1965) or aggregate probabilities and values into expected values (Lichtenstein, Slovic, & Zink, 1969). In fact, the latter study found subjects to also differ widely in the reasons they give for insisting on being inconsistent with a normative expected value model.

Another key dimension in the evaluation of decision making is decision response time or latency. Because real-world decisions must often be made under time constraints, especially in situations like piloting, time-related parameters of performance are important to study; and various experiments have provided useful findings. For example, in a study involving data evaluation and preference specification, Wright (1974) found significant differences in information processing strategies as a function of time stress; in general, high time-pressured subjects used fewer evaluative dimensions (i.e., simpler strategies) in making their decisions than did subjects who had more time to decide. Furthermore, one's time to arrive at a decision has been found to correlate with his/her confidence in the correctness of the decision and whether the decision is being made for the first or second time (Levine & Samet, 1973; Pitz & Geller, 1970), and the data on which these relationships are based have indicated large individual differences among subjects in their speed of response.

Several general sources of individual differences are thus apparent; the speed and accuracy with which a decision is made, of course, but also the particular pattern in which an individual

deviates from a prescriptive model, the strategies employed by the individual to reduce information load in complex decisions, and the consistency with which the individual makes decisions, all contribute to individual differences in decision making tasks. In addition, specific decision making tasks often yield performance differences as a result of their heavy reliance on one decision making subprocess. For the purposes of the present discussion, several representative decision tasks have been selected, on which performance has been observed to vary significantly across individuals. Taken together, these tasks reflect upon the major component processes that appear to be involved in decision making, including information seeking in structuring the problem, probability estimation in evaluating alternatives, confidence-estimation in evaluating responses, and risk taking.

A prototypical task for indexing information seeking behavior is the urn-sampling task. In this task, a subject is informed that Urn A has 70% white balls and 30% black balls, while Urn B has 70% black balls and 30% white balls. One of the urns is chosen at random and its identity is unknown to the subject, who is required to begin sampling by drawing out balls one at a time, without replacement. After each sample, the subject must decide whether to continue sampling or whether there is enough information at that point to correctly identify the urn. Good performance is indicated by a correct decision based on the minimum sampling, but the task may also reveal tendencies to over- or under-sample.

In real-life situations, the acquisition of information is often costly in terms of time or money, and the decision maker must determine whether the value of the information that could be obtained by a data-collection effort is likely to be greater than the cost of obtaining it--which is a decision problem in its own right. Many different experimental paradigms have been employed in the study of information purchasing behavior; the proportion estimation task used by O'Connor, Peterson, and Palmer (1972) provides a simple example. On the basis of randomly sampled data, the subject must decide whether a population proportion is p or 1 - p (e.g., .6 or .4) with the prior probability for each proportion equal to 1/2. The task is structured so that the subject wins x dollars for a correct decision, loses dollars for an incorrect decision, and pays z dollars for each datum sampled from the population. Given these parameters an optimal stopping point can be computed on the basis of maximizing net expected value.

In their experiment, O'Connor et al. varied the stakes, i.e., the amount to win (x) or lose (y) on a trial, and studied the functional relationship between information purchase behavior and risk. In general, the higher the stakes, the greater the amount of information acquired. However, the data were evaluated separately for each subject, and the level of individual differences was observed to be high. In interpreting their results, the authors lent support to the "portfolio theory" of gambling (Coombs, 1975); namely, that each individual has an ideal level of risk to which he/she aspires.

In order to estimate the probability of events occurring when evaluating a potential decision, the decision maker often needs to acquire base rate and diagnostic information, and integrate the diagnostic information. Large individual performance differences have been observed in the base rate phenomenon (Kahneman & Tversky, 1973; Lyon & Slovic, 1976). A sample base rate problem is provided by Bar-Hillel (1977):

> Two cab companies operate in a given city, the Blue and the Green (according to the color of cab they run). Eighty-five percent of the cabs in the city are Blue, and 15% are Green. A cab was involved in a hit-and-run accident at night, in which a pedestrian was run down. The injured pedestrian later testified that though he did not see the color of the cab, due to the bad visibility conditions that night, he remembers hearing the sound of an intercom coming through the cab window. The police investigation discovered that the intercoms are installed in 80% of the Green cabs, and in 20% of the Blue cabs.
>
> What do you think are the chances that the hit-and-run cab was Green?

This is a characteristic Bayesian problem containing two kinds of information: the background or base-rate information (i.e., color distribution of cabs in the city) and the diagnostic information (i.e., the data about the intercom and its relative distribution in blue and green cabs).

The mathematically correct answer to the above problem computed via Bayes theorem is .41; yet subject responses to this problem have been shown to assume a very flat distribution (for N=35) ranging from .10 to .80 with a median of .48 and a mode of .30 (Bar-Hillel, 1977). As suggested by Fischhoff, Slovic, and Lichtenstein (1978), the wide distribution of responses is not due to a misunderstanding of the problem or some artifact, but rather it reflects systematic biases in whether an individual takes base-rate information into account, and if so, how the information is interrated with diagnostic information. Such differences in information processing strategies may reflect reliable differences among individuals in how they approach real-world decisions given a configuration of probabilistic data.

An important task in piloting and in other decision situations is the utilization of cues (which may vary in diagnosticity and reliability) in order to evaluate an hypothesis or to predict a future event. Prediction experiments have received considerable attention in the psychological literature extending over a wide range from simple tasks like probability learning (e.g., Estes, 1972) to complex tasks like extended stochastic forecasting (e.g., Roby, 1968). One of the simplest paradigms employed to investigate cue-utilization behavior involves numerical prediction. For example, Lichtenstein, Earle, and

Slovic (1975) presented subjects on each trial with a pair of numbers (i.e., the predictors -- x_1, x_2) and required them to predict a third number (i.e., the criterion -- y). Prior to beginning this two-cue prediction task, subjects were extensively trained with the criterion (but they were not instructed on how two cues should be combined). Then, under different parameter combinations (e.g., $x_1 = x_2 = y = 50$, $SD_{x1} = SD_{x2} = 10$, and $r_{x1y} = .80$, $r_{x2y} = .40$) and the cue presentation conditions (i.e., simultaneous or successive), individual performance over a series of trials was compared vis-a-vis the appropriate linear regression model for estimating y from x_1 and x_2; namely,

$$Y = b_1 (x_1 - x_1) + b_2 - (x_2 - x_2) + y.$$

Inspection of the subjective beta weights derived for each subject shows considerable individual differences with the range of weights in each case approximately equal to the mean weight (e.g., $b_1 = .26$ with a range from .15 to .40).

Besides concluding that subjects varied much in how they differentially weighted the cues in making their predictions, the investigators suggested that some subects may have been systematically invoking nonoptimal heuristic information processing strategies; some, for example may have simply used $1/2 (x_1 + x_2)$ to predict y. Given similar findings in experiments involving other decision tasks (e.g., Johnson et al., 1973), further study of individual differences in the choice and utilization of a given information processing or decision strategy is promising.

An important aspect of decision making relates to one's probabilistic assessment or confidence that his/her decisions are in fact correct, at least in situations where accuracy can be objectively determined. The validity of such confidence judgments is often referred to as "calibtation" (Lichtenstein, Fischhoff, & Phillips, 1977). That is, an individual is considered well-calibrated if, over the long run, for all decisions assigned a given confidence level (percent probability for being correct), the actual percentage which turns out to be correct matches this same level.

To illustrate the principle of calibration, consider typical almanac questions:

(1) The only bachelor United States president was:

(a) James Madison
(b) James Buchanan

(2) A rudder is located in an airplane's:

(a) Tail
(b) Wings

For a series of such questions, the subject states a confidence (e.g., 50%, 75%, or 100%) that each answer is correct. For all occurrences of each stated confidence level, the observed or actual percentage of correct answers is computed and compared to that confidence level. The resultant calibration, or realism-of-confidence, function (i.e., confidence level versus percent correct) determines the degree of overconfidence or underconfidence expressed by the subject in his/her decisions.

Based on a survey of the relevant literature and their own comprehensive experiments employing almanac questions and other task problems, Lichtenstein and Fischhoff (1978) point to the likelihood of characteristic individual differences with respect to level of calibration. In fact, these investigators found that about half of their subjects were by nature pretty well calibrated whereas the other half required systematic training to achieve reasonable levels of calibration on certain tasks. Although calibration is important in its own right, it can correlate significantly with decision accuracy (Samet, 1969). Furthermore, the source of overconfidence (most commonly found) or underconfidence (less commonly found) appears to result from both internal and external sources (Howell, 1971). Of particular relevance to piloting behavior where the man-machine interaction is so essential, Howell obtained an overwhelming preference among individuals to perform in situations in which the total task uncertainty is more internal or skill-based than external or environment-based. Yet, the specific manner in which people allocate their source of performance uncertainty to skill and chance (luck) factors, respectively (see Cohen, 1960), may in fact be an important discriminator with potential for pilot selection.

One of the most crucial aspects of decision making, especially in the context of piloting and aircraft, is the tendency of the decision makers to attempt lower probabilities of success for their action, i.e., their risk-taking propensity. This aspect of decision making has been widely studied in a variety of paradigms. For example, a simple risk-taking task which reflects decision-making aspects that might occur in an emergency situation (i.e., a pilot's decision when to abandon the aircraft and eject) has been investigated by Slovic (1966). In this task, the subject is presented with a row of 10 switches, 9 of which are "safe(S)" switches and one of which is the "disaster(D)" switch. Prior to each trial, the location of the D switch is randomly determined and unknown to the subject. For each successive pull of an S switch in a given trial, the subject obtains a fixed payoff which accumulates over the course of the trial. However, with the pull of the D switch, all cumulative payoff for the trial is forfeited. A trial is terminated when either the subject elects to stop and retain the payoff up to that point or the D switch is pulled. Performance on this task could provide multiple measures including mean number of switches pulled (an index of risk-taking behavior) and bankroll (i.e., influence of total prior earnings on current performance).

In the Slovic study, 1,047 subjects (735 boys and 312 girls) ages 6 to 16 performed the task -- at a county fair -- on a one-trial basis in order to win candies. The distribution of data showed great variability in risk-taking behavior: the stopping point (i.e., number of switches pulled) was four or less for about 12% of subjects, equal to five for 20%, six for 18%, seven for 30%, and eight for 20%. In addition, except for the 6 to 8 year age bracket, boys were found to be consistently more "bold" (i.e., they stopped later) than girls. These results suggest that there are probably reliable differences in the way specific individuals confront this decision task.

The "bet-preference" or "gamble task" has been employed extensively over the last 25 years in the investigation of such parameters as subjective probability, variance in possible payoffs, expected value/utility, and the way these parameters affect decision making behavior (e.g., Edwards, 1961; Rapoport & Wallstein, 1972; Slovic & Lichtenstein, 1968). Although different versions of the task have appeared in the literature, one of the most popular is the paired-comparison choice. Two simple examples are:

Example 1

BET A -1/6 probability to win $1.80 and 5/6 probability to win nothing.

BET B -4/6 probability to win $.60 and 2/6 probability to win nothing.

Example 2

BET A -3/6 probability to win $1.00 and 3/6 probability to lose $1.00

BET B -2/6 probability to win $1.40 and 4/6 probability to lose $.70

In each example, the subjects' task is to choose the one bet (either A or B) that they would prefer to play (for either real or imaginary payoffs --see Slovic, 1969). Although normative decision theory prescribes that a rational person should always prefer a bet with higher expected value (EV) than a bet with a lower expected value [e.g., Bet B (EV= $.40) over Bet A (EV=$.30) in Example 1], subjects often violate this principle (e.g., Lichtenstein et al., 1969); a subject might prefer Bet B in this example because it is less risky (i.e., has a lower variance) than Bet A.

By appropriate manipulation of bet parameters, this task has provided an efficient way to gain considerable data on individual decision performance. For example, Gilson (1968) successively used 24 paired-comparisons (requiring about 8 minutes of total task time) to assess individual differences on a variety of risk-taking dimensions and their relation to achievement motivation. Furthermore, through

the use of the "duplex gamble" (where each choice of a paired comparison involves two outcomes -- an amount to win and an amount to lose), the application of the bet-preference task has been extended to enable the investigation of individual differences in strategies for processing (Payne, 1973) and weighting (Slovic & Lichtenstein, 1968) risk dimensions (e.g., probability of winning, amount to win, probability of losing, amount to lose). Such differences in weighting schemes are likely to reflect genuine differences in the way individuals react and perform in risky situations.

The "choice dilemma" instrument (Kogan & Wallach, 1967; Stoner, 1968) has been widely used in the study of individual and group risk taking behavior (e.g., Pruitt, 1971a). The task presents hypothetical life-dilemma situations and assesses subjects' attitudes toward a risky course of action. An example of a typical problem follows:

> College X is playing its traditional rival, College Y, in the last game of the season. The game is in its final seconds, and College X, behind in the score, has the ball near the goal line. With time for only one more play, the coach must decide whether it would be best to settle for a tie score with a field goal which would be almost certain to succeed; or, on the other hand, should he try a more risky play for a touchdown which could bring victory if it succeeded, but defeat if not.
>
> Imagine that you are advising the coach. Listed below are several probabilities that the risky play will succeed.
>
> PLEASE CHECK THE LOWEST PROBABILITY THAT YOU WOULD CONSIDER ACCEPTABLE TO MAKE IT WORTHWHILE FOR THE COACH TO TRY THE RISKY PLAY.
>
> ____ The chances are 1 in 10 that the risky play will succeed.
>
> ____ The chances are 3 in 10 that the risky play will succeed.
>
> ____ The chances are 5 in 10 that the risky play will succeed.
>
> ____ The chances are 7 in 10 that the risky play will succeed.
>
> ____ The chances are 9 in 10 that the risky plan will succeed.
>
> ____ Place a check here if you think the coach should not try the risky play no matter what the probabilities.

The response scale provides a numerical index which is usually averaged over a number of different problems to provide a measure of risk-taking tendency. Based on a large number of experiments, variables that affect performance on the choice dilemma problems include person-based, information-based, and situation-based factors (Kogan & Wallach, 1967; Pruitt, 1971b). In general, the personal factors that lead one toward favoring either a "risky" or "cautious" alternative (when no objective criteria exist) reflect on the individual's relative value judgments and degree of social conformity (i.e., congruence or departure with the average level of risk expressed by a peer group). These value judgments and their resistance/tendency to change have a significant effect on individual differences in decision making.

In summary, decision making is one of the processes most widely cited as crucial to piloting. The decomposition of decision making into important subprocesses reveals important sources of variation in performance, as well as variability inherent to the decision making process overall. The speed with which a simple decision may be made is one source of variability, and this may be measured by a simple modification of the perceptual speed paradigm in which more than one potential stimulus and more than one potential response exist.

The primary subprocesses of decision making to be considered were information seeking, probability estimation, confidence estimation, and risk-taking. The urn-sampling problem is an effective means for assessing individual differences in information seeking behavior, and may be easily modified to include sampling costs and payoffs. Probability estimation may be assessed using the base-rate problem or the cue utilization problem, although the base-rate problem may be more easily implemented. The calibrating task for confidence estimation seems appropriate for assessing decision makers' ability to evaluate the accuracy of their responses. Several tasks exist which appear to index an individual's risk-taking propensity, including the sequential gamble, the bet-preference task, and the choice dilemma task. Of these the sequential gamble task appears to be the easiest to implement and most free from verbal bias.

Several other decision tasks were considered for this review, but were not selected because of one or more of the following reasons: (a) their mathematical solutions are relatively complex, (b) the empirical evidence for reliable individual differences in their performance is not clearly established, (c) their administration in a decision task test-battery would not be practical. Such tasks include, among others, cascaded inference (e.g., Johnson, et al., 1973), causal explanation (Fischhoff & Fulero, 1977; Shaklee & Fischhoff, 1978), decision strategy selection (Beach & Mitchell, 1978); McAllister, Mitchell, & Beach, 1978), and dynamic decision tasks (e.g., Rapoport & Wallsten, 1972).

Cerebral lateralization. The role of cerebral lateralization in performance has been discussed briefly in the review of perceptual-motor research, with special concern for the perceptual implications

of lateralization. For example, White (1969) reviewed a number of studies showing that printed verbal information presented to the right visual field (left hemisphere) was recognized better, and Kimura (1961, 1964) has demonstrated a left hemisphere advantage for auditory verbal stimuli and a right hemisphere advantage for melody perception. A study by Bryden (1973) illustrated the complexity of the phenomenon, however. He gave three visual perception tasks (letter recognition, form recognition, and dot localization) to groups of left- and right-handed subjects, balancing sex within each group. He found that letter recognition was better in the right visual field for right-handed subjects. This advantage did not hold for left-handed subjects, nor were any asymmetries apparent in the other two tasks. Moreover, speech lateralization as assessed by dichotic listening did not predict visual performance, indicating that the two phenomena have a different basis.

One possible interpretation of the complex results of cerebral lateralization studies is that hemisphere differences depend on different modes of processing, such as the simultaneous and successive processing modes proposed by Das, Kirby, and Jarman (1975). For example Cohen (1973) suggested that hemispheric differences were due to reliance on serial processing in the left hemisphere and parallel processing in the right hemisphere. This would result in sequential, temporally extended, and analytic materials being processed with a left-hemisphere advantage, whereas holistic, spatial, and patterned material would have a right-hemisphere performance advantage. To test this hypothesis, Cohen (1973) had six right-handed subjects perform an oddity-detection task, in which two, three, or four letters were presented, which were identical or contained all but one letter identical. The sets were presented monocularly to the right eye, and either to the right or left of the fixation point. She found an increase in RT as a function of set size for left-hemisphere processing, but not for right hemisphere processing. She suggested that this indicated parallel processing as the typical mode for the right hemisphere and serial processing for the left. Levy-Agresti and Sperry (1968) suggested that complex configurations are handled integratively by the right hemisphere and analytically by the left. It is clearly possible that the different processing modes underlie hemispheric differences, but it is unclear what predictive advantage this fact might have.

An alternative suggestion is that laterality differences in performance reflect differences in th degree of specialization of the two hemispheres (Levy, 1973). The idea here is that a maturational and perhaps sex-linked process takes place, specializing the left hemisphere for verbal material and the right hemisphere for spatial material. When that specialization is more complete, there is less interference between the hemispheres in processing the different types of material. The evidence for this suggestion is mixed, however, a discussed in the perceptual-motor segment of this review. A further suggestion (Dimond, 1970) is that hemispheric differences in performance reflect temporal differences in the operation of the hemispheres. That is, operations taking place in both hemispheres may

be relevant to performance, but the operation in one hemisphere may be relevant at one moment, while the other hemisphere becomes important later.

It is clear that hemispheric differences may have performance implications, but assessment of those differences may be complex and unclear. Tests for ear advantage in dichotic listening, and visual field advantage in a perceptual task, might be useful if interpreted cautiously.

Cognitive style. Another area of research with implications for human performance considers the topic of cognitive style, which refers to a general or consistent manner for approaching and processing different types of information. A number of dimensions of cognitive style have been specified, and similar to lateralization, the cognitive style elements tend to be bipolar in nature. For example, when considering the dimension of field independence, we do not characterize individuals as having more or less field independence, but rather contrast the field independent individual with the field dependent one. Also, there are performance advantages for different kinds of tasks for both poles of the field independence dimension, as well as the other dimensions of cognitive style.

The most prominent research in the area of cognitive style has been the work of Witkin (1949) concerning field independence. Field independence has been defined as the ability to overcome embedding contexts in perceptual functioning. For field dependent global processors, experience is governed by the organization of the field, while for the analytic field independent person, experiences can be analyzed and structured in other ways to suit the task at hand. Field dependence is easily measured through use of an Embedded Figures Test or some variant of the Rod-and-Frame-Tests.

Goodenough (1976) reviewed the literature on individual differences in field dependence and their relation to learning and memory. He suggested the following conclusions: (a) field-dependent (FD) subjects are dominated by the salient cues in concept attainment, while field-independent (FI) subjects sample the available cue set more fully; (b) FSs use "spectator" approaches to learning, whereas FIs more often use "participant" approaches; (c) FDs are more susceptible to stress effects on learning and memory, and (d) FIs perform more effectively under conditions of intrinsic motivation. The notion that people who develop an "analytical style" tend to employ that style in various learning and problem-solving situations, suggests that an easy index of such a style would be useful for selecting people who would be successful in a variety of learning and problem-solving situations. The general suggestion is that field-independent types may have an advantage in analytic tasks such as troubleshooting an electronic malfunction, whereas field-dependent persons might have an advantage in performance where interpersonal skills assume increased importance.

While field-independence is perhaps the most thoroughly researched cognitive style, a number of other styles have received increased attention (see Ragan et al., 1979, for a review). In addition to field independence, Ragan et al. (1979) have focused on three other styles as having the greatest potential relevance to Air Force selection and training. Those dimensions are impulsivity-reflectivity (Kagan, 1965), visual-haptic (Lowenfeld, 1939), and leveling- sharpening (Holtzman, 1952).

Impulsivity-reflectivity refers to the tendency, in a situation of response uncertainty, to respond with the first hypothesis and risk more errors (impulsive), in contrast with the tendency to consider all hypotheses and make fewer errors (reflective). It is usually measured by the Matching Familiar Figures Test (Kagan, 1969), which presents subjects with a familiar picture and requires them to select its exact duplicate from a set of variants. Ragan et al. (1979) suggested that impulsive styles may be at a disadvantage in Air Force training since too many errors will impede progress and perhaps result in failure.

The visual-haptic dimension of perceptual style refers to the tendency to depend on one sense modality, either visual or kinesthetic sense, as the primary intermediary for interpreting and learning from experience. This dimension is often measured using the Successive Perception Test, in which segments of a visual figure are shown sequentially and subjects are required to select the complete figure from a set of alternatives. Visual types usually integrate the separate impressions and do well on this test, whereas haptics internalize the segments and perform poorly. Ragan et al. (1979) felt that the importance of visual integration and memory for Air Force tasks constituted a significant advantage for persons relying on a visual style.

Leveling-sharpening refers to the tendency to incorporate new information into previous memory images, losing details of the specific event (leveling), as opposed to highlighting new information (sharpening), and making it more easily discriminable (Holtzman, 1952). One technique for assessing this dimension is the Leveling-Sharpening House Test, involving presentation of successive pictures of a house, from which portions are systematically subtracted, or to which they are systematically added. Another, used by Holtzman, is the Schematizing Test in which 10 series of five squares are projected on a screen, with the squares increasing in size regularly within a series, and a systematic increase in size from one series to the next. Unfortunately the reliability of these tests has been inadequate. While people employing a sharpening style might have a distinct advantage in the learning and retention of new information, Ragan et al. (1979) suggested that further refinement of the measurement techniques was necessary to insure reliability and construct validity.

While several other cognitive style constructs have been developed and researched, it seems unlikely that the others would prove useful in an Air Force test selection battery (Ragan et al.,

1979). Nonetheless a review of this area of research does suggest that a number of tests are available which, even though not tied to a specific cognitive or perceptual function, may indeed be used to predict performance in a number of fields. The most reliable predictors appear to be field independence-dependence and impulsivity-reflectivity, which may be assessed using modifications of the Embedded Figures Test and the Matching Familiar Figures Test, respective (or related reliable measures). In addition, measures of visual-haptic style and leveling-sharpening style may prove to be useful even though the selected measures appear somewhat less reliable.

Summary. The literature on cognitive processes is extensive, ranging from the initial mental representation of the physical world to broad patterns or styles for dealing with information. Individual differences in these cognitive processes must of necessity have a wide ranging effect on human performance. Moreover when the task being performed depends heavily on cognitive rather than or in addition to physical functioning, as does the task of piloting, the importance of these cognitive processes is further emphasized.

The processes of perception and attention provide the initial doorway through which information from the environment is screened and picked up. Differences in visual scanning rate and pattern are of evident importance to effective flying, as re the attentional functions of vigilance, selective attention, time-sharing, and resistance to interference. In order for information to be processed, it must be encoded, or represented mentally. The speed with which this representation may take place, as well as the preferred style or format for encoding, has implications for pilot performance.

Several processing alternatives exist once the information has been properly encoded. For piloting, one crucial process involves the ability to transform the data and compare them to some mental standard, as in comparing a dial reading to a stored representation of what the reading should be. This may involve a search process to find the information and to retrieve it from storage, and the ability to carry out these functions may be dependent on the overall processing capacity of the pilots, as well as their organization of memory and the appropriateness of their retrieval cue.

Both in flight training and later flying performance, the ability to learn quickly and to comprehend and integrate new information may be crucial. This comprehension is necessary to allow pilots to become aware of problems quicky as well as to attempt problem solutions. Alternative courses of action have very high stakes for pilots, and their ability to seek information, estimate the probabilities of various alternatives, correctly gauge their confidence in their chosen response, and avoid undue risk without being timid, are all crucial aspects of the decision making process.

Finally, several generalized aspects of cognitive functioning appear to have performance implications for the pilot. Some data suggest that the degree to which individuals are lateralized

determines their resistance to inter-hemispheric interference, and thereby particularly affect their ability to perform spatial tasks. Similarly, dimensions of cognitive style such as field-dependence, impulsivity, leveling-sharpening, and visual-haptic style all have implications for performance in a pilot's task. Within this review, tasks have been identified to tap all of these processes, providing a tool from which tasks assessing the full range of cognitive functions may be selected and employed as part of a test battery for pilot selection.

CANDIDATE TASKS

The portions of the review of the perceptual-motor and cognitive processes literature focused on pilot training and selection have identified a small number of tasks for which there is empirical evidence of their utility as component tasks in a pilot selection test battery. The portions of the review which examined the more general research on perceptual-motor and cognitive processes have revealed a large number of tasks which tap an ability/process assumed or inferred to be involved in flying performance and for which there is evidence of individual differences in task performance that can be reliably assessed. These tasks constitute additional candidate tests for a pilot selection task battery.

Each of the 44 potentially useful tasks revealed by the review of the literature is briefly described below.

Perceptual-Motor Tasks

Perceptual Speed. The subject is presented with a series of four geometric shapes in random order, and instructed to push switches corresponding to each shape in the sequence in which the shapes were presented. Time from presentation of the shapes until the initiation of response by the hand leaving a home key is the measure of perceptual speed.

Movement Speed: Simple. The subject presses a button as rapidly as possible in response to the onset of a light. The score is the average latency of response, over a series of signals.

Movement Speed: Multiple. The subject strikes each of two metal plates with a stylus, back and forth as rapidly as possible. The score is the number of taps recorded in the 15 sec trial.

Ballistic Aiming. The subject is required to tap alternately in small circles as rapidly as possible. The size of the target circles and the distance between them is varied. The subject is scored for number of taps on target during the 15 sec trial, and target width and movement distance are added to the accuracy data to yield a measure of information transmission rate.

Ballistic Aiming: Rotary Hand Movement. This task is similar to the ballistic aiming task except the movement distance is constrained, and the movement executed from the wrist, with the arm and elbow held secure. Rate of information transmission is the dependent measure.

Psychological Refractory Period. Similar to the task employed for movement speed: simple, except that on some trials a second stimulus occurs 200 to 500 msec after the first. In addition to recording response times, the data of special interest concern the increase in response time to the first signal as a result of the occurrence of the second.

Visual Movement Guidance. The subject is required to point with a long stylus to a sequence of visual targets, when the targets are visible but the hand and stylus are not. The stylus closes a circuit to indicate the location of each point, and the subject's score is average error in degrees.

Distance Estimation. A photograph of 15 vertical white stakes, arranged in ascending order from left to right and varying from 27-83 inches in height, is displayed so that the stakes are 14 feet from the camera. In a space in the middle of the 15 standard stakes is another stake 63, 67, 71, or 75 inches high, and 2, 4, 8, 16, or 32 times as far from the camera. Subject must match the distant stake with the standard stake of the same height. The subject's score is the number of stakes correctly identified.

Response Timing. The subject is presented with a regularly oscillating visual signal, and required to time a button pressing response to coincide with the change in direction of the visual signal. Speed and predictability of the oscillation may be varied, and the subject's score is the average error or deviation of the response from perfect coincidence.

Complex Coordination. The subject coordinates stick and pedal controls to match the indicated positions of stimulus light patterns. The score is the number of matches within the time period of the trial.

Two-Hand Coordination. The subject manipulates two hand controls, one of which moves a pointer right and left, while the other moves the pointer toward or away from the subject. The task is to keep the pointer in contact with a target disc as it moves in an eccentric pattern, and the subject's score is the time on target in the time period of the trial.

Pursuit Tracking. The subject attempts to keep a stylus tip in contact with a target set near the edge of a revolving turntable. The score is total time on target during the test period.

Compensatory Tracking. The subject moves a control attempting to keep a pointer in contact with an eccentrically moving target by anticipating direction change. The score is time on target and root mean square error.

Rudder Control. The subject coordinates movement of two foot pedals to align himself and the cockpit with a set of target lights. The score is the total time in the test during which the subject is precisely aligned with the target lights.

Kinesthetic Memory. Four stimuli are presented as for the perceptual speed test, except that a bell rings at stimulus onset. The subject is required to activate a switch sequence in an order corresponding to presentation order of the stimuli. After 12 trials using the same order, the subject is blindfolded, and attempts to produce the activation sequence when the bell sounds, without benefit

of visual guidance. The score is the number of correct trials and average response time in the 12 memory trials.

Spatial Orientation. The subject is required to imagine himself in a familiar location, such as in front of the PX, on a military base. Using a circle as a point of reference, the subject must then indicate the direction of a set of familiar buildings relative to himself. The degree of error in direction indication is the subject's score.

Route-Walking. A map display is presented to the subject, which indicates a starting point and a goal, and the subject must indicate the shortest path from the start to the goal. The subject's performance is indicated by total time taken to arrive at the goal and the number of wrong turn errors committed.

Cognitive Tasks

Monitoring. The subject is required to listen over headphones to detect the occurrence of an auditory signal (a tone) embedded in white noise. The subject's performance is indicated by the percentage of correct detections and false alarm errors over the course of the test period, as well as by the signal detection parameters of d-prime and beta.

Selective Attention. The subject is presented with two different streams of words, one to each ear, over headphones and is required to monitor one ear for the occurrence of two to four digits embedded in the word stream and then to push a button corresponding to the digit when it occurs. Performance is indicated by the number of digits missed, intrusions from the irrelevant ear, and reaction time. In a second part of the study, subjects monitor either the same ear or the other ear. They are presented with three pairs of digits, one digit to each ear, and required to push buttons corresponding to the digits which occur in the relevant ear. Performance is indicated by errors of intrusion and omission. Indications of a performance advantage for one ear over the other may be used cautiously as an indication of degree of cerebral lateralization.

Time-sharing. The subject is required to perform a compensatory tracking task as described above, while also responding to the presentation of one of four geometric figures by pressing a corresponding button as described under decision speed. Instructions may be adjusted so that a certain level of performance is required on both tasks, indicating ability to allocate attention. Or the subject might be instructed to respond only to the figures when it will not interefer with tracking performance, indicating the subject's residual processing capacity.

Interference Susceptibility. The Stroop test employs a Color, or C-card, consisting of nine rows of eight items where the items are four white X's printed on tape of different colors. The subject is required to name the color of the tape for each item as rapidly as

possible and is scored in terms of total time to name the 72 items. On a similar Color-Word card (CW), the X's are replaced by color names different from the color of the tape for each item. The subjects are again required to name the tape color for each item, and the increase in total time for CW cards compared to C cards is an index of interference.

Visual Scan. The subject is required to memorize a target letter, then scan a visual display containing one to six letters. The subject's task is to decide if the target letter is contained in the search set. The score is response time as a function of the search set size, with the slope and intercept of the resulting function being relevant.

Item Recognition. In this task, the subject is required to memorize the search set of one to six letters, then decide if a presented target letter is contained in the search set. The same response measures as employed in the visual scan task are relevant. Additional information may be gained in this paradigm under situations of less than perfect performance. A target set of 12 letters, for example, allows separate evaluation of STM and LTM search, as well as providing error information. Error data may also be obtained by probing for the location of the target letter within the search set, and under these circumstances differing search strategies may be employed (i.e., self-terminating rather than exhaustive).

Perceptual Readout. The subject is presented with a matrix of three rows of four letters each, for 50 msec. After a delay of from 0 to 250 msec, a tone is presented to indicate which row of letters the subject is to report. The subject's score is percentage correct as a function of the delay of the tone, with a greater negative slope to this function indicating slower readout or faster decay of sensory memory.

Encoding Speed. Pairs of letters are presented to the subject, who is required to make a speeded same-different judgment based on either physical identity (AA vs. Aa), name identity (Aa vs. Ab), or catetory identity (Ae vs. AB). The subject's score is the average response time of the judgments, and comparisons between response times for the three different judgment rules are especially relevant.

Encoding Style. Three lists of phrases are presented for meaningfulness judgments; one list contains sensible (S) phrases which look and sound sensible (It's not so) and homophone (H) phrases which sound sensible but do not look sensible (It's knot so). For this list, only phrases which look meaningful are so judged. A second list contains H phrases and nonsense (N) phrases (It's boat so) and meaningfulness is judged on sound. For the third list containing S and N phrases, meaningfulness may be correctly judged using either rule. Reaction times are scored for the judgments on all three lists, and times for the SN list judgment are particularly useful for indicating preferred encoding style.

Mental Rotation. Subjects are presented with a standard letter, followed at .5 sec or 3.0 sec interval by another letter rotated with respect to the standard. The second letter is either the same or a mirror image of the standard, and the subject is required to mentally rotate the second, compare it to the mental image of the standard, and make a speeded same-different judgment. Subjects' reaction times are scored as a function of degree of rotation required, and errors are also noted. The slope of the reaction time function can also be used to discriminate holistic and analytic comparison styles. Three dimensional stimuli may also be substituted for the letter stimuli, since some evidence indicates that the comparison process differs with two- and three-dimensional stimuli.

Imagery. A target list of two to four words is presented, followed by a probe word that the subject must judge whether is a member of the target set. In the second half of the study, subjects are instructed to form an interactive image of the target set items. Reaction time is scored as a function of target set size, and the decrease in slope as a function of imagery instruction is taken as an indication of the effectiveness of the image (since in an interactive image all the items can be searched in parallel, resulting in a zero slope).

Processing Span. Subjects are presented lists of digits, with from 4 to 12 digits per list, at a rate of 2 digits per second. They are required to recall the digits in the order presented. Subjects are scored in terms of the number of digit lists recalled at each list length, given three presentations at each length. That is, if the subject got all the lists correct up to and including a length of six, then two lists right at length, seven, one at length eight, and none higher, span = $6 + \frac{2}{3} + \frac{1}{3} = 7$.

Recall. Subjects are presented a list of 22 unrelated concrete nouns in a multiple study-test free recall procedure until all 22 items have been correctly recalled. Trials to criterion and total study time provide an index of learning rate, while output order consistency over trials indicates the tendency to organize material in LTM.

Recognition. A continuous sequence of concrete nouns (or three digit numbers) is presented with a 2-second inter-stimulus interval. Items in the sequence are repeated at lags of 1, 2, 4, 8, 12, 16, 20, 24, 30, and 36 items, with five exemplars of each lag in the 101 item list. The subject's task is to judge whether each item has occurred previously in the list. Traditional memory measures are used, such as probability of a correct recognition and probability correct as a function of lag. In addition the signal detection parameters of d-prime and beta provide additional recognition performance data.

Immediate/Delayed Memory. A continuous sequence of digits is presented visually to the subject, with either a 2-second

inter-stimulus interval (for immediate memory) or a 5-second inter-stimulus interval. The subject is required to push a button corresponding to the digit which occurred two items prior to current presentation, that is, 6 seconds previously in the immediate condition, 15 seconds previously in the delayed. The score is the number of correct responses made by the subject.

Verbal Comprehension. Subjects are presented a central fixation point for a fixed warning delay, followed by stimultaneous presentation of a sentence and a pictorial representation, and they are required to read the sentence and decide if it's true or false with respect to the picture; for example, "star isn't below cross + --- *
false." Reaction time is measured as a function of grammatical and semantic complexity. Eight sentences are constructed and paired with each of two pictures, resulting in 16 trials per block. The sentences are either positive or negative; use either "above" or "below"; and are either true or false with respect to the picture. The reaction times are used to calculate estimated values for each of Clark and Chase's four parameters, and error data are collected as well.

Analogies. Subjects are presented with the first three items in a verbal analogy with a blank for the fourth item, and choose one of two alternatives to best fill the blank, e.g., LAWYER: CLIENT:: DOCTOR: _____ a) MEDICINE b) PATIENT. A pre-cuing condition is included in which the first two items are presented prior to presentation of the complete problem. Number correct and reaction time are measured in all conditions.

Decision-Making Speed. A visual display of two, four, or eight geometric figures is employed, with a keyboard whose buttons correspond to the figures. When one of the figures is presented, the subject is required to activate the corresponding button. The scores are the latency of this activation, latency as a function of number of alternatives. It is also possible to measure latency changes with practice if so desired.

Probability Estimation. An example of the problem presented in this task is:

> Two cab companies operate in a given city, the Blue and the Green (according to the color of cab they run). Eight-five percent of the cabs in the city are Blue, and 15% are Green. A cab was involved in a hit-and-run accident at night, in which a pedestrian was run down. The injured pedestrian later testified that though he did not see the color of the cab, due to the bad visibility conditions that night, he remembers hearing the sound of an intercom coming through the cab window. The police investigation discovered that intercoms are installed in 80% of the Green cabs, and in 20% of the Blue cabs.

> What do you think are the chances that the errant cab was Green?

The problem contains two kinds of information: the background or base-rate information (color distribution of cabs in the city) and the diagnostic information (the intercom and its relative distribution in blue and green cabs). The response (i.e., probability judgment) reflects whether or not the subject has attended to the base-rate information, and if so, how that information has been integrated with diagnostic information.

Information-seeking. The following representative task is presented to the subject.

> Consider two urns, Urn A with 70% white balls and 30% black balls, and Urn B with 70% black balls and 30% white balls. Select balls one at a time, with replacement, from one of the urns. After observing each sample drawn, decide whether to continue sampling or to make a decision concerning which urn (A or B) is being sampled.

Task performance may be assessed in terms of decision correctness and number of samples drawn prior to decision.

Hypothesis Generation. A common object (e.g., ping pong ball, bobby pin) is named and the subject is given a few minutes to generate as many unusual uses for it as possible. The number of acceptable responses provides a measure of subject creativity related to the decision process of alternative or hypothesis generation.

Risk-taking. The subject is presented with a row of 10 switches, 9 of which are "safe(S)" switches and one of which is the "disaster(D)" switch. Prior to each trial, the location of the D switch is randomly determined and unknown to the subject. For each successive pull of an S switch in a given trial, the subject obtains a fixed payoff which accumulates over the course of the trial. However, with the pull of the D switch, all cumulative payoffs for the trial are forfeited. A trial is terminated when either the subject elects to stop and retain the payoff up to that point or the D switch is pulled. Performance on this task provides multiple measures including mean number of switches pulled (an index of risk-taking behavior), mean reaction time for pull and no-pull decisions, and regret (effect of outcome on previous trial and behavior on current trial) and bankroll effects (i.e., impact of total prior earnings on current performance).

Confidence Estimation. The subject answers several multiple-choice questions. For each question, the subject states a confidence (e.g., 50%, 70%, or 100%) that the answer is correct. For all levels of each stated confidence level, the observed or actual percentage of correct answers is computed and compared to that confidence level. The resultant calibration or realism-of-confidence

function (i.e., confidence level versus percent correct) determines the degree of overconfidence or underconfidence expressed by the subject in the decisions.

Cognitive Style Tasks

Field-independence (embedded figures). As a test for field independence, this procedure involves the presentation of a simple geometric figure followed by two complex figures in one of which the simple figure is embedded. Subjects are required to choose the complex figure that contains the simple geometric figure. Percentage correct and reaction time are measured.

Inpulsivity (matching familiar figures). As a test for the impulsive-reflective cognitive style, this procedure involves presentation of a familiar figure (a standard) followed by a number of variants from whih the subject must choose the one identical to the standard. Twelve trials are given, and measures of latency and errors are recorded.

Visual-haptic (successive perceptions). As a test of the visual haptic dimension, this consists of 38 trials in which the subject is presented with a visual pattern a small portion at a time through a moving slot, and is required to select the presented pattern from five variants. Reaction time and proportion correct are measured.

Leveling-sharpening (leveling-sharpening house test). Sixty different representations of a house are presented in which features are gradually added or subtracted. Each picture is presented for 5 seconds, and the tendency to assimilate the changes versus the tendency to highlight them is observed.

TEST BATTERY SELECTION

The tasks identified in the literature review may be useful as predictors of flying training and flying performance, but a battery of 44 tests is impractical from a psychometric as well as from a cost-effectiveness standpoint. Therefore, further selection among the candidate tasks is required to produce a task battery of manageable size with the highest potential efficiency and predictive value.

A broad spectrum of potential criteria exists for selecting among the tasks in the candidate pool, some of them intrinsic to testing concerns and others specific to the problems of selecting pilot trainees. Two of the most important criteria have already been applied in arriving at the pool of candidate tasks: i.e., the task must assess some ability or process that is demonstrably relevant to flying performance, and the task must show individual differences in performance. The other criteria which were applied to reduce the list of candidate tasks are:

1. feasibility - the ease with which the task may be implemented and the ability to produce a meaningful amount of data in a short time are important positive factors for the inclusion of the task.

2. sensitivity - a task which is more capable of revealing small individual differences is more likely to be included.

3. interest - tasks which are more engaging and tend to hold the subject's interests are considered to have a motivational advantage and preferred for the test battery.

4. independence - wherever possible, no more than one task relevant to a given ability or process is included, and attempts are made to minimize the inclusion of tasks which rely on overlapping or highly related processes.

5. construct validity - tasks for which strong evidence exists that they actually involve the ability or process in question are given preference.

6. verbal bias - tasks which do no rely heavily on verbal materials are given preference in the test battery.

In order to arrive at a preliminary test battery, the criteria were applied iteratively in two stages to the pool of candidate tasks. In the first stage the criteria of feasibility, sensitivity, and interest were applied to the candidate tasks to result in one or more tasks representing each ability/process identified as relevant.

In the second stage, the criteria of independence, validity, and freedom from bias were applied resulting in the selection of a preliminary test battery. The results of the initial stage of criterion application are presented in Table 1.

The second stage of application of criteria to the reduced pool of 26 tasks presented in Table 1 was somewhat more subjective and difficult. A discussion of some of the tasks eliminated and the reasons for elimination illustrates how the criteria were applied. For example, the movement speed task was eliminated on the basis of being marginally less important than other abilities for flying, as well as easily assessed by a slight modification of the perceptual speed paradigm. Similarly, the task employed to assess multilimb coordination (complex coordination) also loads heavily on control precision, and the current evidence supporting complex coordination makes it a preferable task.

The monitoring task requires extensive time for data acquisition and its individual reliability is less than might be desired. There is a large overlap in the selective attention and interference susceptibility processes, but the selective attention paradigm has more evidence of predictive validity and is more flexible, potentially assessing lateralization. Visual scanning is of undoubted importance to flying performance, but the evidence shows a strong correlation between the parameters of this task and those of the memory search task, and the memory search task has stronger reliability evidence.

The visualization ability is assessed by a task whose sensitivity is in doubt and whose reliability is inadequately supported. The processing capacity test has questionable construct validity as demonstrated in the literature review, but processing capacity may be readily assessed through a modification of the time-sharing task instructions. The comprehension and problem-solving functions are also important, but they are very difficult to implement in a manner which avoids verbal bias, and are complex functions which may overlap with other portions of the test battery. The information seeking task measures an important aspect of decision making, but this aspect may also be assessed through modifications of other decision making tasks. Finally the test for the impulsivity style was eliminated due to its near certain overlap with some of the decision-making functions, especially risk taking. The test battery after application of these criteria is presented in Table 2.

Table 1. The Tasks Selected as a Result of the Initial Criterion Application

Task	Ability/Process Related to Flying Performance[a]
Perceptual Speed	Perceptual Speed
Ballistic Aiming	Movement Speed
Pursuit Tracking	Control Precision
Complex Coordination	Multilimb Coordination
Compensatory Tracking	Rate Control
Kinesthetic Memory	Kinesthetic Sensitivity
Route-Walking	Spatial Orientation
Monitoring	Attention
Selective Attention	Attention
Time Sharing	Attention
Interference Susceptibility	Attention
Visual Scan	Perception
Encoding Speed	Encoding
Imagery	Visualization
Mental Rotation	Comparison
Item Recognition	Memory Search
Processing Span	Memory Capacity
Immediate/Delayed Memory	Memory Retrieval
Verbal Comprehension	Comprehension
Analogies	Problem Solving
Decision Making Speed (Choice Reaction Time)	Decision Making
Information Seeking	Decision Making
Probability Estimation	Decisiod making
Risk Taking	Decision Making
Selective Attention	Cerebral Lateralization
Embedded Figures	Cognitive Style: Field-Independence
Matching Familiar Figures	Cognitive Style: Impulsivity

[a]From Table A1, Appendix A.

Table 2. The Test Battery as a Result of the Secondary Criterion Application

Task	Ability/Process Related to Flying Performance[a]
Perceptual Speed	Perceptual Speed
Complex Coordination	Multilimb Coordination
Compensatory Tracking	Rate Control
Kinesthetic Memory	Kinestheteic Sensitivity
Route-Walking	Spatial Orientation
Selective Attention	Attention
Time Sharing	Attention
Encoding Speed	Encoding
Mental Rotation	Comparison
Item Recognition	Memory Search
Immediate/Delayed Memory	Memory Retrieval
Decision Making Speed (Choice Reaction Time	Decision Making
Probability Estimation	Decision Making
Risk Taking	Decision Making
Embedded Figures	Cognitive Style: Field-Independence

[a]From Table A1, Appendix A.

SPECIFICATION OF TASK PARADIGMS

In this section detailed descriptions of this task comprising the pilot selection battery are provided. Inasmuch as possible, these descriptions are empirically based; that is, they are taken directly from the literature which provided evidence for the reliability and validity of the task. (See Table 3 for representative sources.) Where necessary, the task descriptions have been modified to allow easy adaptation to a computerized testing station. These descriptions are intended to provide a software designer with information necessary to implement the tests without requiring detailed knowledge of them. Nonetheless, several caveats should be noted prior to the implementation and testing of the task battery.

A number of task parameters (i.e., inter-stimulus interval, inter-trial interval, etc.) are specified in the task paradigms, based on those used in previous research. These parameters are meant to be guidelines rather than rigid requirements. Empirical testing will be required to determine parameter values that result in maximum reliabilty, and considerations regarding the amount of time consumed by the total task battery must be taken into account. A similar situation exists for the suggested number of trials per task. Enough trials should be employed to provide reliable data within a reasonable time; the number of trials is a matter for empirical determination.

It is important to note that, for many of the tasks described, the particular form of implementation chosen is only one of many alternative means for implementing that task. Several criteria influenced the choice of the particular form of implementation specified, including: (a) demonstrated reliability and validity for the task using the selected implementation, (b) application of common practices in testing (such as providing rest periods between trials), to tasks not originally designed for testing purposes, and (c) the feasibility and ease with which the chosen implementation may be adapted to the Air Force Automated Measurement System. Nonetheless the specific paradigms for some tasks could be changed substantially without altering the task's potential value as a selection device.

For example, the time-sharing task requires performance of a primary compensatory tracking task and employs the secondary decision-making speed task to assess time-sharing. While this approach has demonstrated validity (Damos, 1978), other tasks have been used as both primary and secondary tasks, and different task instructions employed (North & Gopher, 1976), with equally convincing evidence of valid prediction. The choice of the final paradigm for the time-sharing task (and other tasks such as those related to decision making) is flexible and may be somewhat altered, if necessary, without reducing the value of the battery.

In their present format, the tasks are specified independently, and no suggested order of testing is provided. Nonetheless an examination of the paradigms reveals similarities between tasks in

Table 3

Exemplary Reference Sources for the Tasks in the Pilot Selection Battery

Task/Process	References
Perceptual Speed/ Perceptual Speed	Hunter, 1975 Sperling, 1960
Complex Coordination/ Multilimb Coordination	Fleishman, 1964 McGrevy & Valentine, 1974
Compensatory Tracking/ Rate Control	Fleishman, 1964 Adams & Creamer, 1962
Kinesthetic Memory/ Kinesthetic Sensitivity	Hunter, 1975 Fleishman & Rich, 1963
Route-Walking/ Spatial Orientation	Koslowsky & Bryant, 1977 Allen, Siegel, & Rosinski, 1978
Selective Attention/ Attention	Gopher & Kahneman, 1971 Jensen, 1965
Time Sharing/ Attention	North & Gopher, 1976 Damos, 1978
Encoding Speed/ Encoding	Posner & Mitchell, 1967 Baron, 1973
Mental Rotation/ Comparison	Shepard & Metzler, 1971 Cooper, 1976
Item Recognition/ Memory Search	Sternberg, 1966 Juola & Atkinson, 1971
Immediate/Delayed Memory/ Memory Retrieval	Hunter, 1975 Shepard & Teghtsoonian, 1961
Decision Making Speed/ Decision Making	Fleishman, 1964 Mowbray & Rhoades, 1959
Probability Estimation/ Decision Making	Kahneman & Tversky, 1973 Lyon & Slovic, 1976
Rick Taking/ Decision Making	Slovic, 1966 Slovic, 1969
Embedded Figures Field Independence	Witkin, 1949 Goodenough, 1976

terms of visual displays employed, stimulus materials provided, and response modes required. Ideally these similarities may be taken advantage of in developing and programming the tasks, but caution and professional guidance are urged to ensure that the tasks are not adversely affected by modifications in the quest for programming ease. Similarly, the tasks might be ordered to take advantage of display and response similarities, but care must be exercised to note the influence of learning when the tasks are thus ordered. The determination and evaluation of potential task order effects is also a matter for empirical determination through pilot testing of the battery.

The task battery samples a broad range of abilities and processes of importance to flying and emphasizes the processes most often discussed in pilot training research (i.e., attention and decision making). Nonetheless, in a battery of 15 tasks, some are unlikely to demonstrate individual predictive validity, which might suggest either replacing the task with an alternative from the list of candidate tasks or deleting it from the battery without replacement. The worth of the battery must be evaluated in terms of the additional predictive value gained by each task added in contrast to the costs incurred. The composition of the ultimate task battery will very likely be similar to that presented here, but the number of tasks may be fewer.

Perceptual Speed

The subject is presented with a sequence of four digits all at once and in random order, and required to respond by actuating keyboard numbers in the same order as the presented digits. In addition to noting the number of sequences recalled correctly and the response time, a measure of perception time is taken. The subject is required to keep a home key pressed down prior to each trial. The time that passes between the onset of the digit sequence and the time at which the home key is released to initiate the response is a measure of the subject's perceptual speed.

For this task the stimulus ensemble is defined in terms of the digits 0 to 9. From this ensemble, a subset of four digits is drawn and presented in random order on each trial. For any trial, the subset of digits is drawn by sampling without replacement, so that no digit occurs more than once on that trial. A total of 22 trials are defined in terms of these stimuli. The first 2 trials are designated for practice, and the final 20 trials are scored.

The subject has two primary respones in this task: initiating each trial by depressing a "home key" and after the stimuli have been presented, releasing the home key prior to actuating the keyboard numbers in the desired sequence. The first response, then, is the release of the home key, and the time taken between onset of the digits and this response is a measure of perceptual speed (i.e., how quickly the subject can recognize the digits and organize an appropriate motor response). The second response is the actuation of this sequence of keyboard numbers, measured as the time from the release of the home key to the actuation of the fourth numbers. After each trial, the subject is provided feedback on the correctness of the response.

The sequence of events in this task is as follows: A warning signal appears for 2 seconds, and disappears, followed 1 second later by the sequence of four digits. The four digits remain present until the subject has released the home key, and then they disappear. A 5-second inter-trial interval feedback is provided. After this interval, the warning signal appears again to signal the beginning of the next trial. The digits for each trial are presented with a size and separation sufficient to assure discriminability for the subject.

For each trial, the computer must keep track of two intervals -- the time from the appearance of the digits to the release of the home key, and the time from the release of the home key to actuation of the fourth digit key. In addition, the computer must compare the response sequence with the presented sequence for accuracy of the response. The primary response measures are mean and median latency for perceptual speed, as well as the variability (standard deviation) of the latency. In addition total response time and accuracy are recorded. Two trials are presented initially as practice, followed by the block of 20 trials from which response measures are compiled.

Sample instructions are as follows:

> In this task you will be making a sequence of simple responses to correspond to a sequence presented to you. We are concerned with the speed and accuracy of your response. A warning signal will appear on the screen, followed shortly by a sequence of four digits. You must have your hand pressing down on this home key for the trial to begin. When the digits appear, your task will be to release the home key and press the keyboard buttons with the same numbers as the digits in the same sequence as presented. For example, if the digits appeared 4 1 7 3, you would press key 4 first, then 1, then 7, and finally 3. The digits will disappear when you release the home key, so do not release the key until you are ready to respond. It is important that you respond as quickly as possible without making mistakes. Are there any questions?
>
> The first two trials will be to practice to acquaint you with the task and will not be scored. If there are no questions following practice, we will then continue with the task.

Complex Coordination

In the revised version of the complex coordination task, the subject manipulates two independent hand controls to make continuous corrections of three axes. A display of a vertical and a horizontal row of dots intersecting in the center of the screen is presented, along with two response symbols. One symbol is controlled by a joystick, and the subject attempts to counteract its movement and keep it at the point of intersection of the two rows of dots (essentially a two-dimensional compensatory tracking task). At the same time, a bar marker moves left to right at the bottom of the screen, and the subject uses a rudder control with the other hand to attempt to align this marker with the vertical row of dots. The sum of error in terms of distance from the desired location of both symbols on all three axes serves as an indicator of the subject's multilimb coordination ability.

For this task, a cathode ray tube screen might provide a visual display which consists of a horizontal and vertical row of dots intersecting in the center of the screen. On this display, two symbols are defined; an X serves as one response cursor and is moveable in two dimensions (i.e., left to right and up and down), and at the bottom of the display, a bar marker serves as the other response cursor and moves in the left to right dimension only. Both response symbols move with a direction and velocity corresponding to the direction and displacement of their respective controls.

A central viscously damped joystick, which can be adjusted to the length of the subject's arm and can be moved in both dimensions, is used to control the X response cursor. That cursor is driven by a forcing function to move away from the central intersection point and moves in two dimensions. The subject moves the joystick opposite the movement of the cursor in order to keep it at the central intersection point. A second control, defined as the rudder control, is mounted to one side and can be moved back and forth only. This control moves the bar marker cursor at the bottom of the display. The bar cursor is also driven by a forcing function to move left to right, and the subject moves the control to attempt to negate this movement and keep the bar cursor aligned with the vertical row of dots.

A trial in this task is defined in terms of 60 seconds of performance during which the subject attempts to coordinate the control movements to minimize the distance of both response cursors from their desired location. Five trials constitute the test. The total sequence for this task includes instructions, 3 minutes of directed practice on the task, and five 60-second trials with a brief rest between the practice and task and between each trial.

Performance measurement in this task requires that some unit of distance be defined in terms of the visual display. Error is then measured throughout the trial by noting the distance between the cursors and their desired locations. For the X-response cursor, two

error measurements are taken, corresponding to error distance in the X-axis and the Y-axis. For the bar marker, a single error measurement -- distance from the vertical line -- is collected. The test station samples the location of each cursor 20 times a second, records the three error measurements, and cumulates them over the 60-second period of the trial. Learning appears to stabilize by the fourth trial, so the total error on trials 4 and 5 for each dimension constitutes the score.

A sample of the instructions for this task is as follows:

> In this task we are interested in how well you can coordinate movements from the two arms to control independently fluctuating symbols on the display. The display consists of a horizontal and vertical row of dots crossing in the middle of the screen. There is also an X near the middle of the screen, and a line near the bottom.
>
> The X will be moving erratically about the display. By moving the joystick directly in front of you, you can counteract the movement. Use that control to try to keep the X at the center where the lines of dots cross. At the same time, the bar marker at the bottom of the display will be moving back and forth. The control to the side controls the bar marker's movement. Use that control to try to keep the bar directly aligned with the vertical row of dots. Remember, both the X and the bar will be moving at the same time, so you must operate both controls continuously and simultaneously to minimize the distance of each from its desired location. Any questions?
>
> First you will practice the task for a few minutes to get used to the controls and the nature of the task. Then after a brief rest you will be tested on the task, in five 60-second trials. Remember, use both controls to minimize the error distance of both the X and the bar.

Compensatory Tracking

A pointer or error cursor is defined and moves eccentrically within a fixed defined field. The subject attempts to null the position of the error cursor -- keep it as close as possible to a defined central point by manipulating a control stick which changes the direction of the pointer. For this task, the error cursor moves only along a single axis -- left and right. Parallel left and right manipulations of the control stick may be used to change the direction of the cursor's movement. The average error in terms of distance from the central or null point serves as an index of the rate control ability of that subject.

The stimulus ensemble for this task is defined in terms of a fixed field (for example, a 7.5 cm. circular space) with a defined central null point. Within this field, an error cursor is defined by a marker which can be moved along the horizontal left-to-right axis within the fixed field. The speed and eccentricity of the movement of the cursor is determined by a forcing function. The forcing function may be defined in several ways. Damos (1978) provides an example: "The forcing function was generated by random noise filtered by a fourth-order Butterworth filter with an upper cutoff of 0.40 Hz. The control dynamics were of the form Y = K/S." The difficulty of the tracking task may be varied by changing the form of the control dynamics, but even a random function may provide a reliable test.

The subject responds by moving a control stick which controls the *movement* of the error cursor, attempting to anticipate and counteract the movement of the cursor and thereby keep it as close to the central null point as possible. The control stick is mounted centrally in front of the subject and operated by the subject's dominant hand. The position of the stick may be adjusted for the length of the subject's arm. The movement of the control stick is viscously damped, and left-to-right manipulations of the stick control corresponding movements of the error cursor.

A trial is defined in terms of the length of time during which continuous tracking performance is recorded. For example, tracking performance might be divided into 1-minute trials, with a brief rest between trials. The compensatory tracking task is then defined in terms of a block of six 1-minute trials. If desired, more than one block of trials might be employed, so that the difficulty of tracking could be varied by changing the form of the control dynamic.

Performance is assessed on the compensatory tracking task by measuring the absolute error (without regard for direction of the error) in distance from the central null point accumulated over the 1-minute trial. The root-mean-square error for each trial may be calculated for each subject, and changes in error over trials (learning) and a function of task difficulty may also be examined. The test station is required, therefore, to sample and record the distance of the cursor from the null point on-line and continuously and to cumulate the absolute value of those distances. The first

trial of the block of trials may serve as a practice trial to ensure that the subject understands the task, and performance on that trial is not scored.

A sample of instructions for the compensatory tracking task follows:

> In this task we are interested in how well you can anticipate the movement of a point on the screen and move to counteract it. You can see that a small circular field with a point in the center is displayed on the screen, and a marker or dot is moving irregularly back and forth inside that field. The control stick in front of you can be used to control the movement of that marker. That is, when you move the control to the left, the marker will move to the left, and when you move the control to the right, the marker also moves to the right.
>
> Your task is to move the control stick to keep the moving dot aligned with or on top of the central point in the field. In order to do so, you will have to anticipate the movement of the marker and continuously adjust the control in order to counteract the movement. You must also be careful not to make too large a control movement, which will cause you to overshoot the central point. The object is to minimize the distance between the central point and the moving cursor. Any questions?
>
> Each trial will last 1 minute, and will be followed by a brief rest. You will have six trials in this block. The first trial will accustom you to the controls and the task, and will not be scored. Remember to try to minmize the distance between the central point and the moving marker.

Kinesthetic Memory

The kinesthetic memory task is presented in two parts. In the first part, a warning tone is sounded, followed by the presentation of a sequence of four digits. The subject must respond by leaving a home key and activating four switches, which correspond to the digits presented, in the same sequence as the digits are presented. Twelve such trials occur, the same digit sequence being used on each trial. The subject is then required to wear opaque goggles, and for the next 12 trials, attempts to activate the learned sequence of switches without visual guidance at the sound of the warning tone. The speed and accuracy with which these blind activation sequences are performed serve as indices of the kinesthetic sensitivity of the subject.

The stimulus ensemble is defined in terms of four digits, which are randomized into a single sequence and thereafter serve as the stimulus on all trials. The response required of each subject is to leave a home key which the subject holds down prior to stimulus onset, and to activate the switch sequence in the prescribed order as quickly as possible.

For any given trial, a tone sounds simultaneously with the onset of the stimulus sequence, and the four digits appear all at once, in a horizontal row. The subject controls the length of time the stimulus is visible; when the home key is released to initiate the response, the screen is erased so the subject cannot refer to it while responding. The subject then activates the switch sequence as quickly as possible and returns to the home key. A 5-second inter-trial interval occurs during which feedback concerning the accuracy of the response is provided. Twelve learning trials are presented in this format.

In the second part of the test, the procedure is identical, but the subject is required to wear opaque lenses so that the display and the response switches cannot be seen. When the tone occurs, the subject attempts to duplicate the learned activation sequence without visual guidance. Twelve such kinesthetic trials are presented. The visual stimulus in part one must be large enough and far enough apart to be easily discriminable, and the response switches must also be adequately spaced.

The primary measures collected in this task are response time, measured from release of the home key to completion of the activation sequence, and response accuracy. These measures may be used at the completion of the test to compute mean, median, and standard deviation of response time for both parts of the test, as well as accuracy. Part one parameters reflect on kinesthetic learning, whereas part two parameters are especially relevant to sensitivity to kinesthetic activity in the absence of visual feedback.

Sample instructions for this task are as follows:

> In this task we are interested in the speed and accuracy with which you can learn to complete a switch activation sequence. This task consists of two parts. In part one, a sequence of four digits will appear on the screen, and simultaneously you will hear a warning bell sound. Your task will be to activate the switches corresponding to the digits in the same sequence as they appear in on the screen. For example, if the sequence is 4 1 2 3, you must activate first switch 4, then 1, then 2, and finally 3. Do not leave the home key until you know the sequence, for the digits will disappear as soon as you release the home key. Activate the switch as quickly as you can without making mistakes. The digits will be presented in the same sequence on each trials, so the sequence of movements may be learned.
>
> In part two of the task, the same switch activation sequence will be required. However, you will be wearing opaque lenses and will not be able to see the response switches. When you hear the tone, activate the switch sentence exactly as you did in part one, working as quickly as possible without making mistakes. Are there any questions?

Route Walking

The subject is presented with a schematic representation of a city map. On this map, a starting point and end point are indicated, and a relatively complex route drawn from the start to end. The subjects are asked to imagine themselves at the start point, and to decide at each turn whether they would be turning to the right or left. They indicate the direction of the turn by pressing on appropriate response keys, and continue as quickly as possible until each turn in the route has been negotiated. The total time taken to traverse the route serves as an index of the subjects' spatial orientation ability.

There are three key elements in the stimulus ensemble for this task. First a schematic city map must be developed, of sufficient complexity that it does not consist of rectangular blocks. Second, end points for the route must be defined, perhaps by employing an X to indicate a starting point and an arrow to indicate the end point. Third, the route to be traversed must be depicted.[1] The route should be clearly discriminable from the schematic map, either by being in a different color or by clearly distinguishable graphic style. Moreover, having the route disappear behind the subject as each turn is negotiated might reduce the chance of "getting lost." The subject's response is defined in terms of two keyboard responses, one defined to indicate a right turn, the other a left turn. The computer must keep track of these responses on-line, and be able to compare them to correct responses. In addition, the computer starts a clock simultaneously with the appearance of the route, and stops it when the final turn response is actuated.

The procedure in this task requires that the schematic map be displayed, with only a brief route visible. The instructions are read to the subject, and a brief practice is given on the route. The experimental trial begins when the 32-turn route is displayed all at once on the map, and requires that the subject make 32 key press responses to indicate left or right turns. A single trial may be adequate, but additional routes can easily be generated, keeping in mind the requirement for 32 turns, with half occurring in each direction.

The measurement requirements for this task require only that total response time be recorded, and that the protocol for left and right turns be recorded and compared to a standard to detect any errors.

---

[1]Further information on a form from which this task is derived may be obtained in: Money, J., Alexander, D., & Walker, H. T., Jr. A standardized road-map test of the direction sense. Baltimore, Md.: Johns Hopkins Press, 1965.

The instructions for this task are as follows:

> In this task we are concerned with the speed and accuracy with which you can determine direction with respect to your own position. On the display before you is a map of a city. Imagine you are taking a walk along the route indicated by the red path in the upper left-hand corner. You start at point X and go to the arrow. Your task is to decide at each turn which direction (left or right) you turned. If you turn left, press the response key on the left; if you turned right, press the right hand response key. Continue along the walk, indicating the direction of each turn until you reach the arrow. Work as quickly as you can without making mistakes.
>
> First we will try a practice on this short walk. Remember to indicate the direction of each turn, and to work as quickly as you can.
>
> Now we are ready for the actual task. A much longer walk will appear on the city map. Your task, though, stays the same. Imagine yourself at point X, and walk along the route, indicating the direction of each turn as quickly as possible until you reach the arrow.

Selective Attention

The subject is presented with two messages simultaneously, employing a dichotic recording to present a different message to each ear. In one part of the task the messages consist of series of words with digits embedded in the series, and in a later part the messages are pairs of digits, with one member of each pair presented to each ear at the same time. An indicator tone tells the subject which ear to attend to, and the subject's task is to indicate which digits occur in the relevant ear by pressing the corresponding digit key. Number of errors in terms of missed digits (omissions) and digits reported from the irrelevant ear (intrusions) serves as an index of the selectivity of the subject's attentional processes.

The stimulus ensemble for this task consists of a set of common English nouns (nearly 1,000 words would be required if none were repeated, but a smaller pool of 300 or fewer words should allow adequate spacing of repetitions), and the digits 0 to 9. On any given trial two messages are recorded using subsets of the stimulus ensemble. One message consists of 10 of the nouns and 6 of the digits, while the other message consists of 12 nouns not included in the first message, and the remaining 4 digits. The first part of these messages are arranged so that digits do not occur in the same ordinal location on two messages. In the interval between the first and second parts of the message, an additional noun is presented to each ear. In the second part of the messages, digits are presented simultaneously in each ear, with three pairs or six different digits constituting the second part of the message. In addition, a 250 Hz tone and a 2,500 Hz tone are required to signal the first and second parts and to indicate the relevant ear. The subject's response is to listen to the relevant ear and press a key corresponding to any digit which occurs on that ear (note that previous work required the subject to vocalize a response rather than press a key, and some speculation indicated that the vocalization itself interfered with the performance. The key press response is preferable for automated scoring, but other alternatives may have to be explored if error variance is too low.)

A trial in this task takes place in the following sequence: (a) a warning tone occurs for 200 msec, 250 Hz tone indicating right ear relevance, 2,500 Hz tone indicating left ear relevance. The tone is recorded monaurally, heard in the relevant ear only, (b) 1.5 seconds after the indicator tone, a series of 16 pairs of items are delivered to both ears at a rate of two items per second (the items are words and digits as described previously, and are synchronous, so that one member of each pair is presented to each ear simultaneously), (c) .5 second after the last pair in part one of the messages, an indicator tone occurs on the channel that was relevant for part one. The high and low tones have the same significance as at the beginning, so the tone will identify whether the relevant ear stays the same or switches for part two, (d) a pair of words is presented .5 second after the indicator tone, (e) .5 second after the last pair of words, three

pairs of digits are presented at a rate of two pairs per second, with one member of the pair presented to each ear, and (f) a 5-second interval occurs between the last pair of digits and the onset of the tone indicating the next trial. Twenty-four trials are employed per subject. The relevant ear may be either left or right, and either same or different in parts one and two, so the sequence of relevant ears for each trial could be right-right, right-left, left-right, and left-left. Six trials are presented in each of these sequences to constitute a 24-trial task. A brief rest is provided after 12 trials. Four additional trials are constructed, one for each sequence type. These are presented at the beginning of the test as unscored practice and may be repeated as necessary to ensure comprehension of the task.

The primary measurement concerns in this task involve the determination of th types of errors. The computerized station should have a record of the appropriate responses to the entire series of trials, and should record the subject's responses on-line. A comparison of these two protocols should allow quantification of errors and categorization of errors as omissions and intrusions. The types of errors can then be subcategorized according to which part of the sequence they occurred in and according to whether or not ear relevance had switched and what ear was actually relevant at the time of error.

A sample of instructions for this task follows:

> In this task we are interested in how well you can detect relevant information when irrelevant information is present. When you put on the headphones, you will hear a tone. If it is a high tone, it will be heard only on your left ear, and it indicates that you should attend to the message in the left ear and ignore the one in your right ear. Similarly, if it is a low tone, attend to the right ear and ignore the left. The tones sound like this (demonstrate).
>
> Shortly after the tone occurs, messages will begin in each ear. The messages will consist of rapidly presented words with occasional digits mixed in. When a digit occurs on the ear in which you are attending only, press the key corresponding to that digit. Do not respond to digits occurring in the other ear.
>
> After several seconds, another tone will occur, which you will hear only in the ear you are attending to. Again, if it is a high tone, attend to the left ear, and if it's a low tone, attend to the right ear. This may require you to change the ear to which you are attending in the middle of the trial.

> In this second part of the trial the messages will continue to both ears, with digits being presented simultaneously to both ears. Again, press the keys corresponding to the digits which occur in the relevant ear only, and ignore the other ear. When the messages stop, there will be a 5-second break, then the tone will sound again indicating the beginning of the next trial. Any questions?
>
> First we will do a few practice trials to give you a feeling for the task, and these will not be scored. Once you are familiar with the task, you will be scored on 24 trials, with a brief rest after the first 12. Any questions?

In addition to the computerized testing station, this task as presently implemented requires a tape recorder which can be controlled by the testing station, audio headphones through which the subject receives the stimuli, and a taped recording of the stimulus materials.

Time-Sharing

The subject in this task is required to perform a compensatory tracking task within specified limits of error, and to react as quickly as possible to light signals in a decision speed task. To perform the compensatory tracking task, the subject must anticipate the movement of a marker on a visual display, and operate a control stick to counteract the movement and keep the marker aligned with a fixed central point. When tracking performance is within acceptable limits, digits are presented at varying intervals. The subject is instructed to press a key corresponding to the lighted digit as quickly as possible, when this can be done without affecting performance on the tracking task. It is assumed that the subject has a fixed processing capacity and that a stable percentage of that capacity is devoted to the primary tracking task. A subject with a larger overall capacity, then, should react more quickly to the digits in the secondary decision speed task. Reaction time in the secondary task is therefore an index of the subject's time-sharing ability.

The stimulus ensemble for the compensatory tracking task consists of a visual display with a fixed center point and a moving marker or error cursor. The error cursor moves along the horizontal axis in the middle of the display, with a speed and eccentricity determined by the forcing function employed. A centrally mounted joystick is employed by the subject to control movement of the cursor. The subject's task is to keep the cursor aligned with the fixed center point. To do so, the subject must anticipate movement of the cursor and move the joystick in the direction opposite the cursor's movement.

For the secondary decision speed task, a light panel with 10 lights numbered 0 through 9 is employed. The stimulus ensemble consists of the digits 1 through 8, although on trial one, only 1 and 2 are employed. Trial two employs the digits 1, 2, 3, and 4, and trial three the complete stimulus ensemble. Below the lighted digits is a response keyboard consisting of 13 keys. One key is below each light and labelled to correspond to the light. An additional key is added at each end of the keyboard to prevent the subject from feeling the way to the end response keys. The thirteenth key is in a separate space and a central distance from the response keys. This "home key" is kept depressed until a response to a lighted digit is initiated.

In the first phase of this task, the subject practices on the tracking task until reaching performance within pre-specified error limits. After the subject has reached stable performance on the task, it can be made easier or more difficult through gradual adjustments of the forcing function. Once performance on the primary task is within tolerance, the secondary task can be introduced and time-sharing trials conducted.

A time-sharing trial consists of 40 secondary task stimuli presented during continuous primary tracking performance. The digits are randomly presented with the constraint that each digit occurs

equally often over the 40 presentations. The digits are presented at intervals of 5 to 15 seconds (with a rectangular distribution of inter-stimulus intervals) to prevent the subject from anticipating the presentation. After 20 stimuli have been presented, a 2-minute rest is provided prior to completing the trial. The three trials differ in the difficulty of the secondary task; trial one consists only of digits 1 and 2; trial two uses digits 1, 2, 3, and 4; and trial three uses digits 1, 2, 3, 4, 5, 6, 7, and 8.

During each trial, performance on the primary task is continuously monitored in a cross-adaptive technique. That is, if a 5-second running average error on the tracking task falls below the specified minimum error, the average inter-stimulus interval for presentation of secondary task stimuli is shortened. On the other hand, if average error becomes too high, no secondary stimuli are presented until the average returns to the specified limits. If the subject fails to respond to a secondary stimulus for 10 seconds, a warning buzzer sounds.

The dependent variable of primary concern in this task is the response time to secondary task stimuli. These response times must be recorded on-line, and the median, mean, and standard deviation of response time calculated for each trial. Also, the time during which tracking error exceeds the maximum specified tolerance may be recorded on-line, and total time in excess of the error limit calculated. To record response times, the test station must start a clock simultaneously with presentation of the secondary task stimulus, terminate the clock with actuation of the correct response key, and store the time recorded by the clock. Primary task performance is monitored by sampling the distance of the cursor from the center point and cumulating that distance over the trial (that sampling may occur 20 times per second, or less depending on availability of core). In addition, the 5-second running average of the error must be maintained and continuously compared to the specified performance tolerances, and decisions made to adjust secondary task presentation conditions based on primary task performance.

A sample of instructions for this task is as follows:

> In this task we are interested in how well you can perform two tasks simultaneously. For the primary task you will see on the screen a fixed center point and a marker moving to the left or right from the center point. The control stick in front of you changes the direction of the movement of the marker. The task is to move the control stick to keep the marker aligned with the center point. That is, if the marker is moving to the left, moving the control stick to the right will bring the marker back toward the center.

equally often over the 40 presentations. The digits are presented at intervals of 5 to 15 seconds (with a rectangular distribution of inter-stimulus intervals) to prevent the subject from anticipating the presentation. After 20 stimuli have been presented, a 2-minute rest is provided prior to completing the trial. The three trials differ in the difficulty of the secondary task; trial one consists only of digits 1 and 2; trial two uses digits 1, 2, 3, and 4; and trial three uses digits 1, 2, 3, 4, 5, 6, 7, and 8.

During each trial, performance on the primary task is continuously monitored in a cross-adaptive technique. That is, if a 5-second running average error on the tracking task falls below the specified minimum error, the average inter-stimulus interval for presentation of secondary task stimuli is shortened. On the other hand, if average error becomes too high, no secondary stimuli are presented until the average returns to the specified limits. If the subject fails to respond to a secondary stimulus for 10 seconds, a warning buzzer sounds.

The dependent variable of primary concern in this task is the response time to secondary task stimuli. These response times must be recorded on-line, and the median, mean, and standard deviation of response time calculated for each trial. Also, the time during which tracking error exceeds the maximum specified tolerance may be recorded on-line, and total time in excess of the error limit calculated. To record response times, the test station must start a clock simultaneously with presentation of the secondary task stimulus, terminate the clock with actuation of the correct response key, and store the time recorded by the clock. Primary task performance is monitored by sampling the distance of the cursor from the center point and cumulating that distance over the trial (that sampling may occur 20 times per second, or less depending on availability of core). In addition, the 5-second running average of the error must be maintained and continuously compared to the specified performance tolerances, and decisions made to adjust secondary task presentation conditions based on primary task performance.

A sample of instructions for this task is as follows:

> In this task we are interested in how well you can perform two tasks simultaneously. For the primary task you will see on the screen a fixed center point and a marker moving to the left or right from the center point. The control stick in front of you changes the direction of the movement of the marker. The task is to move the control stick to keep the marker aligned with the center point. That is, if the marker is moving to the left, moving the control stick to the right will bring the marker back toward the center.

While you are tracking the moving marker, occasionally one of these digits will light up. You will keep your left (non-dominant) hand on this center (home) key while tracking, but when the light comes on try to leave the home key and press the button corresponding to the digit which occurs. While it is important that you do this as quickly as possible, do not respond until you are sure it will not affect performance on the tracking task.

First we will practice on the tracking task alone, until you have learned the task and feel comfortable performing it. Then we will do three trials of combined task performance. You must maintain performance on the tracking task at the level we establish when you are practicing the tracking task alone. If you track too much better, the digits in the secondary task will be presented faster, whereas if tracking is too poor, no digits are presented in the secondary task. If you ignore the secondary task, a warning buzzer will sound. Since the trial will continue until a fixed number of digits have been presented, it is in your best interest to maintain tracking performance and respond to the digits, as quickly as possible without disrupting tracking.

On trial one, only the digit 1 or the digit 2 will be presented.

On trial two, the digits 1, 2, 3, or 4 will be presented.

On trial three, the digits 1, 2, 3, 4, 5, 6, 7, or 8 will be presented.

Remember to maintain tracking performance and then respond to the presented digits as quickly as possible without disturbing tracking performance.

Encoding Speed

Subjects are presented simultaneously with two letters and required to make a same-different judgment on the letter pair. This judgment may be based on physical identity (AA vs. Aa), name identity (Aa vs. AH), or category identity (vowels vs. consonants -- Ae vs. AH). The latency of the encoding judgment provides a measure of the speed of the encoding process. Moreover, latency differences indicate the speed of recoding; that is, the reaction time for the name identity judgments minus reaction time for physical identity judgments indicates the speed with which physical stimuli may be recoded to the level at which their name may be accessed.

A stimulus ensemble must be defined, consisting of both uppercase and lowercase examples of the four letters A, E, H, and T. On any given trial, two of these letters are selected for presentation. The letters selected for presentation are constrained so that 50% of the pairs result in a same judgment -- e.g., for the physical identity judgment, there are eight potential identical pairs -- AA, EE, HH, TT, aa, ee, hh, tt. Moreover each of these letters may be paired with one of the seven other items in the stimulus ensemble, resulting in 56 pairs to which the "different" response is appropriate (for the name identity judgment, there are 16 "same" pairs and 48 "different," and for the category judgment, 32 "same" pairs and 32 "different").

A block of trials might consist of 64 trials. For the physical identity judgment, this will include four repetitions of each of the eight identical pairs, plus 32 of the 56 possible different pairs. The order of presentation of the pairs is constrained so that the same pair does not occur twice consecutively, and so that the same response does not occur more than three consecutive times if the subject is responding correctly. Similarly, for the name identity judgment, each of the 16 same pairs is presented twice, plus 32 of the possible 48 different pairs, with the order constrained as above. The category identity judgment presents the 32 same pairs and the 32 different pairs once each.

The subject is seated approximately 2 feet from a cathode ray tube display, with a keyboard on which one key is designated for a same response and the other for a different response. At the beginning of a trial, the display will show two small Xs, one an inch to the right at the central fixation point, the other an inch to the left. Two seconds after the Xs appear, they will be replaced with the letter pair from the stimulus ensemble for that trial, one member of the pair replacing each X. The letters will be sufficiently large to be easily discriminated from the distance at which the subject is seated (perhaps about 1-inch high for subjects at a 2-foot distance from the display). When the letter pair is presented, a reaction time clock is started simultaneously and runs until the subject terminates it with his response. Similarly, the letters remain visible until the subject responds. When the subject has responded, a 5-second inter-trial interval begins, during which some form of feedback is

provided (perhaps as simple as a check to indicate correct response or an X to indicate incorrect response). Then the two Xs appear again, signalling the beginning of the next trial. Each block of trials requires 6 minutes, and each subject will complete three blocks -- physical, name, and category judgments. With a brief rest between blocks to explain the next decision rule, the task should require a total of about 25 minutes per subject.

Within each trial, reaction time to same or different judgments must be stored. When a block of trials has been completed, the mean, median, and standard deviation of the reaction time to same and to different judgments may be computed for correct responses, as well as the number of errors and error latency. When all three blocks have been concluded, reaction time differences may be computed (i.e., mean RT category - mean RT name).

At the beginning of the testing period, subjects are presented with 12 pairs of digits, 6 same and 6 different pairs, as a practice. These pairs are constrained only in that the same digit does not occur twice as a "same" pair, and the same response is not required on more than two consecutive trials. The other display characteristics are as described above.

Sample instruction are as described below.

> In this task you will be making a series of simple judgments. In the first condition we are concerned with how accurately and rapidly you can decide whether two displays are physically the same or different. You will see pairs of letters, one to the left side of the screen and one to the right side. Your job is to judge whether the two letters are physically identical or different. For example, the letter pair AA would be judged the same, the pairs AB or Aa different. If they are the same, press the right key, if they are different, press the left key.
>
> Before the letters appear two Xs will be displayed for a few seconds as a warning signal, and the letters will appear where the Xs are located. Please respond as rapidly as you can without making errors. The display will indicate whether your judgment was correct, and after a few seconds the Xs will reappear to signal that the next trial is ready to begin. Any questions?
>
> The first block of trials will be for practice, and will use digits instead of letter pairs.
>
> Name. In this second condition you will again see pairs of letters. This time you must decide whether the letters have the same <u>name</u>. For

example, the pair Aa would be the same, but Ae would be different. Please make the same-different response as quickly as possible while avoiding errors, using the same keys as before. Any questions?

Category. In this third condition you will decide whether the pair of letters belong to the same category. For example if they are both vowels -- Ae -- or both consonants -- Ht -- respond "same," but if one letter is from each category, respond "different." Any questions?

Mental Rotation

Subjects are presented sequentially with a pair of letters and asked to make a speeded same-different judgment. The letter pair may be either identical or mirror-images, and the pair may be either in the same orientation, or rotated in space with respect to each other. In order to perform the task, the subject must form a mental image of the first letter and perform a point-by-point comparison with the second. In addition, when the letters are rotated with respect to each other, the subject must mentally rotate the mental image into congruence with the second letter before undertaking the comparison. Reaction time and accuracy data are collected, and reaction time may be plotted as a function of degrees of rotation required to bring the letters into congruence. In the resulting monotonic function, the slope indicates the speed with which the subject accomplishes the rotation, and the intercept is indicative of the speed with which the other processes involved are performed.

The stimulus ensemble may be defined in terms of the three uppercase letter, G, F, and R, and their mirror-images. The total pool from which stimuli may be drawn is then elaborated by considering the orientation of the letters within the stimuli pairs. For each stimulus pair, the first letter may be in one of six positions, clockwise with respect to the vertical: 0^{o}, 60^{o}, 120^{o}, 180^{o}, 240^{o}, 300^{o}. The second letter may be oriented at one of four rotations with respect to the first: 0^{o}, 60^{o}, 120^{o}, or 180^{o}. Thus for each pair of stimulus letters, 24 pairs exist which may be rotated into congruence, and an additional 24 pairs exist which use the mriror-image letter and cannot be rotated into congruence. Using the three letters, a total of 144 pairs are defined, 50% of which require a same judgment.

The stimulus pairs are presented to the subject on a cathode ray tube display. The subject is seated about 2 feet from the display, with a response keyboard at the left of the central fixation point and remains on for 2 seconds to serve as a warning signal. This marker is replaced by the first letter of the stimulus pair, large enough to be easily visible, and displayed for 2 seconds. After a 1-second inter-stimulus interval the second letter appears 1 inch to the right of the central fixation point. The second letter remains visible until the subject's response terminates the display. A 3-second inter-trial interval ensues, during which the display gives feedback regarding the accuracy of the subject's judgment. This is followed by the reappearance of the warning marker signalling the beginning of the next trial. A block of trials will consist of 48 trials, including the 24 same pairs and 24 different pairs for each letter. The order of the pairs will be randomly determined with the constraint that the same response is not required on more than three consecutive trials. A block of trials should require just under 8 minutes, so with instructions, practice, and rest between blocks, this task should require about 30 minutes to complete.

Measurement in this paradigm requires that the identity of each pair be preserved, in terms of same-different judgments and in terms of degree or rotation. In addition, the individual subject reaction times must be stored for each pair. At the onset of the second letter of the stimulus pair, a reaction time clock is started, and is terminated when the subject makes a response. The reaction time and accuracy of the response are stored, and feedback provided. When the task is completed, mean reaction times may be plotted as a function of degree of rotation, and the slope and intercept of this function computed. In addition, accuracy of the judgments overall may be computed.

The subject should be provided with a brief practice, in which the stimuli consist of pairs of the singe capital letter R. The initial letter of each practice pair is oriented 0^0, 60^0, or 120^0 clockwise with respect to the vertical, and the second letter of each pair is rotated 60^0 or 120^0 with respect to the first. These six same pairs and six pairs utilizing the mirror-image are presented in random order as a practice block

The subject's instructions for performing this task are as follows:

> In this task, we are concerned with your ability to handle and compare letters in various orientations. You will see a warning marker, followed by the appearance of one letter, of which you should form a visual image to preserve its orientation. When it disappears, a second letter will be presented, and you must decide if its the same as the first. If the image of the first can be slid over the second to correspond, or mentally turned and then slid over the second to correspond, they are the same. For example, the pairs R and R are both the same, but the pairs R and $\overline{R}$ are different. If they are the same press the button on your right as quickly as possible. If they are different, press the button on the left. Remember to respond as quickly as possible without making mistakes.
>
> When you have responded, the second letter will disappear and the display will indicate whether or not your judgment was correct. After a brief delay, the warning signal will again appear, signalling the beginning of the next trial. Any questions?
>
> The first 12 trials will be a practice and will not be scored. Remember to form the image, then to mentally rotate it if necessary and slide it over the second letter to compare. Work as quickly as possible without errors.

Item Recognition

In the item recognition paradigm, a series of one to six digits is presented in a row on a cathode ray tube display, followed after a brief delay by a single digit. The subject is instructed to remember the initial series of digits, then to decide if the single digit is one of those presented in the initial series. The subject is instructed to push one button if the digit was in the series, another if not, and is instructed to make a response as quickly as possible without errors. The subject's reaction time is recorded and serves as an index of the efficiency with which the subject searches memory. That is, when reaction time is plotted as a function of the number of items to be searched, the slope of that function indicates the speed of search, and the intercept indicates the time taken by other processes.

A stimulus ensemble must be defined, consisting of the 10 digits 0 to 9. From this ensemble, a positive set is selected for each trial. The positive set consists of one, two, three, four, five, or six digits randomly selected from the stimulus ensemble. For each trial, a probe or test stimulus is also selected from the stimulus ensemble. Constraints on the selection of the probe stimulus are as follows: (a) on 50% of the trials, the probe must be a member of the positive set, (b) when the probe is a member of the positive set, each serial position within the positive set should be equally likely to contain the probed digit, so that over a series of trials each serial position is probed equally often, and (c) on trials when the probe is not a member of the positive set, it may be randomly selected from the negative set portion of the stimulus ensemble.

The digits which constitute the positive set for any given trial are presented to the subject all at once in a horizontal row, and remain visible for 5 seconds. A 2-second delay ensues, followed by a warning tone which occurs for 1 second, followed .5 second later by the probe or test stimulus. The test stimulus stays visible until the trial is terminated by the subject's response. The size of the digits presented on the CRT should be large enough to be easily visible from where the subject is seated, and should be presented sequentially in space across the screen. A warning tone occurs 3 seconds after a trial has been terminated, followed 1 second later by the presentation of the next positive set.

Two distinct buttons on a keyboard must be defined as the response buttons. The computer must start the reaction time clock simultaneously with the presentation of the probe stimulus, and terminate the reaction time clock at the press of the response button. For each trial, reaction time is recorded in milliseconds and stored along with the size of the positive set and the accuracy of the response.

A trial consists of the presentation of the positive set, followed by presentation of the test stimulus, and terminated by the

subject's response. The trials are presented in blocks of 24, with a brief rest between blocks. Each block contains four trials at each of the six positive set sizes, with two of those trials requiring a positive response and two a negative one. Within a block, trials are constrained so that the same positive set size and response (positive or negative) occur no more than twice consecutively. At least two blocks of trials should be presented (and or may be necessary) to insure that reliable data are collected.

The instructions presented to the subject for this task might be as follows:

> This task measures how quickly and accurately you can recognize items that you have just seen. On each trial, one or more digits will appear in a row on the screen. Try to remember each of the digits. A few seconds after they have disappeared, a tone will occur, followed immediately by a single digit on the right of the screen. You must decide whether this digit is one of those you just memorized. If the probe number was present in the memory list, press the button on your right. If not, press the button on your left. Please make your response as quickly as possible without making errors.
>
> The first 12 trials will be practice for you, and following the practice you will receive five blocks of 24 trials, with a short rest in between. Remember to work quickly while avoiding errors. Are there any questions?

Immediate/Delayed Memory

In this task, the subject is presented with a sequence of digits, and required to push a button corresponding to the item which occurred two digits previously. The task is presented in two parts. In the first part the digits are presented for 2 seconds followed by a 2-second inter-stimulus interval, so that 6 seconds pass between the offset of a digit and the response to that digit. In the second part, the inter-stimulus interval is 5 seconds so 12 seconds pass between the offset of a digit and the response to that digit. (That is, part one deals with immediate memory, part two with delayed memory). The number of correct responses in each part is an index of memory retrieval facility.

A stimulus ensemble is defined in terms of the digits 1 to 9. The actual arrangement of the stimuli within any part of the test is random except with the following restrictions: (a) no digit repeats without at least two intervening digits, and (b) each of the nine digits occurs exactly three times in each part of the test. Thus each part of the test contains 27 digits and requires 25 responses.

The response for the subject is defined in terms of a keyboard; that is, the keys corresponding to the digits 1 to 9 are indicated, and the subject is told to press the key corresponding to the digit occurring two items previously. For part one of the tasks, a warning signal appears for 2 seconds, followed by a digit presented 025 2 seconds. A 2-second inter-stimulus interval is followed by the next digit for 2 seconds, and so on until all 27 digits have been presented. At the occurrence of the third digit, subject presses the key corresponding to the first, and continues in that manner for the rest of the trial.

Part two of the study is exactly analogous to part one, except that the interval between digits is increased from 2 seconds to 5 seconds. The digits are of sufficient size to be easily discriminable to the subject, and the reponse keys easy to reach and distinguish. A brief rest is allowed on completion of part one before initiating part two, and during this rest, some feedback concerning accuracy of part one responses is provided.

The principal measure in this task is the accuracy of response (number correct out of 25 in each part), but, in addition, a reaction time measure could be used. This could measure time from the onset of the third digit to the activation of the appropriate response key. To obtain such a measure, subjects must be required to keep their hand on a home key until responding, and return to the home key on completion of the response.

Prior to the actual task subjects are given a practice, consisting of five digits and three responses, in order to insure that the instructions are understood. A sample of instructions for this task follows:

> In this task we are interested in how well you can remember digits in a complex memory situation. The task will take place in two parts, with a brief rest in between. In both parts you will see a warning signal followed by a sequence of digits presented one at a time. Your task is to press the button on this keyboard which corresponds to the item which occurred two previously. For example, if the sequence 6 4 8 2...occurs, you would make no response to 6 or 4, press 6 when 8 occurs, 4 when 2 occurs, and so on. The only difference between part one and part two is that a longer time separates the digits in part two, and you must remember the appropriate response for a longer time. Work as quickly as you can without making mistakes. Any questions?
>
> First we will try a brief practice with just a few digits. Remember, respond with the item that occurred two digits previously.

Decision-Making Speed

In the decision-making speed task, one of a number of alternative signals is presented to the subject, who is required to respond to the signal with the matching response as quickly as possible. The key to this task is the amount of uncertainty that must be resolved in order to make the response decision. When more alternative signals may potentially be presented, greater uncertainty exists and the decision is made more slowly. This task evaluates the extent to which the individual's decision-making speed is affected by uncertainty and is thus distinct from a perceptual speed task where uncertainty remains the same. This task consists of three parts: in part one, two potential signals and two responses are defined in part two, four potential signals and responses and in part three, eight potential signals and responses are defined. Reaction time as a function of the number of potential signals serves as an index of the subject's decision-making speed.

A stimulus ensemble is defined in terms of the eight digits 1 to 8. For part one, the digits 1 and 4 represent the potential signals; for part two, the digits 1, 3, 5, and 7 represent the potential signals; and for part three, the entire stimulus ensemble -- 1, 2, 3, 4, 5, 6, 7, and 8 -- represent potential signals. A trial is defined by the occurrence of a stimulus coupled with a response by the subject. In each part of this task, 24 trials are performed. The determination of which signal occurs on a given trial is random, with the constraint that each signal occurs an equal number of times (12 for each alternative in part one, 6 for each alternative in part two, and 3 for each alternative in part 3).

A 13-key response panel is employed, with 12 keys arranged in a shallow half-circle and the other key equidistant from each of those 12. Ten of the keys correspond to the digits 0 to 9, and an additional blank key is provided at each end. The subject is required to keep the central "home key" depressed until a stimulus occurs. At this point the subject must release the home key and activate the key which corresponds to the digit presented.

For each trial, a warning signal is presented 2 seconds after the home key is depressed, followed after a brief interval (2 seconds) by the stimulus. When the subject leaves the home key to respond, the signal is erased. The subject responds by depressing the key representing the signal which occurred and is then presented with immediate feedback concerning the accuracy of the response. When the subject returns to the home key, the next trial is initiated. The signal must be large enough to be easily distinguished from where the subject is seated, and is always presented at the central fixation point on the CRT screen.

A large number of response measures may be collected from the task. In addition to error data, latency measures may be gathered from onset of the signal to release of the home key, and from onset of

the signal to the occurrence of the response. These latency measures must be collected and stored in real time -- a reaction time clock starting with the onset of the signal and terminating with the release of the home key, and another starting with the signal and terminating with the response. Over the entire task, some measure of the latency increase as a function of number of alternatives may be calculated (e.g., the slope of the function relating median response time to number of alternatives).

A brief practice of two trials is provided prior to each part of this task, with the number of alternative potential signals appropriate to the upcoming part. A brief rest is provided between each part, during which the subject is reminded of how many alternatives will be provided in the next part. A sample of the subject's instructions follows:

> In this task we are interested in how quickly you can recognize which of several potential signals has occurred and decide which response is appropriate. The task will have three parts, with a brief rest between each. In part one, the signal will be either the digit 1 or the digit 4. When you press the home key, a warning signal (asterisks) will appear, followed briefly by the signal -- either 1 or 4. You must leave the home key and press the numbered button which corresponds with the signal. Do not leave the home key until you are sure of the signal's identity, since it will disappear when the home key is released. When you have responded, the CRT will display feedback about the correctness of your response. Then return to the home key to initiate the next trial. The first two trials in each part are practice and will not be scored, but are to ensure that you understand the task. Work as quickly as possible while avoiding errors. Any questions?
>
> The procedure for part two is identical to that for part one, except the potential signals include the digits 1, 3, 5, and 7, and the responses are the numbered keys corresponding to those signals.
>
> The procedure for part three is identical to that for part two, except the potential signals include the digits 1, 2, 3, 4, 5, 6, 7, and 8, and the responses are the numbered keys corresponding to those signals.
>
> The procedure for part three is identical to that for part two, except the potential signals include the digits 1, 2, 3, 4, 5, 6, 7, and 8, and the responses are the numbered keys corresponding to those signals.

The subject is presented with a verbal decision problem containing two types of information. Base rate information defines the distribution of two mutually exclusive events, such that one of the two events has to occur (i.e., their summed probability is equal to 1). Diagnostic information defines the likelihood of a third event happening when either of the first two events is true. The subject's task is to combine the base rate and diagnostic information to estimate the probability of one of the mutually exclusive base rate events having occurred. The accuracy of this estimate serves as an indicator of the subject's ability to estimate probabilities. A sample problem is presented below.

> Two cab companies operate in a given city, the Blue and the Green (according to the color of cab they run). 85% of the cabs in the city are Blue, and 15% are Green. A cab was involved in a hit-and-run accident at night, in which a pedestrian was run down. The injured pedestrian later testified that though he did not see the colo of the cab, due to the bad visibility conditions that night, he remembers hearing the sound of an intercom coming through the cab window. The police investigation discovered that the intercoms are installed in 80% of the Green cabs, and in 20% of the Blue cabs.
>
> What do you think are the chances that the hit-and-run cab was Green?

In this sample, the base rate information concerns the distribution of cab colors in the city -- 85% Blue and 15% Green. This generalizes to the form $p(A) = p$, $p(B) = 1 - p$. The diagnostic information concerns the likelihood that an intercom has been installed for each color of cab -- 80% of the Green cabs and 20% of the Blue cabs have intercoms. The diagnostic information generally takes the form $p(C/A) = m$, $p(C/B) = n$, where m and n are free to vary from 0 to 1. The pedestrian in the sample confirms the presence of an intercom, which essentially generalizes to the statement $p(C) = 1$. The statement of the problem, the probability that the hit-and-run cab was Green, essentially asks the subject to estimate $p(B/C)$. All of the problems have this same general form.

The stimuli consist of brief scenarios like the one just described. The base rate data for each problem specify two alternatives which constitute the complete universe of possible outcomes and the frequency with which each of those potential outcomes occurs. The probability that one of the two events occurs is equal to one. These base rate probabilities are free to vary from .10 to .90, with the constraint that the two probabilities must total 1.0. The diagnostic information gives conditional probabilities for a third or diagnostic event; that is, it states the likelihood of the diagnostic

event if the first of the base rate alternatives is true and the likelihood if the second alternative is true. The diagnostic event is then confirmed, and the subject asked to estimate the conditional likelihood of one of the base rate alternatives. The diagnostic information need not be of the form p(C/A) + p(C/B) = 1.0; it would be permissible for 80% of the Green cabs and 40% of the Blue cabs to have intercoms. The diagnostic probabilities may range from .2 to .8, may change in increments of .05, and must be different by at least .20. (That is, probabilities of .40 and .80 are permissible, but not .90 and .10, .42, and .78, or .60 and .50). The subject responds by punching a keyboard with a number from 0 to 100, indicating the best estimate that the event in question occurred.

The actual content of the problems may be amenable to some more face valid situations. For example, "A pilot hears a change in engine pitch. Eighty-five percent of the time that change in pitch is meaningless, but 15% of the time it indicates a potential problem. When that problem occurs, a dash light blinks 80% of the time, whereas it blinks only 20% of the time when no problem occurs. The dash light is blinking. What is the probability that a problem has developed?"

A problem trial consists of a brief warning, followed immediately by the presentation of the problem in a text format like the example above. The problem remains visible for 1.5 minutes or until the subject terminates the trial by responding with the estimate. (The actual upper limit for time of display may vary due to the length of the problem, but must be long enough to allow the problem to be comprehended while forcing the subject to estimate rather than compute probabilities). After a 5-second inter-trial interval, the warning signal occurs for the next trial. Eleven different trials are constructed. One serves as an unscored practice, and after a pause for questions, the 10 test problems are presented without interruption.

The primary dependent measure for this task is the accuracy of the subject's probability estimation. For each problem specified, the computer must have a record of the mathematically correct probability. It then records the subject's response to each problem, and computes two parameters for the test. For one, the absolute value of the difference between the subject's response and the correct solution is cumulated over problems to indicate total error without regard to over- or under-estimation. For the second parameter, the signed difference values are algebraically summed over trials to indicate a central tendency and direction of error of estimation.

The instructions for this task are as follows:

> In this task we are interested in how well you can estimate probabilities in different situations. For each trial a short problem will be displayed on the screen. Read the problem and estimate the answer to the problem as quickly as possible. Enter your answer on the keyboard before you by punching a number from 0 to 100. Zero indicates no

chance that the questioned event occurred, 50 indicates a 50% chance that it occurred, and 100 indicates a 100% likelihood that the event occurred.

The first problem is a practice and will not be scored. When you are sure you understand the task, we will continue. You will see a warning signal on the screen, followed immediately by the problem. Read the problem, decide on an appropriate response, and respond as quickly as possible. If you wait too long, the problem will disappear, so make your response quickly. Once you have responded, the problem will disappear, and after a brief interval the warning signal will occur, signalling the onset of the next problem. You will do 10 problems in all. Any questions?

Risk-Taking

In the risk-taking task, the subject is presented with a matrix of 10 boxes (in two rows of 5) and is told that 9 of the boxes contain a reward, whereas the other box is a disaster box. The subject is allowed to select the boxes, one at a time. If the selected boxes contain a payoff, the subject gets to keep it, but if it is the disaster box, the subject loses all of the payoff acquired. The average number of boxes selected provides an index of the subject's propensity for taking risks when making decisions.

The stimulus ensemble for this task is defined in terms of the 10 boxes. The boxes should be large enough to contain an identifying number and be readily discriminable, but small enough to be readily displayed on a cathode ray tube. They may be arranged in a 2 x 5 matrix, with the top row numbered 1, 2, 3, 4, and 5, and the bottom row 6, 7, 8, 9, and 0 (use 0 for 10 so a single keystroke can identify each box). The stimuli for feedback are a cumulative total of the amount won to that point; i.e., if a subject chooses box one and it is not a disaster box, it is replaced by $10. If box two is safely chosen, it is replaced by $20, and so on. If the disaster box is selected, an X appears, and the preceding cumulative payoff is erased. Prior to each trial, the test station randomly selects a digit to determine which box is identified as the disaster box. The subject responds in this task by pressing the key identified with each box selected. Or the subject may opt to select no further boxes, and keep any payoff acquired to that point, and can indicate this decision by pressing a carriage return.

A trial in this task is defined by a warning signal, followed immediately by presentation of the matrix of numbered boxes. The subject selects a box (since at that point there is no risk of loss since no payoff has been accumulated). The contents of the box are immediately displayed, and the subject must decide whether to continue the trial by selecting another box, or end the trial by pressing carriage return (unless, of course, the first box contained the disaster on that particular trial). The trial continues until termination either by carriage return or selection of the disaster box. Accumulated payoff is displayed in the box matrix within each trial, and total accumulated payoff displayed between trials (selection of the disaster box affects only payoff accumulated within that trial, and does not eliminate payoff accumulated on previous trials). The next trial begins after a brief (10-second) inter-trial interval, until a block of 10 trials has been completed. The block of 10 trials is followed by a brief rest, and then another block of 10, for a total of 20 trials.

The measurement of primary interest is the average number of boxes selected, which indexes the level of risk the subject would tolerate (the likelihood of selecting the disaster box increases form 10 percent to 11.1, 12.5, 14.3, 16.7, 20.0, 25.0, 33.0, to 50 percent, while simultaneously the potential for loss increases from 0 to 80, in

$10 increments, for each successive selection). Additional measures with potential validity could examine the effect of the outcome of the previous trial on number of boxes selected, and the effect of total accumulated payoff on number of boxes selected. In addition, the latency of each decision can be recovered. The testing station must keep a record on-line of the number of boxes selected within each trial, and store that number at the end of each trial.

A sample of instructions for the risk-taking task follows.

> In this task we are concerned with how you make decisions. Following a brief warning signal, a set of boxes numbered from 1 to 10 will appear on the screen (box 10 is labeled with a 0). Nine of these boxes are worth $10 each, but the tenth box is a disaster box. You must select a box, and receive its contents, then decide whether to quit and keep those contents, or select another box and risk losing the contents of the first. That is, if you select boxes 7 and 3 and are safe, you will have $20. You can quit and keep the $20. If you select box 5 and it contains the disaster, you lose the $20 and the trial is over.
>
> If you decide to select a box, press the key indicated for selection. If you decide to keep what you have accumulated, terminate the trial by pressing the carriage return key. The box which contains the disaster is randomly determined. Your objective is to accumulate as much money as possible. This requires you to balance the number of boxes selected against the risk of selecting the diaster box.
>
> When you select a box, either an amount will appear telling you how much you've won so far in the trial, or an X will appear, the dollars will be erased, and the trial will end. Whether the trial ends by disaster or by your electing to select no more boxes and keep the payoff, a short interval will be followed by the recurrence of the warning signal denoting the beginning of the next trial. In that interval, your payoff on that trial will be added to your accumulated payoff on previous trials, and your total winnings displayed. A block of 10 trials will occur in this manner, followed by a brief rest and another block of 10 trials. Are there any questions?

Embedded Figures

The subject is presented with a simple geometric figure and two complex geometric figures. This task is to decide which of the two complex figures has the simple figure embedded within it and to indicate a choice by pressing the button corresponding to that figure. The speed and accuracy of the response serves as an indicator of that subject's level of field-independence of field-dependence.

The stimulus ensemble is defined in terms of a set of simple geometric figures and pairs of corresponding complex geometric figures. For each pair of complex geometric figures, one has the simple figure embedded within it and the other does not.[2]

The figures for each presentation must be labeled; for example, the simple geometric figure is labeled B, and the alternative complex figures are labeled B-1 and B-2. The response required of the subject is to push a button -- one or two -- indicating which complex figure contains the simple figure. The subject must be seated in front of the display screen, at a distance to allow easy discriminability of the figures. Also, the distance must allow the subject to rest one finger of each hand comfortably on the defined response buttons.

A trial in this task consists of the presentation of a warning signal on the screen, followed by a display including a single simple figure and its corresponding alternative embedding figures. The three figures and their labels are displayed simultaneously and remain on the screen until the subject actuates one of the response buttons. When a response is made, the display is erased, and a 5-second inter-trial interval ensues during which feedback concerning the accuracy of the response is provided. The warning signal then appears again to signal the next trial. A total of 26 such trials are defined. The first two serve as practice, are not scored, and are followed by a brief break during which questions may be answered. The final 24 trials are then presented continuously and scored.

The primary dependent measures are the accuracy and speed of the subject's response. To measure accuracy, the test station must have a stored list of the correct response for each trial, keep track of the subject's response, and compare it to the stored correct response. The total correct responses are then cumulated over the 24 testtrials. To measure the speed of response, the test station must start a reaction time clock synchronously with presentation of the stimulus array, terminate the clock with the actuation of the response button, and record the elapsed time. Mean, median, and standard deviation of the response times may then be computed over the 24 test trials.

---

[2]For examples of embedded figures, see Witkin, H. A., Oltman, P. K., Raskin, E., and Karp, S. A. A Manual for the Embedded Figures Test. Palo Alto, CA: Consulting Psychologists Press, 1971.

A sample of the instructions for this task follows:

> In this task we are interested in how well you can pick out simple geometric figures when they are embedded in more complex figures. A warning signal will occur prior to each trial, followed shortly by the display. The display will consist of a simple geometric figure labeled with a letter (A) and two complex figures labeled correspondingly (A-1 and A-2). You must decide which complex figure contains the simple figure for each display.
>
> Place the index finger of your left hand on the key labeled 1, and the index finger of your right hand on the key labeled 2. When you have decided which complex figure contains the simple figure, press the corresponding button. That is, if A-1 contained the simple figure, press 1 with your left hand. When you have responded, the figures will disappear and the screen will indicate whether or not you were correct.
>
> The first two trials are practice, after which any questions will be answered. The next 24 trials will be scored, and are presented without any interruption. Are there any questions?

SUMMARY AND IMPLICATIONS

The objective of this research was to identify a batter of performance tasks tapping perceptual-motor abilities and cognitive processes, which could be useful for the selection of candidates for undergraduate pilot training. The activities undertaken are represented schematically in Figure 1. A large number of candidate tasks were identified, and a framework relating flying performance to specific abilities and processes was developed. A preliminary task battery was then generated, conceptually linked to flying performance. Psychometric and economic considerations mandated that this initial battery be reduced to a more manageable size. A final battery of 15 tasks was identified and is presented below.

1. Perceptual Speed
2. Complex Coordination
3. Compensatory Tracking
4. Kinesthetic Memory
5. Route Walking
6. Selective Attention
7. Time Sharing
8. Encoding Speed
9. Mental Rotation
10. Item Recognition
11. Immediate/Delayed Memory
12. Decision Making Speed
13. Probability Estimation
14. Risk Taking
15. Embedded Figures

Although the objective of identifying and specifying the task battery was satisfied, there is much more research required prior to implementing the battery. First, the mechanics of each task paradigm must be carefully tested to insure that the task parameters maximize the discriminability of each task as a selection device. Parameters such as display time, inter-stimulus interval, inter-trial interval, and number of trials should be empirically established, and alternative performance measures (accuracy, latency, etc.) examined to determine which are most useful. Considerations about total test time and ease of implementing the tasks may require changes in the task, to insure that the test is cost-effective for the Air Force. Similarly, the sequencing of the tasks may affect both the validity of the battery and its motivational level. The tasks could be arranged according to similarity of the response required, but the ultimate testing sequence, like the other points raised here, should be empirically determined by careful pilot testing of the task battery.

Once the task battery has been implemented and refined as described, its validity must be evaluated. First of all, an adequate experimental design for testing the predictive validity of the battery must be constructed. An important consideration in testing the validity of the battery is the selection of criteria against which the

battery will be validated. The criteria should be clearly relevant to pilot performance and should show enough variability in performance that real differences between pilot trainees may actually be distinguished. The validity of each task as well as that of the battery as a whole must be examined. While the battery does contain 15 tasks, it is unlikely that each will demonstrate significant predictive validity. Some tasks may have to be replaced with alternative tasks from the pool of candidate tasks, and others merely omitted because of lack of validity. The final task battery resulting from the pilot testing and validity testing is likely to contain fewer tasks than does the battery presented here.

If either cost-effectiveness considerations or empirical considerations should mandate a further reduction of the size of the task battery, several options should be considered. Although each task is selected to tap a single ability or process, several have similar components. For example, a measure of perceptual speed may be obtained from the decision-making speed task. There are other similarities among these tasks which may be used to advantage in implementing them; the displays for compensatory tracking and complex coordination may be made very nearly identical. It should be noted that one task - time-sharing - is actually a combination of the compensatory tracking and decision-making speed tasks; the compensatory tracking task may be used as the baseline for time-sharing performance. The complex coordination task might be considered a form of perceptual-motor time sharing. All of these factors may be useful in streamlining or reducing the task battery.

In addition to the implications for refining and validating this battery, several considerations emerge from the unique nature of this task battery. While judgment and decision-making processes are consistently rated as extremely important for pilot performance, few, if any, attempts have been made to directly test a candidate's decision-making processes. The inclusion of several decision-making tasks in ths battery is unique; indeed, decision-making tasks have less prior history as testing devices for any purpose than any of the other included tasks. Because of the importance of the decision making process to flying, and the relative lack of previous testing of these tasks, the decision-making tasks require special attention. Careful development and testing of these tasks, and considerations of alternative decision-making tasks, is warranted, even if initial results are not conclusive.

A general feature of this task battery is its concern with cognitive processes and their role in limiting human performance. There is considerable evidence for the predictive potential of cognitive tasks, and emphasis on such tasks is consonant with the increasingly heavy cognitive processing load in piloting an aircraft. However, the study of individual differences in cognitive processes has lagged behind more developed studies of perceptual-motor and physical abilities, and indeed the processes themselves are less well specified and differentiated. A great deal of basic research is beginning to be devoted to individual differences in cognitive

processes, and a large number of applications to testing and training are apparent. While this report represents the state of the art in the study of individual differences in cognitive processes, the amount of research currently being devoted to the field ensures that refinement of the theory of cognitive processes will continue. Moreover, the role of different processes in the performance of a variety of tasks will continue to be specified. This suggests that a periodic updating of the present data base, to take into account the rapid expansion of research in the field, would be of considerable value in the evolution of this task battery.

The present task battery combines traditionally effective tests from the perceptual-motor domain and attention tasks, with state-of-the-art thinking about cognitive psychology and cognitive tasks, and their ability to reflect important differences in human performance. These tasks typically lack a psychometric developmental history, so considerable care, patience, and insight are required to maximize the potential of this task battery.

REFERENCES

Adams, J. A., & Creamer, L. R. Anticipatory timing of continuous and discreet responses. Journal of Experimental Psychology, 1962, 63, 84-90.

Allen, G. L., Siegel, A. W., & Rosinski, R. R. The role of perceptual context in structuring spatial knowledge. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 617-630.

Allen, T. W., Rose, A. M., & Kramer, L. J. An information processing approach to performance assessment: III. An elaboration and refinement of an information processing performance battery. AIR-58500-11/78-TR. Washington, D. C.: American Institutes for Research, November 1978.

Alluisi, E. A., & Chiles, W. D. Sustained performance, work-rest scheduling, and diurnal rhythms in man. Acta Psychologica, 1967, 27, 436-442.

Ambler, R. K., & Guedry, F. E., Jr. The validity of a brief vestibular disorientation test in screening pilot trainees. NAMI-974, NASA Order R-47. Pensacola, Florida: Naval Aerospace Medical Institute, 1965.

Anders, T. R. Retrospective reports of retrieval from short-term memory. Journal of Experimental Psychology, 1971, 90, 251-257.

Anderson, L. W. An empirical investigation of individal differences in time to learn. Journal of Educational Psychology, 1976, 68, 226-233.

Atkinson, R. C., Holmgren, J. E., & Juola, J. F. Processing time as influenced by the number of elements in a visual display. Perception and Psychophysics, 1969, 6, 321-327.

Bale, R. M., & Ambler, R. K. Application of college and flight background questionnaires as supplementary noncognitive measures for use in the selection of student naval aviators. Aerospace Medicine, 1971, 42, 1178-1181.

Bar-Hillel, M. The base-rate fallacy in probability judgment. Technical Report PTR-1042-77-4/1. Eugene, Oregon: Perceptronics, April 1977.

Baron, J. Phonemic stage not necessary for reading. Quarterly Journal of Experimental Psychology, 1973, 25, 241-246.

Baron, J., & McKillop, B. J. Individual differences in speed of phonemic analysis, visual analysis, and reading. Acta Psychologica, 1975, 39, 91-96.

Baxter, T. D. Predicting undergraduate pilot training (UPT) performance for Air Force Academy graduates. SLR-TR-78-004. USAF Academy, CO: Frank J. Seiler Research Laboratory, July 1978.

Beach, L. R., & Mitchell, T. R. A contingency model for the selection of decision strategies. Academy of Management Review, 1978, 3.

Beggs, W. D. A., Sakstein, R., & Howarth, C. I. The generality of a theory of intermittent control of accurate movements. Ergonomics, 1974, 17, 757-769.

Berkshire, J. R., & Ambler, R. The value of indoctrination flights in the screening and training of Naval aviators. Aerospace Medicine, 1963, 34, 420-423.

Berlucci, G., Crea, F., diStefano, M., & Tassineri, G. Influence of spatial stimulus-responses compatibility on reaction time of ipsilateral and contralateral hand to lateralized light stimuli. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 505-517.

Bloom, B., & Broder, L. The problem solving processes of college students. Chicago: University of Chicago Press, 1950.

Bransford, J. D., Barclay, J. R., & Franks, J. J. Sentence memory: A constructive versus interpretive approach. Cognitive Psychology, 1972, 3, 193-209.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Brecke, F. H., Gerlach, V. S., & Shipley, B. D. Effects of instructional cues on complex skill learning. AFOSR-TR-75-0201. Arlington, Virginia: Air Force Office of Scientific Research, August 1974.

Broadbent, D. E., & Gregory, M. Donders B and C-reactions and S-R compatability. Journal of Experimental Psychology, 1962, 63, 575-578.

Broadbent, D. E., & Gregory, M. Vigilance considered as a statistaical decision. British Journal of Psychology, 1963, 54, 309-323.

Brown, J. S., & Slater-Hammel, A. T. Discrete movements in the horizontal plane as a function of their length and direction. Journal of Experimental Psychology,. 1949, 39, 84-95.

Bryden, M. P. Perceptual asymmetry in vision: Relation to handedness, eyedness, and speech lateralization. Cortex, 1973, 9, 419-435.

Carpenter, P. A., & Just, M. A. Sentence comprehension: A psycho-linguistic processing model of verification. Psychological Review, 1975, 82, 45-73.

Carroll, J. B. Problems of measurement related to the concept of learning for mastery. Educational Horizons, 1970, 48, 71-80.

Carroll, J. B. Psychometric tests as cognitive tasks: A new "structure of intellect." Princeton, NJ: Educational Testing Service, 1974.

Carroll, J. B. Psychometric tests as cognerpretive approach. Cognitive Psychology, 1972, 3, 193-209.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Brecke, F. H., Gerlach, V. S., & Shipley, B. D. Effects of instructional cues on complex skill learning. AFOSR-TR-75-0201. Arlington, VA: Air Force Office of Scientific Research, August 1974.

Broadbent, D. E., & Gregory, M. Donders B and C-reactions and S-R compatability. Journal of Experimental Psychology, 1962, 63, 575-578.

Broadbent, D. E., & Gregory, M. Vigilance considered as a statistical decision. British Journal of Psychology, 1963, 54, 309-323.

Brown, J. S., & Slater-Hammel, A. T. Discrete movements in the horizontal plane as a function of their length and direction. Journal of Experimental Psychology, 1949, 39, 84-95.

Bryden, M. P. Perceptual asymmetry in vision: Relation to handedness, eyedness, and speed lateralization. Cortex, 1973, 9, 419-435.

Carpenter, P. A., & Just, M. A. Sentence comprehension: A psycho-linguist processing model of verification. Psychological Review, 1975, 82, 45-73.

Carroll, J. B. Problems of measurement related to the concept of learning for mastery. Educational Horizons, 1970, 48, 71-80.

Carroll, J. B. Psychometric tests as cognitive tasks: A new "structure of intellect." Princeton, NJ: Educational Testing Service, 1974.

Carroll, J. B. Psychometric tests as cognitive tasks: A new "structure of the intellect." In L. Resnick (Ed.), The nature of intelligence. Hillsdale, NJ: Lawrence Erlbaum Associates, 1976.

Cavanaugh, J. P. Relation between immediate memory span and memory search rate. Psychological Review, 1972, 79, 525-530.

Cherry, E. C. Some experiments on the recognition of speech, with one and with two ears. Journal of the Acoustical Society of America, 1953, 25, 975-979.

Chiang, A., & Atkinson, R. C. Individual differences and interrelationships among a select set of cognitive variables. Memory and Cognition, 1976, 4, 661-672.

Christina, R. W. Influence of enforced sensory and motor sets on reaction latency and movement speed. Research Quarterly, 1973, 44, 483-487.

Clark, H. H., & Chase, W. G. On the process of comparing sentences to pictures. Cognitive Psychology, 1972, 3, 472-517.

Clifton, C., & Birenbaum, S. Effects of serial position and delay of probe in a memory scan task. Journal of Experimental Psychology, 1970, 86, 69-76.

Cohen, G. Hemispheric differences in serial versus parallel processing. Journal of Experimental Psychology, 1973, 97, 349-356.

Cohen, J. Chance, luck, and skill. Baltimore, MD: Penguin, 1960.

Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 240-247.

Conrad, R. Very brief delays of immediate recall. Quarterly Journal of Experimental Psychology, 1960, 12, 45-47.

Coombs, C. H. Portfolio theory and the measurement of risk. In M. F. Kaplan, & S. Schwartz, (Eds.), Human judgment and decision processes. New York: Academic Press, 1975.

Cooper, L. A. Individual differences in visual comparison processes. Perception and Psychophysics, 1976, 19, 433-444.

Cooper, L. A., & Shepard, R. N. Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

Craig, A. Broadbent and Gregory revisited: Viligance and statistical decision. Human Factors, 1977, 19, 25-36.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.

Crolotte, A., & Saleh, J. Development of a decision taxonomy for the Marine command and control environment: Volume I. Technical Report PATR-1069 79-6. Woodland Hills, CA: Perceptronics, Inc., June 1979.

Crosby, J. V. Cognitive pretraining: An aid in the transition from instrument to composite flying. AFHRL-TR-77-62, AD-A048 816. Williams AFB, AZ: Air Force Human Resources Laboratory, October 1977.

Damos, D. C. Residual attention as a predictor of pilot performance. Human Factors, 1978, 20, 435-440.

Das, J. P., Kirby, J., & Jarman, R. F. Simultaneous and successive synthesis: An alternative model for cognitive abilities. Psychological Bulletin, 1975, 82, 87-103.

Dimond, S. Hemispheric refractoriness and control of reaction time, Quarterly Journal of Experimental Psychology, 1970, 22, 610-617.

DiVesta, F. J., Ingersoll, G., & Sunshine, P. A factor analysis of imagery tests. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 471-479.

Dorfman, R. W., & Goldstein, I. L. The effects of task coherency, preview, and speed-stress on timing and anticipation. Journal of Motor Behavior, 1975, 7, 45-55.

DuBois, P. H. (Ed.) The classification program (AAF Aviation Psychology Program Research Report No. 2). Washington, D. C.: U.S. Government Printing Office, 1947.

Earhard, M. Individual differences in subjective organization. Psychonomic Science, 1970, 18, 220-221.

Earhard, M. Free recall, transfer, and individual differences in subjective organization. Journal of Experimental Psychology, 1974, 103, 1169-1174.

Earhard, M., & Endicott, O. Why are there individual differences in subjective sequential organization during free recall memorization? Journal of Verbal Learning and Verbal Behavior, 1969, 8, 316-319.

Edwards, W. Behavioral decision theory. Annual Review of Psychology, 1961, 12, 473-498.

Egan, D. E. Testing based on understanding: Implications from studies of spatial abilities. Intelligence, 1979, 3, 1-16.

Emurian, H. H. A multiple-task performance battery presented on a cathode ray tube. JSAS Catalog of Selected Documents in Psychology, 1978, 8, 102.

Ernest, C. H., & Paivio, A. Imagery and verbal associative latencies as a function of imagery ability. Canadian Journal of Psychology, 1971, 25, 83-90.

Estes, W. K. Research and theory on the learning of probabilities. Journal of the American Statistical Association, 1972, 67, 81-102.

Fennel, E., Satz, P., Abell, T. V. D., Bowers, D., & Thomas, R. Visuospatial competency, handedness, and cerebral dominance. Brain and Language, 1978, 5, 206-214.

Fernandes, K., & Rose, A. M. An information processing approach to performance assessment: II. An investigation of encoding and retrieval processes in memory. AIR-58500-11/78-TR. Washington, D. C.: American Institutes for Research, November 1978.

Fischer, G. W. Multidimensional utility models for risky and riskless choice. Organizational Behavior and Human Performance, 1976, 17, 127-146.

Fischhoff, B., & Fulero, S. What makes a good explanation? Technical Report PTR-1060-77-11. Eugene, OR: Perceptronics, November 1977.

Fischhoff, B., Slovic, P., & Lichtenstein, S. Improving intuitive judgments by subjective sensitivity analysis. Technical Report PTR-1060-78-7. Eugene, OR: Perceptronics, July 1978.

Fitts, P. M. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 1954, 47, 381-391.

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

Fleishman, E. A. Testing for psychomotor abilities by means of apparatus tasks. Psychological Bulletin, 1953, 50, 241-262.

Fleishman, E. A. Evaluations of psychomotor tests for pilot selection: The direction control and compensatory balance tests. Technical Report 54-131. Lackland AFB, TX: USAF Personnel Training and Research Center, 1954.

Fleishman, E. A. Psychomotor selection tasks: Research and application in the United States Air Force. Personnel Psychology, 1956, 9, 449-467.

Fleishman, E. A. Factor structure in relation to task difficulty in psychomotor performance. Educational Psychological Measurement, 1957, 17, 522-532.

Fleishman, E. A. Dimensional analysis of movement reactions. Journal of Experimental Psychology, 1958, 55, 430-453.

Fleishman, E. A. The structure and measurement of physical fitness. Englewood Cliffs: Prentice-Hall, 1964.

Fleishman, E. A. On the relation between abilities, learning, and human performance. American Psychologist, 1972, 27, 1017-1032.

Fleishman, E. A. Toward a taxonomy of human performance. American Psychologist, 1975, 30(12), 1127-1130.

Fleishman, E. A., & Helpel, W. E. A factor analysis of dexterity tests. Personnel Psychology, 1954, 7, 19-32.

Fleishman, E. A., & Hempel, W. E. The relation between abilities and improvement with practice in a visual discrimination reaction task. Journal of Experimental Psychology, 1055, 49, 301-312.

Fleishman, E. A., & Hempel, W. E. Factorial analysis of complex psychomotor performance and related skills. Journal of Applied Psychology, 1956, 40, 96-104.

Fleishman, E. A., & Ornstein, G. N. An analysis of pilot flying performance in terms of component abilities. Journal of Applied Psychology, 1960, 44, 146-155.

Fleishman, E. A., & Rich, S. Role of kinesthetic and spatial-visual abilities in perceptual-motor learning. Journal of Experimental Psychology, 1963, 66, 6-11.

Flowers, K. Handedness and controlled movement. British Journal of Psychology, 1975, 66, 39-52.

Forsyth, G. A., & Huber, R. J. Selective attention in ambiguous figure perception: An individual difference analysis. Bulletin of the Psychonomic Society, 1976, 7, 498-500.

Forsyth, G. A., & Shor, R. E. An individual difference analysis of interdimensional additivity in multidimensional scaling. Perception and Psychophysics, 1974, 15, 563-567.

Fournier, B. A., & Stager, P. Concurrent validation of a dual-task selection test. Journal of Applied Psychology, 1976, 61, 589-595.

Gabriel, R. F., & Burrows, A. A. Improving time-sharing performance of pilots through training. Human Factors, 1968, 10, 33-40.

Gerathewohl, S. J. Psychophysiological effects of aging--developing a functional age index for pilots: II. Taxonomy of psychological factors. Technical Report FAA-AM-78-16. Washington, D.C.: FAA Office of Aviation Medicine, March 1978.

Gerlach, V. S. Cues, feedback and transfer in undergraduate pilot training. Phase III. Tempe, AZ: Arizona State University Department of Education Technology, October 1974.

Gerlach, V. S. Cues, feedback, and transfer in undergraduate pilot training. AFOSR-TR-76-0056. Tempe, AZ: Department of Education Technology, August 1975.

Gilbert, C. Non-verbal perceptual abilities in relation to left-handedness and cerebral lateralization. Neuropsychologia, 1977, 15, 779-791.

Gilson, C. R. Individual differences in risk-taking. Technical Report 13. New Haven, CT: Yale University, June 1968.

Goodenough, D. R. The role of individual differences in field dependence as a factor in learning and memory. Psychological Bulletin, 1976, 83, 675-694.

Gopher, D., & Kahneman, D. Individual differences in attention and ther prediction of flight criterion. Perceptual and Motor Skills, 1971, 33, 1335-1342.

Griffith, D., & Johnston, W. A. An informatin processing analysis of visual imagery. Journal of Experimental Psychology, 1973, 100, 141-146.

Girard, Y., & Requin, J. Effects of a preparatory period on RT followed by a pointing movement either guided or not. Perceptual and Motor Skills, 1973, 37, 980-982.

Guilford, J. P., & Lacey, J. I.l Printed classification tests. (AAF Aviation Psychology Program Research Report No. 5). Washington, D.C.: U.S. Government Printing Office, 1947.

Guion, N., Vitola, B. M., & Leisey, S. A. Background and interest measures as predictors of success in undergraduate pilot training. AFHRL-TR-76-9. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, May 1976.

Hawkins, H. L., Rodriguez, E., & Reicher, G. M. Is time-sharing a general ability? Technical Report No. 3. University of Oregon: Office of Naval Research, 1979.

Hick, W. E. On the rate of gain of information. Quarterly Journal of Experimental Psychology, 1952, 4, 11-26.

Hill, J. W., Gardiner, K. W., & Bliss, J. C. Design study of a tactile cuing system for pilot training. AFHRL-TR-69-12, AD-697 991. Wright-Patterson AFB, OH: Training Research Division, Air Force Human Resources Laboratory, August 1969.

Hitch, G. J. The role of short-term working memory in mental arithmetic. Cognitive Psychology, 1978, 10, 302-323.

Hogaboam, T. W., & Pellegrino, J. W. Hunting for individual differences in cognitive processes: Verbal ability and semantic processing of pictures and words. Memory and Cognition, 1978, 6, 189-193.

Holtzman, P. D. Cognitive attitudes of leveling and sharpening in time-error assimilation tendencies. Doctoral Dissertation, University of Kansas, 1952.

Howell, W. C. Uncertainty from internal and external sources: A clear case of overconfidence. Journal of Experimental Psychology, 89, 240-243.

Howell, W. C. Compounding unceartainty from internal sources. Journal of Experimental Psychology, 1972, 95, 6-13.

Hunt, E., & Lansman, M. Cognitive theory applied to individual differences. In W. K. Estes (Ed.), Handbook of learning and cognitive processes (Vol. 1). Hillsdale, NJ: Lawrence Erlbaum Associates, 1976.

Hunt, E., Lunneborg, D., & Lewis, J. What does it mean to be high verbal? Cognitive Psychology, 1975, 7, 194-227.

Hunter, D. R. Development of an enlisted psychomotor/perceptual test battery. AFHRL-TR-75-60, AD-A020 544. Lackland Air Force Base, TX: Air Force Human Resources Laboratory, Personnel Research Division, November 1975.

Hunter, D. R., Maurelli, V. A., & Thompson, N. A. Validation of a psychomotor/perceptual test battery. AFHRL-TR-77-28, AD-A044 525. Lackland Air Force Base, TX: Air Force Human Resources Laboratory, Personnel Research Division, July 1977.

Huttenlocher, J., & Burke, D. Why does memory span increase with age? Cognitive Psychology, 1976, 8, 1-31.

Isaac, P. D. Dissimilarities as indices of individual perceptual structure. Perception and Psychophysics, 1970, 7, 229-233.

Jensen, A. R. Scoring the Stroop Test. Acta Psychologica, 1965, 24, 398-408.

Jensen, A. R. Individual differences in visual and auditory memory. Journal of Educational Psychology, 1971, 62, 123-131.

Jensen, R. S., & Benel, R. A. Judgment evaluation and instruction in civil pilot training. Illinois University at Urbana-Champaign: Savoy Aviation Research Laboratory, December 1977.

Johnson, E. M., Cavanaugh, R. C., Spooner, R. L., & Samet, M. G. Utilization of reliability measurements in Bayesian inference: Models and human performance. IEEE Transactions on Reliability, 1973, r-22, 176-183.

Juola, J. F., & Atkinson, R. C. Memory scanning for words versus categories. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 522-527.

Just, M. A., & Carpenter, P. A. Eye fixations and cognitive processes. Cognitive Psychology, 1976, 8, 441-480.

Kagan, J. Impulsive and reflective children: Significance of conceptual tempo. In J. D. Krumboltz (Ed.), Learning and the educational process. Chicago: Rand-McNally, 1965.

Kagan, J. Matching familiar figures test, 1969.

Kahneman, D,, & Tversky, A. On the psychology of prediction. Psychological Review, 1973, 80, 237-251.

Karlins, M., Lee, R. E. III, & Shroder, H. M. Creativity and information search in a problem-solving context. Psychometric Science, 1967, 8, 165-166.

Karlovich, R. S., & Graham, J. T. Auditorily paced keytapping performance during synchronous, delayed, and decreased visual feedback. Perceptual and Motor Skills, 1968, 26, 731-743.

Kay, B. R. Intraindividual differences in sensory channel preference. Journal of Applied Psychology, 1958, 42, 166-167.

Keele, S. W. Movement control in skilled motor performance. Psychological Bulletin, 1968, 70, 387-403.

Kennedy, R. S., & Bittner, A. C., Jr. Progress in the analysis of a performance evaluation test for environmental research (PETER). Report No. TM-78-02. New Orleans, Louisiana: Naval Aerospace Medical Research Lab, October 1978.

Kimura, D. Cerebral dominance and the perception of verbal stimuli. Canadian Journal of Psychology, 1961, 15, 166-171.

Kimura, D. Left-right differences in the perception of melodies. Quarterly Journal of Experimental Psychology, 1964, 16, 355-358.

Kirby, J. R., & Das, J. P. Information processing and human abilities. Journal of Educational Psychology, 1978, 70, 58-66.

Klee, H., & Eysenck, M. W. Comprehension of abstract and concrete sentences. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 522-529.

Kogan, N., & Wallach, M. A. Risk-taking as a function of the situation, the person, and the group. In G. Mandler, P. Mussen, N. Kogan, & M. A. Wallach (Eds.), New directions in Psychology III. New York: Holt, Rinehart, and Winston, 1967.

Koslowski, L. T., & Bryant, K. J. Sense of direction, spatial orientation, and cognitive maps. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 590-598.

Kvalseth, T. O. Comparison between information rates generated by rotary hand and arm movements. Perceptual and Motor Skills, 1976, 42, 1115-1118. (a)

Kvalseth, T. O. Distribution of movement time in a target-aiming task. Perceptual and Motor Skills, 1976, 43, 507-513. (b)

Lackner, J. R. Proprioceptive facilitiation of open-loop visual pointing. Perceptual and Motor Skills, 1974, 39, 263-265.

Laszlo, J. I. Training of fast tapping with reduction of kinesthetic, tactile, visual, and auditory sensations. Quarterly Journal of Experimental Psychology, 1967, 19, 344-349.

Laszlo, J. I., & Manning, L. C. The role of motor programming, command, and standard in the central control of skilled movement. Journal of Motor Behavior, 1970, 2, 111-124.

Lee, W. Decision theory and human behavior. New York: Wiley, 1971.

LeMaster, W. D., & Gray, T. H. Ground-training devices in job-sample approach to UPT selection and screening. AFHRL-TR-74-86 AD-A009 995. Williamns Air Force Base, AZ: Air Force Human Resources Laboratory, Flying Training Division, December 1974.

Leshowitz, B., Parkinson, S. P., & Waag, W. L. Visual and auditory information processing in flying skill acquisition. AFHRL-TR-74-103, AD-A009 636. Williams Air Force Base, AZ: Flying Training Division, Air Force Human Resources Laboratory, December 1974.

Levine, J. M., & Samet, M. G. Information seeking with multiple sources of confusing and unreliable information. Human Factors, 1973, 15(4), 407-419.

Levy, J. Lateral specialization of the human brain: Behavioral manifestations and a possible evolutionary basis. In J. A. Kiger, Jr. (Ed.), The biology of behavior. Corvallis, Oregon: Oregon State University Press, 1973.

Levy-Agresti, J., & Sperry, R. W. Differential perceptual capacities of major and minor hemispheres. Proceedings of the National Academy of Sciences, 1968, 61, 1151.

Lewis, C. H., & Griffin, M. J. The effects of vibration on manual control performance. Egronomics, 1976, 19, 203-216.

Lichtenstein, S., Earle, T. C., & Slovic, P. Cue utilization in a numerical prediction task. Journal of Experimental Psychology: Human Learning and Memory, 1975, 104, 77-85.

Lichtenstein, S.,, & Fischhoff, B. Training for calibration (ARI Technical Report TR-78-A32). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, November 1978.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. Calibration of probabilities: The state of the art. In H. Jugermann and G. de Zeeuw (Eds.), Decision making and change in human affairs. Dordrecht, Holland: D. Reidel, 1977.

Lichtenstein, S., & Slovic, P. Reversals of preferences between bids and choices in gambling decisions. Journal of Experimental Psychology, 1971, 89, 46-55.

Lichtenstein, S., Slovic, P., & Zink, D. Effect of instruction in expected value on optimality of gambling decisions. Journal of Experimental Psychology, 1969, 29, 236-240.

Lowenfeld, V. The nature of creative activity: Experimental and comparative studies of visual and non-visual sources. London: Harcourt, 1939.

Lyon, D. R. Individuial differences in immediate serial recall: A matter of mnemonics? Cognitive Psychology, 1977, 9, 403-411.

Lyon, D., & Slovic, P. Dominance of accuracy information and neglect of base rates in probability estimation. Acta Psychologica, 1976, 40, 287-298.

Mandler, G. Organization in memory. In K. W. Spence, & J. T. Spence (Eds.), The psychology of learning and motivation Volume 1. New York: Academic Press, 1967.

Mandler, G., Worden, P. E., & Graesser, A. C. II. Subjective disorganization: Search for the locus of list organization. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 220-235.

Marks, D. F. Individual differences in the vividness of visual imagery and their effect on function. In P. W. Sheehan (Ed.), The function and nature of imagery. New York: Academic Press, 1972.

Marteniuk, R. G., & Roy, E. A. The codability of kinesthetic location and distance information. Acta Psychologica, 1972, 36, 471-479.

Martin, E. Short-term memory, individual differences, and shift performance in concept formation. Journal of Experimental Psychology, 1968, 76, 514-520.

Martin, M. Memory span as a measure of individual differences in memory capacity. Memory and Cognition, 1978, 6, 194-198.

Matheny, W. G., Gray, T. H., & Waters, B. D. AFHRL/FT capacities in undergraduate pilot training simulation research: Executive Summary. AFHRL-TR-75-26(I), AD-A017 168. Williams Air Force Base, AZ: Flying Training Division, Air Foce Human Resources Laboratory, August 1975.

Matheny, W. G., Lowes, A. L., Baker, G., & Bynum, J. N. An investigation of visual, aural, motion, and control movement cues. Technical Report NavTrnDevCen 69-C-0304-1. Fort Worth, TX: Life Sciences Inc., April 1971.

Mavrides, C. Selective attention and individual differences in judgmental responses to multifeature patterns. Psychonomic Science, 1970, 21, 67-68.

McAllister, D. W., Mitchell, T. R., & Beach, L. R. The contingency model for the selection of decision strategies: An empirical test. Technical Report 78-17. Seattle, Washington: University of Washington, October 1978.

McCaughan, L. R. Stability/instability and change of expectancy, a test for cognitive determinants of psychomotor performance. Perceptual and Motor Skills, 1978, 46, 219-225.

McGee, M. G. Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. Psychological Bulletin, 1979, 86, 889-915.

McGrevy, D. F., & Valentine, L. D., Jr. Validation of two aircrew psychomotor tests. AFHRL-TR-74-4, AD-777 830. Lackland Air Force Base, TX: Air Force Human Resources Laboratory, Personnel Research Division, January 1974.

McLaurin, W. A. Validation of a battery of performance tests for prediction of aerospace ground equipment course grades. AFHRL-TR-73-20, AD-774 586. Lackland Air Force Base, Texas: Air Force Human Resources Laboratory, November 1973.

McLeod, C. M., Hunt, E. G., & Mathews, N. M. Individual differences in the verification of sentence-picture relations. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 493-507.

Meisel, P., & Wapner, S. Interaction of factors affecting space localization. Journal of Experimental Psychology, 1969, 79, 430-437.

Melton, A. W. (Ed.). Apparatus tests. Air Force Aviation Psychology Program Research Report No. 4. Washington, D. C.: U.S. Government Printing Office, 1947.

Mendelsohn, G. A., & Lindholm, E. P. Individual differences and the role of attention in the use of cues in verbal problem-solving. Journal of Personality, 1972, 40, 226-241.

Mensh, I. N. Pilot selection by phychological methods. Report No. ONRL-R-26-70. London, England: Office of Naval Research, May 1970.

Meyer, D. E., Schaneveldt, R. W., & Ruddy, M. G. Functions of graphemic and phonemic codes in visual word-recognition. Memory and Cognition, 1974, 2, 309-321.

Meyer, R. P., Laveson, J. I., Weissman, N. S., & Eddowes, E. E. Behavioral taxonomy of undergraduate pilot training tasks and skills: Surface task analysis, taxonomy structure, classification rules, and validation plan. AFHRL-TR-74-33(II), AD-A008 771. Williams Air Force Base, AZ: Flying Training Division, Air Force Human Resources Laboratory, December 1974.

Miller, A. C., III, Rice, T. R., & Metcalfe, M. R. An analytic characterization of Navy command and control decisions. Menlo Park, CA: Applied Decision Analysis, March 1979.

Moore, J., & Massaro, D. Attention and processing capacity in auditory recognition. Journal of Experimental Psychology, 1973, 99, 49-54.

Moroney, W. F., Pruitt, R., & Lau, C. Utilization of energy maneuverability data in improving in-flight performance in air combat maneuvering. Proceedings of the Human Factors Society, 1979, 23, 503-507.

Mowbray, G. H., & Rhoades, M. V. On the reduction of choice reaction times with practice. Quarterly Journal of Experimental Psychology, 1959, 11, 16-23.

Navon, D., & Gopher, D. On the economy of the human-processing system. Psychological Review, 1979, 86, 214-255.

Neisser, U. Visual search. Scientific American, 1964, 210, 94-102.

Neisser, U. Cognitive Psychology. New York: Appleton-Century-Crofts, 1967.

Newell, A., & Simon, H. Human problem solving. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

Nickerson, R. S., & Feehrer, C. E. Decision making and training: A review of theoretical and empirical studies of decision making and their implications for the training of decision makers. Technical Report NAVTRAEQUIPCEN 73-C-0128. Cambridge, MA: Bolt, Beranek, and Newman, July 1975.

North, R. A., & Gopher, D. Measures of attention as predictors of flight performance. Human Factors, 1976, 18, 1-14.

O'Connor, M. F., Peterson, C. R., & Palmer, T. J. Stakes and probabilities in information purchase., Organizational Behavior and Human Performance, 1972, 7, 43-52.

Over, R. Individual differences in figural aftereffects. Psychological Bulletin, 1970, 74, 405-410.

Paivio, A. Imagery and verbal processes. New York: Holt, Rinehart, and Winston, 1971.

Parker, J. F., & Fleishman, E. A. Ability factors and component performance measures as predictors of complex tracking behavior. Psychological Monographs, 1960, 74, No. 503.

Passey, G. E., & McLaurin, W. A. Perceptual-psychomotor tests in aircrew selection: Historical review and advanced concepts. PRL-TR-66-4, AD-636-606. Lackland Air Force Base, TX: Personnel Research Laboratory, June 1966.

Payne, J. W. Alternative approaches to decision-making under risk: Moments versus risk dimensions. Psychological Bulletin, 1973, 80, 439-453.

Payne, J. W. Task complexity and contingent processing in decision making: An information search and protocol analysis. Organizational Behavior and Human Performance, 1976, 16, 366-387.

Peterson, C. R., Ulehla, Z. J., Miller, A. J., Bourne, L. E., & Stilson, D. W. Internal consistency of subjective probabilities. Journal of Experimental Psychology, 1965, 70, 526-533.

Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. Journal of Experimental Psychology, 1959, 58, 193-198.

Pew, R. W., & Adams, M. J. Development of procedures for assignment of subjects in ASUPT experiments: Background and data collection plan. Report No. 3204. Cambridge, MA: Bolt, Beranek, and Newman, Inc., November 1975.

Pew, R. W., Rollins, A. M., Adams, M. J., & Gray, T. H. Development of a test battery for selection of subjects for ASPT experiments. Report No. 3585. Cambridge, MA: Bolt, Beranek, and Newman, Inc., November 1977.

Pitz, G. F., & Geller, E. S. Revision of opinion and decision tasks in an information-seeking task. Journal of Experimental Psychology, 1970, 83, 400-405.

Posner, M. I. Abstraction and the process of recognition. In G. H. Bower and J. T. Spence, (Eds.). The psychology of learning and motivation: Advances in theory and research. New York: Academic Press, 1969.

Posner, M. I., & Keele, S. W. Decay of visual information from a single letter. Science, 1967, 158, 137-139.

Posner, M. I., & Mitchell, R. F. Chronometric analysis of classification. Psychological Review, 1967, 74, 392-409.

Prophet, W. W. Long-term retention of flying skills: A review of the literature. Technical Report FR-ED(p)-76-35. Alexandria, VA: Human Resources Research Organization, October 1976.

Pruitt, D. G. Choice shifts in group discussion: An introductory review. Journal of personality and Social Psychology, 1971, 20, 339-360. (a)

Pruitt, D. G. Conclusions: Toward an understanding of choice shifts in group discussion. Journal of Personality and Social Psychology, 1971, 20, 495-510. (b)

Ragan, T. J., Back, K. T., Stansell, V., Ausburn, L. J., Ausburn, F. B., Butler, P. A., Huckabay, K., & Burkett, J. R. Cognitive styles: A review of the literature. AFHRL-TR-78-90(1), AD-A069 435. Lowry Air Force Base, CO, Technical Training Division, Air Force Human Resouces Laboratory, May 1979.

Rapport, A., & Wallsten, T. S. Individual decision making. Annual Review of Psychology, 1972, 23, 131-176.

Rigal, R. A. Hand efficiency and right-left discrimination. Perceptual and Motor Skills, 1974, 38, 219-224.

Roby, T. B. Small group performance. Chicago: Rand-McNally, 1968.

Rose, A. M. Human information processing: An assessment and research battery. Technical Report No. 46. Ann Arbor, Michigan: University of Michigan Human Performance Center, January 1974.

Samet, A. M., & Fernandes, K. An information processing approach to performance assessment: I. Experimental investigation of an information processing performance battery. AIR-58900-TR. Washington, D. C.: American Institutes for Research, November 1977.

Samet, M. G. Quantitative interpretation of two qualitative scales used to rate military intelligence. Human Factors, 1975, 17, 192-202.

Schaller, M. J., & Oziadosz, G. M. Individual differences in adult foveal visual asymmetries. Journal of Experimental Psychology: Human Perception and Performance, 1975, 1, 353-365.

Schrenk, L. P. Aiding the decision maker -- a decision process model. Ergonomics, 1969, 12, 543-557.

Seamon, J. G. Imagery codes and human information retrieval. Journal of Experimental Psychology, 1972, 96, 468-470.

Semjen, A., & Requin, J. Movement amplitude, pointing accuracy, and choice reaction time. Perceptual and Motor Skills, 1976, 43, 807-812.

Shaklee, H., & Fischhoff, B. Discounting in multicausal attribution: The principle of minimal causation. Technical Report PTR-1060-78-8. Eugene, OR: Perceptronics, August 1978.

Sheehan, P. W. Functional similarity of imaging to perceiving: Individual differences in vividness of imagery. Perceptual and Motor Skills, 1966, 23, (Monograph Supplement 6-V23), 1011-1033. (a)

Sheehan, P. W. Accuracy and vividness of visual images. Perceptual and Motor Skills, 1966, 23, 391-398. (b)

Shepard, R. N. The mental image. American Psychologist, 1978, 33, 125-137.

Shepard, R. N., & Metzler, J. Mental rotation of three-dimensional objects. Science, 1971, 171, 701-703.

Shepard, R. N., & Teghtsoonian, M. Retention of information under conditions approaching a steady state. Journal of Experimental Psychology, 1961, 62, 302-309.

Sinacori, J. B. Piloted aircraft simulation concepts and overview. (STI-TR-1074-2). Mountain View, California: Systems Technology Inc., March 1978.

Slovic, P. Risk-taking in children: Age and sex differences. Child Development, 1966, 37, 169-176.

Slovic, P. Differential effects of real versus hypothetical payoffs upon choices among gambles. Journal of Experimental Psychology, 1969, 79, 139-145.

Slovic, P. Towards an understanding and improving decision. In Science, Technology, and the modern Navy, Thirtieth anniversary, 1946-1976. Arlington, Virginia: Office of Naval Research, 1976.

Slovic, P., Fischhoff, B., & Lichtenstein, S. Behavioral decision theory. Annual Review of Psychology, 1977, 28, 1-39.

Slovic, P., & Lichtenstein, S. The relative importance of probabilities and payoffs in risk-taking. Journal of Experimental Psychology Monograph Supplement, 1968, 78, No. 3, Part 2.

Slovic, P., & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Smode, A. F., Hall, E. R., & Meyers, D. E. An assessment of research relevant to pilot training. AMRL-TR-66-196. Arlington, VA: Biotechnology, Inc., November 1966.

Sperling, G. The information available in brief visual presentations. Psychological Monographs, 74, No. 498.

Sternberg, R. J. Component processes in analogical reasoning. Psychological Review, 1977, 84, 353-378.

Sternberg, R. J. Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

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

Sternberg, S. Two operations in character recognition: Some evidence from reaction-time measurements. Perception and Psychophysics, 1967, 2, 45-53.

Sternberg, S. Memory scanning: New findings and current controversies. Quarterly Journal of Experimental Psychology, 1975, 27, 1-32.

Sterner, R. T., & Carpp, L. Psychomotor rehearsal: Enhancement of rotary pursuit tracking using a massed-training procedure. Perceptual and Motor Skills, 1977, 44, 243-248.

Stewart, W. A., & Wainstein, E. S. Rand symposium on pilot training and the pilot career. Santa Monica, CA: Rand Corporation, December 1970.

Stoner, J. A. F. Risky and cautious shifts in group decisions: The influence of widely held values. Journal of Experimental Social Psychology, 1968, 4, 442-459.

Stroop, J. R. Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 1935, 18, 643-662.

Svenson, O. Process descriptions in decision making. Organizational Behavior and Human Performance, 1979, 23, 86-112.

Syllabus of instruction for undergraduate pilot training (T-37/T-38). Randolph Air Force Base, TX: Air Training Command, January 1979.

Tulving, E. Subjective organization in free recall of "unrelated" words. Psychological Review, 1962, 69, 344-354.

Tulving, E., & Osler, S. Transfer effects in whole/part free recall learning. Canadian Journal of Psychology, 1967, 21, 253-262.

Tversky, A. Intransitivity of preferences. Psychological Review, 1969, 76, 31-48.

Tversky, A., & Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.

Tversky, B. Encoding processes in recognition and recall. Cognitive Psychology, 1973, 5, 275-287.

Underwood, B. T., Boruch, R. F., & Malmi, R. A. The composition of episodic memory. Evanston, IL: Northwestern University, 1977.

Walker, N. K., & Walker, M. M. Dual-task performance (ZITA/ADT) on ANG and USAF pilots and a comprison with results of cockpit simulator tests and with dive bombing and strafing results. Rockville, MD: Normal K. Walker Associates, May 1979.

Wallach, H., & Smith, A. Visual and proprioceptive adaptation to altered oculo-motor adjustments. Perception and Psychophysics, 1972, 11, 413-416.

Wailsten, T. S. Conjoint-measurement framework for the study of probabilistic information processing. Psychological Review, 1972, 79, 245-260.

Ware, J.R., & Barnhill, R.C. Effects of field articulation and feedback on perception of immediate visual-kinesthetic position. Perceptual and Motor Skills, 1975, 40, 875-878.

Weitzman, D. O. Individual differences in geographic orientation and spatial memory representation. Paper presented at the 87th Annual Convention of the American Psycnological Association, New York, 1979.

Welford, A. T. The "psychological refractory period" and the timing of high-speed performance - A review and a theory. British Journal of Psychology, 1952, 43, 2-19.

Whimbey, A. E., & Leiblum, S. L. Individual differences in memory span with and without activity intervening between presentation and recall. Journal of Educational Psychology, 1967, 58, 311-314.

White, M. J. Laterality differences in perception: A review. Psychological Bulletin, 1969, 72, 387-405.

Whitely, S. E. Information-processing on intelligence test items: Some response components. Applied Psychological Measurement, 1977, 1, 465-476.

Wilkinson, D. A. Visual-motor control loop: A linear system? Journal of Experimental Psychology, 1971, 89, 250-257.

Witkin, H. A. Perception of body position and the position of the visual field. Psychological Monographs, 1949, 63, 1-46.

Wood, M.E., & Gerlach, V.S. Transfer from audiovisual pretraining to a continuous perceptual motor task. AFHRL-TR-74-8, AD-778 078. Williams Air Force Base, AZ, Flying Training Division, Air Force Human Resources Laboratory, March 1974.

Wright, P. The harassed decision maker: Time pressures, distractions, and the use of evidence. Journal of Applied Psychology, 1974, 59(5), 555-561.

APPENDIX A: CONCEPTUAL FRAMEWORK

This appendix considers a variety of approaches to the specification of the abilities and processes relevant to flying. In categorizing piloting tasks, we have maintained the distinction between abilities and processes as discussed in the introduction. The term ability tends to be used in reference to perceptual-motor tasks, whereas the complex cognitive tasks are discussed in terms of processes or operations without attempting to specify underlying abilities.

A number of potential sources are available from which the abilities and processes crucial to piloting may be inferred. One possibility consists of analysis of the behavior of successful pilots, as well as the complementary analysis of unsuccessful pilot behavior. A related approach would involve the analysis of the Syllabus for Undergraduate Pilot Training, along with analyzing the behaviors that lead to success or failure in pilot training. A third approach might consider an analysis of specific piloting tasks to allow the inference of abilities and processes crucial to the performance of that maneuver. All of these approaches have proven useful, and by correlating and interpreting the results of these approaches, we have been able to construct a framework characterizing pilot performance in terms of underlying abilities and processes. The tasks identified in the literature review may then be related to those abilities and processes to provide a conceptual framework within which a battery of those tasks may be identified as a pilot-selection device.

Gerathewohl (1978) reviewed the literature analyzing the causes of successful and unsuccessful pilot behavior. He combined the results of a variety of studies by subsuming similar factors across studies under the same name and by inferring requisite factors that were not explicitly stated. Using this approach, he specified nine factors which are common to most studies of successful pilot behaviors. Those factors included perception, reaction time and response, vigilance and attention, sensorimotor abilities and skills, motor activities, learning, cognition, personality dependent behavior, and social behavior. When the factors relevant to pilot error were considered as well, Gerathewohl (1978) was able to specify a set of 14 major "psychological factors" which appeared to be characteristics of an essential to pilot performance.

This analysis considered the job elements of major piloting tasks and the psychological factors associated with those job elements. For example, mission and flight planning require planning and anticipation, comprehension of input, and application of previous knowledge, implicating the psychological factors of perception, attention, cognition, and learning. The cruise phase of flight requires coordination of activities, closed-loop manual tracking, decision making, systems monitoring, and computation, implicating additional factors of sensorimotor skill, social relations, personal

adjustment, and stamina. This approach thus allows specific factors to be tied directly to major piloting tasks and subtasks.

The 14 psychological factors specified by Gerathewohl are useful, but have two major limitations for the identification of a test battery. First of all, the factors are much more general than the abilities and processes dealt with here, often subsuming several quite different processes within a single factor. Secondly, many of the factors specified are personality and social factors, and are thus outside the scope of this project. The 14 factors are described below.

1. Perception. This factor includes sensing and perceiving visual, auditory, tactual, and other stimuli, signals, and information as well as the observation, detection, and visualization processes.

2. Attention. This factor includes alertness, vigilance, watchkeeping, span, channel capacity, and time-sharing functions.

3. Reaction. This factor includes reaction time and discrete, serial, and multiple task responses.

4. Orientation. This factor includes bodily, spatial, and geographic orientation.

5. Sensorimotor. This factor includes eye-hand coordination, finger dexterity, speed and accuracy of muscular activities, tracking, and precise multiple control.

6. Stamina. This factor includes body strength, physical and emotional endurance, acceleration tolerance, work capacity, resourcefulness, and stress and fatigue tolerance.

7. Cognition/Mentation. This factor includes acquisition and processing of information, thinking, concept formation, deductive and inductive reasoning, finding and establishing of relations, judgment, foresight, planning, and problem solving.

8. Experience. This factor includes memory, conditioning, habit formation, situational and personal adjustment, management, and procedural actions.

9. Interpersonal Relations. This factor includes communication, working with others, accepting personal and organizational responsibility, supervision, living and working with others, and crew coordination.

10. Personality. This factor includes self-confidence, self-sufficiency, self-discipline, calmness, composure, risk-taking, thoroughness, attitudes, leadership, and morale.

11. Learning. This factor includes memory functions (both short- and long-term), remembering written and verbal material, objects, courses of action and relationships; as well as acquiring information from various sources and following procedures based on acquired and learned information.

12. Decision-Making. This factor consists of selecting and formulating from a variety of possibilities or a limited number of alternatives a course of action with the intent of executing it. Hence, this factor can be considered independent of cognition/mentation, since decisions can be made for other than logical reasons and contain an intent component beyond the reasoning and judgment state.

13. Mechanical Aptitude. This factor includes mechanical comprehension, handling tools and equipment, visualization of mechanical relations, detecting and locating malfunctions in instruments, and fabricating, assembling, and repairing (faulty) equipment.

14. Flight Motivation. This factor includes the intention to become a pilot, to fly and be active in aviation, to overcome difficulties, hardships, and risks involved in flying, and to succeed as an aviator under all circumstances ("keep my license").

The results obtained by Gerathewohl (1978) using the correlative summary approach are supported by an examination of the Syllabus for Undergraduate Pilot Training and the Record of Training employed by instructor pilots. For example, the first instructional unit (B4501) in the T-51 instrument flight simulator contains practice items including crosscheck, airspeed control, altitude control, heading control, straight-and-level, turns to headings, changes of airspeed straight and turning, constant airspeed climb and descent, rate climb and descent, military power climb, enroute descent, level-off, use of trim, steep turn 45 and 60 degrees, and unusual altitude recoveries. Even this single unit requires many of the piloting subtasks of various phases of flight given by Gerathewohl. In addition, the Record of Training form indicates that student pilots are graded not only on performance of specific maneuvers, but also on judgment, emergency procedures, and ATC voice procedures. There seems to be little doubt that the perceptual-motor abilities and cognitive processes tied to piloting success by Gerathewohl are also implicated in the successful completion of UPT.

Another approach to determining the abilities and processes crucial to flying is to analyze specific flying tasks and construct a taxonomy of behavioral skills required to perform these tasks. Meyer, Leveson, Weissman, and Eddowes (1974) constructed such a taxonomy for tasks required in UPT. They began with a simple model involving Cues from both external and system sources which, when detected and interpreted accurately led to Mental Actions. The Mental Actions led to Motor Actions upon the system, as well as further Mental Action comparing

the Motor Action and resulting Cue changes to some performance standard. They then performed surface task analyses in terms of these Cues, Mental Actions, and Motor Actions. This analysis was very specific: Cues were specified in terms of kind (visual, aural, motion, and control cues), complexity (number of cue input sources), and total inputs. Mental Actions were classified according to complexity (number of cue inputs and number and coordination of required responses), type of information processing required (specific cue, memory recall, multicue, or iterative processing), and type of decision processing required (simple judgment based on fact or complex based on estimation). Motor Actions were similarly classified according to their continuity, control output, and complexity. Within this system, then, a behavioral skill could be defined in terms of a Cue, Mental Action, and Motor Action.

It seems likely that a taxonomy such as that of Meyer et al. (1974) would be extremely useful for identifying specific behavioral skills as topics for training. However, for the purpose of building a selection battery, the taxonomy is both too specific (so that Motor Actions are specified in terms of specific output controls rather than underlying abilities) and too vague (so that the specific cognitive and information processing operations involved are difficult to determine). The relevance of certain psychomotor and perceptual-motor abilities to Cue pickup and Motor Action as defined by the taxonomy seem apparent, and is supported by the history of validity for predicting pilot success shown by tests of these abilities. Similarly, it is easy to reconceptualize multi-cue processing of visual and aural inputs in terms of encoding, selective attention, time-sharing, comparison, memory search, and decision-making (and probably other processes). The model's emphasis on sensory cues, information-processing, and perceptually guided motor response as the key elements of flying coincide with the emphasis in the development of this test battery, even if the level of analysis does differ.

In summary, all three approaches to specifying the abilities and processes requisite for flying agree in their emphasis on perceptual-motor and cognitive processes, whether analyzing a specific flight maneuver or the series of tasks required for successful mission completion. While the level at which performance has been analyzed varies from "major psychological factors" to "specific behavior skills," a careful examination of these lists provides support for the presumption that the list of abilities and processes described below is in fact explicitly related to successful piloting. Moreover, these abilities and processes have the advantage of support and empirical research within the fields of perceptual-motor and cognitive psychology.

In nearly every phase of flight, the pilot must monitor information sources from a number of modalities, including visual, aural, and tactile and motion cues. Moreover, especially visual cues arise from different sources, both in terms of location in the external environment and in terms of the maze of dials, meters, and other information sources within the cockpit. It is important for a pilot to perceive these sources quickly (perceptual speed), for changes in information about direction or altitude, or emergency warnings must be perceived and reacted to (reaction time) quickly. Since there are multiple input sources, the pilot must be able to focus on the crucial one (selective attention) and resist inference from extraneous inputs. Moreover, since often more than one input or action will be crucial at the same time, the pilot must be able to divide attention between different inputs and activities (time-sharing). Since flights are extended in time, the pilot's monitoring activities must also be maintained. A variety of sensorimotor sensitivities and facilities (kinesthetic sensitivity; spatial orientation) contribute to the ease of information pickup.

These information inputs lead to responses designed to maintain flight control, carry out the flight mission, and deal with emergency situations. The response is contingent upon which of the possible inputs occurs, and the speed with which a pilot can decide which response is appropriate (decision speed) may be crucial. Moreover, since multiple responses are required often in rapid succession, sheer motor movement speed (speed of arm movement) is important. In addition to speed, however, accuracy of the movement response is essential. The task of steering an airplane requires the pilot to make fine, controlled adjustments of the controls (control precision) to alter heading, altitude, speed, etc. Moreover, the pilot must often coordinate both hands or a hand and foot (multilimb coordination) in making these adjustments. Furthermore, steering the aircraft requires the pilot to make anticipatory movements timed to coincide with predicted future locations (rate control; also the sensorimotor sensitivities are helpful here).

In many cases, the pilot's task is more complex than merely perceiving the appropriate input quickly and making a relatively automatic response quickly and accurately. In those cases where the response is not so automatic, the pilot must search the input sources for additional information (visual scanning) and encode the inputs for further information processing (encoding). For example, an altimeter reading of itself may not indicate a specific response, so the pilot must search memory for the appropriate reading at that stage of flight (memory search), perhaps form an image of the appropriate reading (visualization), and compare the image or mental standard to the actual reading (comparison). The pilot may also seek additional information from the crew or ground sources; and must understand and integrate their inputs (verbal comprehension) with previously acquired information. The pilot might also draw on knowledge of the mission and the aircraft, and integrate that information (retrieval) as well as mapping relationships from previous knowledge onto the present situation (problem solving). The variety of alternatives and

contingencies possible point to the crucial role of decision making in normal flight performance.

When an unexpected situation or an emergency arises, the burden on the pilot becomes even greater. First of all, the pilot must detect the emergency. The speed in doing so is determined by input scanning efficiency and attentional processes, as well as the overall amount of information with which the pilot can deal (processing capacity). Visualization functions may be especially helpful in estimating the nature and location of the emergency. For example, if a warning light indicated an overheating engine, once the pilot became cognizant of the problem, the location of the engine could be visualized and the probable cause and degree of danger estimated. In order to evaluate alternative courses of action, the pilot must then seek further information and estimate the likelihood of alternative explanations for the problem based on the new information. The pilot must also estimate the probability of success of alternative courses of action and decide what probability of success is required for the mission to continue. Dimensions of cognitive and perceptual style, which deal with preferences and facilities with a combination of the relevant abilities and processes, are also important to pilot success.

Table A1 presents an outline of the relationship between piloting subtasks and abilities and processes required for successful performance of the tasks.

Table A1. A Representation of Flight Tasks of Subtasks and the Abilities or Processes Required for Performance

Flight Task/Subtask	Ability/Process
Flight planning	
a. seek weather and mission information	Decision Making
b. understand and integrate information	Comprehension
c. estimate success potential of alternative flight plans	Decision Making
d. use previously stored knowledge to aid in evaluation	Memory Retrieval
e. check visual instrument displays	Visualization
f. check movement controls	Kinesthetic Sensitivity
g. compare instrument displays to a stored standard	Comparison
h. search memory for stored standard for instrument display	Memory Search
i. focus on a single instrument reading and avoid interference	Attention
j. time-share between multiple instrument inputs	Attention
Takeoff	
a. perceive changes in instrument readings quickly	Perceptual Speed
b. react to effect control changes in response to instrument reading changes	Movement Speed
c. make small but precise instrument and control adjustments	Control Precision

Table A1 (continued)

Flight Task/Subtask	Ability/Process
Flight Task/Subtask	Ability/Process
d. coordinate hand and foot movements in control adjustments	Multi-Limb Coordination
e. anticipate location and direction of aircraft movement	Spatial Orientation
f. coordinate control and steering movement to anticipated location	Rate Control
Cruise - in addition to maintaining the control functions described under Takeoff, the pilot must:	
a. maintain constant monitoring of instrument and environmental input	Attention
b. identify inputs quickly and accurately	Encoding
c. scan input sources for new or confirming information	Perception
d. allocate attention to multiple inputs	Attention
e. avoid interference from irrelevant inputs	Attention
f. maintain large amounts of information for immediate processing	Memory Capacity
g. diagnose system failures and errors	Visualization
h. map previously learned relations onto current system failure	Problem Solving

Table A1 (continued)

Flight Task/Subtask	Ability/Process
Emergency situations	
a. detect a significant change in information inputs	Perceptual Speed
b. compare cue to standard to diagnose difficulty	Comparison/Visualization
c. seek confirming or additional information	Decision Making
d. evaluate alternative courses of action	Decision Making/Memory Retrieval
e. decide quickly on a course of action and initiate response	Decision Making