

DTIC FILE COPY

1

AD-A190 876

THE REPRESENTATIONAL CODE OF THE INTERNAL MODEL OF DYNAMIC SYSTEMS:
AN INDIVIDUAL DIFFERENCES AND DUAL TASK APPROACH

Christopher D. Wickens and Annette Weingartner
University of Illinois

for

Contracting Officer's Representative
Michael Drillings
Contracting Officer's Representative
Michael Drillings

BASIC RESEARCH LABORATORY
Michael Kaplan, Director



U. S. Army

Research Institute for the Behavioral and Social Sciences

December 1987



Approved for public release; distribution unlimited.

88 1 25 048

U. S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the
Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

WM. DARRYL HENDERSON
COL, IN
Commanding

Research accomplished under contract
for the Department of the Army

University of Illinois

Technical review by

Dan Raqland



| | |
|--------------------|-------------------------------------|
| Accession For | |
| NTIS GRA&I | <input checked="" type="checkbox"/> |
| DTIC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | |
| By | |
| Distribution/ | |
| Availability Codes | |
| Dist | Avail and/or Special |
| A-1 | |

This report, as submitted by the contractor, has been cleared for release to Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or other reference services such as the National Technical Information Service (NTIS). The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|--|-----------------------|---|
| 1. REPORT NUMBER ARI Research Note 87-78 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) The Representational Code of the Internal Model of Dynamic Systems: An Individual Differences and Dual Task Approach | | 5. TYPE OF REPORT & PERIOD COVERED Interim Report September 85 - September 86 |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) Christopher D. Wickens and Annette Weingartner | | 8. CONTRACT OR GRANT NUMBER(s) MDA903-83-K-0255 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology, University of Illinois 603 East Daniel Street Champaign, IL 61820 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q161102B74F |
| 11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600 | | 12. REPORT DATE December 1987 |
| | | 13. NUMBER OF PAGES 37 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) -- | | 15. SECURITY CLASS. (of this report) Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE -- |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) -- | | |
| 18. SUPPLEMENTARY NOTES Michael Drillings, contracting officer's representative | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Internal Model Individual Differences Mental Model Multiple Resources Process Control Dual Task Performance Abilities, | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) - When a human operator monitors and controls complex dynamic processes, it is assumed that an internal representation of the process directs the operator's actions. This internal model is presumed to lie at some point along a verbal-spatial continuum. In order to determine the point on this continuum, nine subjects with high verbal and low spatial abilities, and nine with low verbal and high spatial abilities performed a multi-element failure detection task, either by itself, or concurrently with either a verbal or spatial secondary (OVER) | | |

ARI RESEARCH NOTE 87-7820. Abstract (continued)

memory task. Patterns of interference between the maintaining and updating of the internal model and the performing of the memory tasks were used to infer the mode of internal model employed by the subjects.

Interference results confirm that the failure detection task is spatial, and, as expected, verbal subjects performed better on the verbal secondary task and spatial subjects performed better on the spatial one. Both ability groups demonstrated similar failure detection abilities, and generated similar patterns of dual task interference. These results indicated that all subjects adopted the same strategy for failure detection.

THE REPRESENTATIONAL CODE OF THE INTERNAL MODEL OF DYNAMIC
SYSTEMS: AN INDIVIDUAL DIFFERENCES AND DUAL TASK APPROACH

Christopher D. Wickens and Annette Weingartner

Department of Psychology

University of Illinois at Urbana-Champaign

Champaign, Illinois

U.S.A.

ABSTRACT

When the human operator monitors and controls complex, dynamic processes, it is assumed that an internal representation of the process directs the operator's control actions. This internal model is proposed to lie at some point along a verbal-spatial continuum. In order to determine the point on this continuum, nine subjects with high verbal and low spatial abilities and nine with low verbal and high spatial abilities performed a multi-element failure detection task by itself or concurrently with either a verbal or spatial secondary memory task. Patterns of interference between maintaining and updating the internal model and performing the memory tasks, were used to infer the mode of the internal model adopted by the subjects.

Interference results confirmed that the failure detection task was spatial and as expected, verbal subjects performed better on the verbal secondary task and spatial subjects performed better on the spatial one. Both ability groups demonstrated similar failure detection abilities and generated similar patterns of dual task interference. These results indicated that all subjects adopted the same strategy for failure detection.

INTRODUCTION

A number of researchers have recently invoked the concept of an "internal model" as essential to the task of process control; a task in which operators must monitor and control slowly changing multi-element dynamic systems (Umbers, 1979; Sheridan, 1981; Landeweerd, 1979; Rasmussen, 1979, 1981; Genter & Stevens, 1983). In such a task, the internal model can be expected to serve at least two functions (Wickens, 1984a). It may allow the controller to predict the process response to control input, and therefore will serve as the basis for open loop control (Kragt and Landeweerd, 1974), and it may create a set of expectancies for how the process variables should respond to inputs and covary with each other. When these expectancies are violated, the model may serve as the basis for detection and diagnosis of process failures (Rasmussen, 1981).

Despite the importance of the internal model to control and failure detection, the operator is not necessarily aware of the model he adopts (Moray, 1980). Further, the internal model is largely an uninvestigated concept, subject primarily to speculation rather than experimental analysis (Rouse & Morris, 1985). Because it is an inferred construct empirical, investigations of the internal model must analyze the operator's overt behavior in attempting to understand it.

Jagacinski and Miller (1978) felt that the human operator's pattern of behavior in a control task could represent the internal model. Subjects performed a tracking task in which they could select one of two discrete control actions: exerting a positive force or a negative force. The task was to bring a moving dot to rest at a specified target position from an arbitrary initial position in a minimum amount of time. The subject's ability to predict the motion, or path, of the dot would be overtly

expressed in the control behavior (switching from a positive or negative force). There were optimal points at which to switch. Thus by comparing deviations of actual switch points to the optimal points, Jagacinski and Miller were able to quantify the operator's internal model. However, the representational form of the model was not obvious. Gill et al. (1982), and Eberts and Schneider (1985) have further pursued this line of approach using a continuous rather than discrete tracking task, and inferring the operator's understanding of system dynamics respectively, through the timing at which discrete controls were implemented, and through a transfer of training design.

One important defining characteristic of the model of dynamic systems is the extent to which it is represented in terms of verbal propositions as opposed to analog spatial images. While the typical process itself is clearly analog involving continuous changes in variable states over time, there is no reason why the mental model might not be represented in terms of a series of verbal propositions of the form: "if the pressure is high at X, the temperature should decline at Y."

Using this perspective, Bainbridge (1974, 1981) asserted that knowledge of a system's dynamic relations could be modelled by conditional verbal statements or program-like routines. Thus in her investigations of the internal model, she used verbal protocols, obtained by asking controllers to think aloud while executing the task (Bainbridge, 1974). Crossman and Cooke (1962) and Brigham and Laios (1975) have also attempted to determine an operator's mental processes through analysis of verbal communication.

Umbers (1979) has criticized this method by noting that a controller's explanation does not necessarily represent his actual thinking. Rasmussen (1981) elaborated on this point by arguing that protocols give very little

information concerning the underlying processes but rather are a sequence of statements indicating "states of knowledge" concerning the operational state of the plant, operator tasks and actions, etc. The operator spontaneously knows where to direct his attention. It is what directs this behavior that has yet to be analyzed. In order to explore more objectively the nature of this representation on a verbal-spatial continuum, two approaches are contrasted below: that based upon individual differences in spatial-verbal ability, and that based upon differences in dual task interferences.

Spatial-verbal ability differences. One approach to examining the verbal-spatial contrast is to examine the effects of individual differences in these abilities on performance of a task which is assumed to be supported by an internal model. Individual differences are, by virtue of their definition, reflected in task performance. Hunt, Frost, and Lunneborg (1973), in noting the wide range of individual differences in performance of information processing tasks, asked if it was possible to identify high and low verbal ability using such tasks. They contrasted the performance of high and low verbal subjects, as designated by the verbal composite from the Washington Pre-College Test (equivalent to the SAT), on a number of information processing tasks. Reaction times for name identification in the letter name identity task, developed by Posner, Boies, Eichelman, and Taylor (1969), were significantly lower for high verbal subjects. Thus, certain performance parameters could distinguish individuals with high scores on psychometric tests from those with low scores.

Klee and Eysenck (1973) investigated comprehension latencies of sentences varying in concreteness and meaningfulness that were obtained under conditions of visual and verbal interference. Subjects of high and low image ability were employed. Results indicated that imagery ability was

not a predictor of performance over all conditions (i.e., there was not a main effect for imagery ability). However, they found a significant interaction across interference conditions between sentence concreteness and imagery ability, such that high imagers produced shorter comprehension latencies than low imagers with abstract sentences, but not concrete ones.

Individual differences also are associated with the selection of information processing strategies. MacLeod, Hunt, and Mathews (1978) gave 70 subjects a sentence-picture comprehension task developed by Clark and Chase (1972) in which the subjects had to verify agreement between sentence-picture pairs. Reaction times of subjects were described in terms of the Carpenter and Just's (1975) constituent comparison model that requires the use of a linguistic strategy. Two groups were isolated. One group fit the model well. The second group fit the model poorly and analyses suggested that they employed a pictorial-spatial strategy. Psychometric measures (scores on the Washington Pre-College Test) indicated that subjects using the pictorial-spatial strategy had significantly higher spatial ability. This demonstrates that different comprehension strategies can be used consistently by different subjects and that the choice of strategies can be predicted from psychometric measures of certain types of cognitive ability. Data such as these support the assertion made by Paivio (1971) that high verbal or high spatial ability allows for those respective processes to be more available for processing information.

Research on individual differences in process control and monitoring has been scarce. However, in one investigation of this sort, Landeweerd (1979) examined the process controller's tasks from the perspective of individual differences in ability along the verbal-spatial continuum and has concluded that those with good verbal-causal understandings of the process

(what leads to what) make better controllers, while those with a better visual-spatial image (what is located where) are better at fault diagnosis.

Dual task performance differences. A second approach to investigating the code of representation of an internal model is through dual task methodology (Wickens, 1984a,b). The assumption is made that process monitoring depends upon active comparisons of values in working memory. The research of Baddeley and his colleagues (Baddeley and Hitch, 1974; Baddeley and Lieberman, 1980) indicates that there appear to be two working memory systems, one involving a verbal phonetic rehearsal loop and the other described by the metaphor of a visual spatial scratchpad. From the present perspective, the important characteristic of these two systems is that they appear to depend upon different processing resources (Wickens, 1984b; Friedman et al., 1982; Polson et al., in press). Therefore tasks depending upon verbal working memory will be more likely to be disrupted by concurrent phonetic and verbal activity, while tasks employing spatial working memory will be more disrupted by concurrent spatial and manual tasks (Baddeley and Lieberman, 1980; Brooks, 1968). The dual task methodology thus offers another means for assessing the code of representation of the internal model of the process. If the model is spatial, concurrent performance of the primary monitoring task with a secondary spatial task should be more difficult than concurrent performance with a secondary verbal task. If the model is verbal, then the converse results should be obtained.

The objective of the present study is to combine the individual differences and dual task methodology to infer the code of representation of the internal model of a complex dynamic system for individuals of high spatial or high verbal ability. This objective is to be accomplished by examining the level of performance on a multivariate process monitoring task

obtained by the two groups, and the pattern of dual task interference of each group, as each was required to time-share system monitoring with verbal and with spatial secondary tasks. Three alternative patterns of data can be predicted, each corresponding to different characteristics of the internal model and of the capacity of the two working memory systems that might be exhibited by the two groups.

(1) Flexible model. Each ability group (verbal-spatial) employs an underlying model of the dynamic system that capitalizes upon that group's respective strength. Hence, performance of each group is expected to be most disrupted by the concurrent task that corresponds to the group's label (i.e., verbal task for the verbal group, spatial task for the spatial group). In addition, there might potentially be a disadvantage for the verbal group on primary task performance because they deploy a model code that is less compatible with the inherently spatial analog system that is to be monitored.

(2) Fixed model-fixed capacity. Both ability groups possess the same underlying internal model code. Because the system is inherently a spatial analog one, it is logical to assume that the internal model code would also be spatial. As a consequence, the spatial secondary task should disrupt primary task performance (or be disrupted by the primary task) to a greater extent than the verbal. In addition to this main effect however, there are two nested predictions of how the spatial and verbal groups might differ in their interference patterns. These predictions are based upon different assumptions of how spatial subjects differ from verbal in terms of their working memory capacity. According to a fixed capacity model, both groups possess the same spatial memory capacity, but the spatial group deploys this capacity to spatial tasks with greater efficiency. Hence, the hypothetical

performance resource function (Norman & Bobrow, 1975) relating performance on the task (vertical axis) to resources invested in the task (horizontal axis) would appear as shown in Figure 1a. The spatial group, indicated by the dashed line requires fewer resources to perform the primary task effectively, and hence would be less disrupted by allocating a given amount of those resources to the spatial side task than would be the verbal group. Such a model might also predict that the spatial group would produce somewhat better primary task performance. Model (2) is consistent with assumptions made by Hunt & Lansman (1982) to the extent that single task ability differences will be directly manifest in dual task time-sharing differences.

(3) Fixed model-variable capacity. This model also assumes that both ability groups deploy a spatial model, and so again predicts that both will be more disrupted by performing the spatial secondary task. However, unlike model 2 it predicts that the spatial group differs from the verbal by having more total spatial resources to utilize, even though these might not be deployed in a more efficient manner as in model (2). Such a difference, shown in the hypothetical performance resource functions of Figure 1b, would suggest that the difference between groups is a general one related to the resources available and is not task specific related to the resources consumed. It would predict an equal level of performance loss of the two groups as the same amount of resources are withdrawn, but both the baseline and the dual task level of performance would presumably be higher for the spatial ability group.

In the present experiment subjects were given the primary task of monitoring 5 continuously varying indicators of the state of a slowly changing closed loop dynamic system. Concurrently each of two secondary

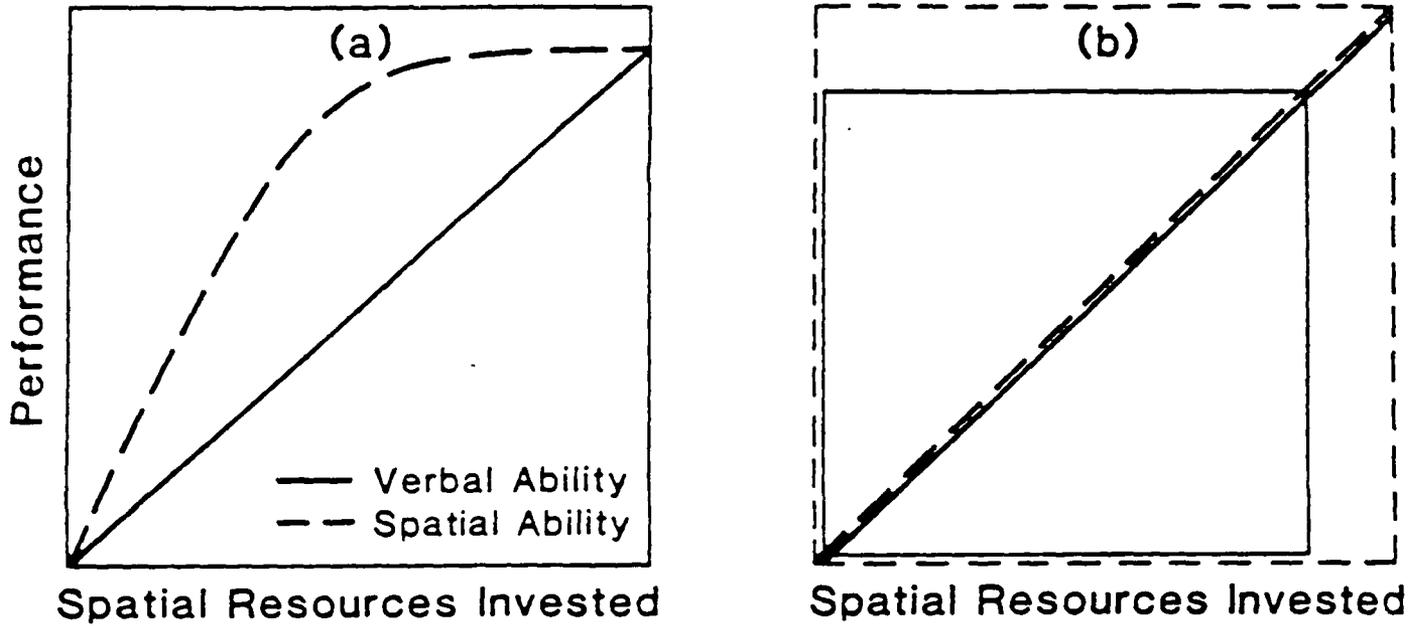


Figure 1: The performance resource function of two models: (a) fixed capacity for both ability groups, (b) variable capacity between ability groups.

tasks were imposed, one requiring spatial working memory and the other verbal. The two groups of subjects were selected from a larger population to have either high spatial and low verbal abilities (relative to the population mean), or low spatial and high verbal abilities. These groups will be referred to as "spatial" and "verbal", respectively.

METHOD

Subjects

Eighteen subjects were chosen from a pool of seventy-seven Psychology 103 students who expressed interest in participating in psychology experiments. The seventy-seven students were given \$1.50 for taking four paper-and-pencil tests which would discriminate verbal and spatial ability. The vocabulary test of the Nelson-Denny Reading test (Nelson & Denny, 1960) and the Grammatical Reasoning task were used to identify students with relatively high and low verbal ability. The Card Rotation test of the Kit of Reference Tests for Cognitive Factors (French, Ekstrom, & Price, 1963) and the Rotated Letters task (Cooper, 1980) were used to identify students with relatively high and low spatial ability. Quartiles from the distribution of scores for these seventy-seven students were used as guidelines for subject selection. As nearly as possible, nine students whose scores fell in the fourth quartile of the verbal tests and the first quartile of the spatial tests were used as subjects with relatively high verbal-low spatial abilities. Three males and six females were chosen. The same procedure was used for selecting nine subjects with relatively high spatial-low verbal abilities. Six males and three females were chosen. Table 1 presents the test scores, in percentages, of the selected subjects. Figure 2 presents a bivariate distribution of the combined verbal and spatial scores so as to demonstrate group separateness.

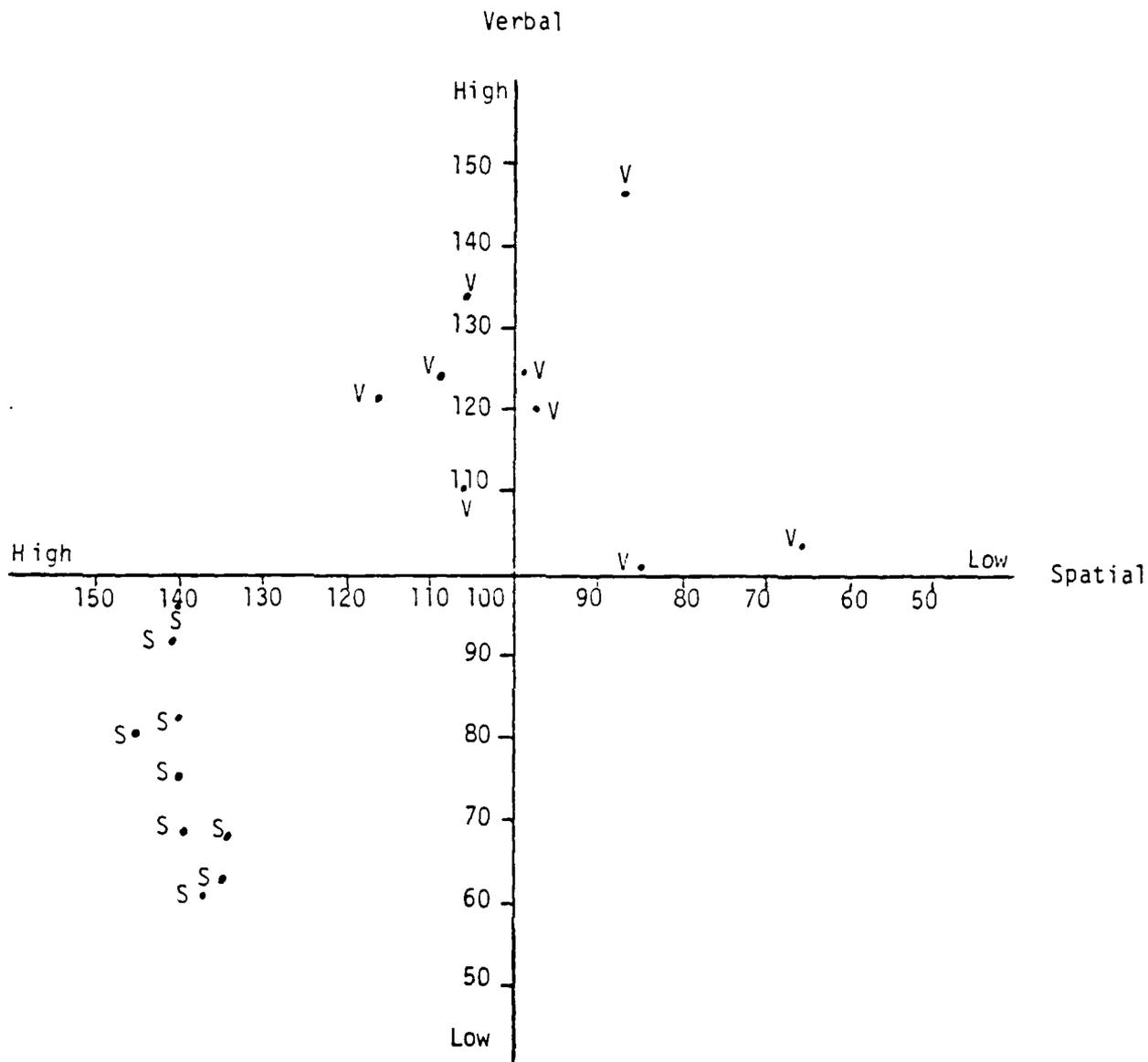


Figure 2. Bivariate distribution of scores on tests of verbal and spatial ability for the high verbal-low spatial (V) and the low verbal-high spatial (S) groups.

Subjects' Ability Test Scores in Percentages

| Group | TESTS | | | |
|------------------------------|------------|-----------------------|----------|-----------------|
| | Vocabulary | Grammatical Reasoning | Rotation | Rotated Letters |
| High Verbal - Low Spatial | 99 | 47 | 62 | 25 |
| | 74 | 50 | 64 | 35 |
| | 73 | 47 | 68 | 29 |
| | 70 | 34 | 46 | 20 |
| | 68 | 56 | 71 | 39 |
| | 67 | 69 | 71 | 35 |
| | 66 | 44 | 71 | 34 |
| | 62 | 59 | 84 | 32 |
| | 45 | 56 | 63 | 22 |
| High Spatial- Low Verbal | 43 | 25 | 93 | 45 |
| | 52 | 31 | 100 | 40 |
| | 32 | 31 | 100 | 35 |
| | 31 | 38 | 99 | 40 |
| | 43 | 25 | 99 | 35 |
| | 46 | 34 | 94 | 51 |
| | 50 | 25 | 90 | 50 |
| | 54 | 38 | 90 | 51 |
| | 59 | 38 | 88 | 52 |
| Quartiles Fourth First | 62 | 56 | 94 | 44 |
| | 43 | 31 | 64 | 32 |

Table 1. Test scores of the selected subjects.

These eighteen subjects were paid \$3.00 an hour for participation in the experiment. All subjects were right-handed.

Apparatus

The failure detection task was displayed on a Hewlett-Packard 1310a 39 x 27.5 cm. CRT display. A PDP11/40 digital computer and an IMLAC PDS4 graphic display provided the input to the CRT. Subjects' reaction times were processed by and stored on the PDP11/40. The two memory tasks were recorded on a Sony TC-654-4 tape recorder. Subject responses for the secondary tasks were recorded by hand.

Subjects were seated in a light attenuated room. The subject sat directly in front of the CRT screen, about 82 cm. from the display. A panel rested on the subject's lap, on top of which was the microphone and button pressing mechanism with which the subject made responses. For a schematic representation of the experimental setup and display see Figure 3.

Task Description

Failure detection task. The failure detection task required the subject to monitor and detect failures in a simulated, highly simplified, five-variable, nuclear power plant. The five variables were: 1) the temperature of water in the reactor core, 2) the depth of the rods in the reactor core, 3) the pressure of water in the cooling system, 4) the temperature of the steam in the steam generator, and 5) the pressure of the steam driving the turbine. The variables were displayed as five bars which varied in length over time. The subject's goal was to understand the overall pattern of relations between variables under normal conditions so as to be able to detect, as soon as possible, a departure from normal. Failure detection was indicated by pressing a button on a panel which rested in the subject's lap. Reaction time was recorded in milliseconds.

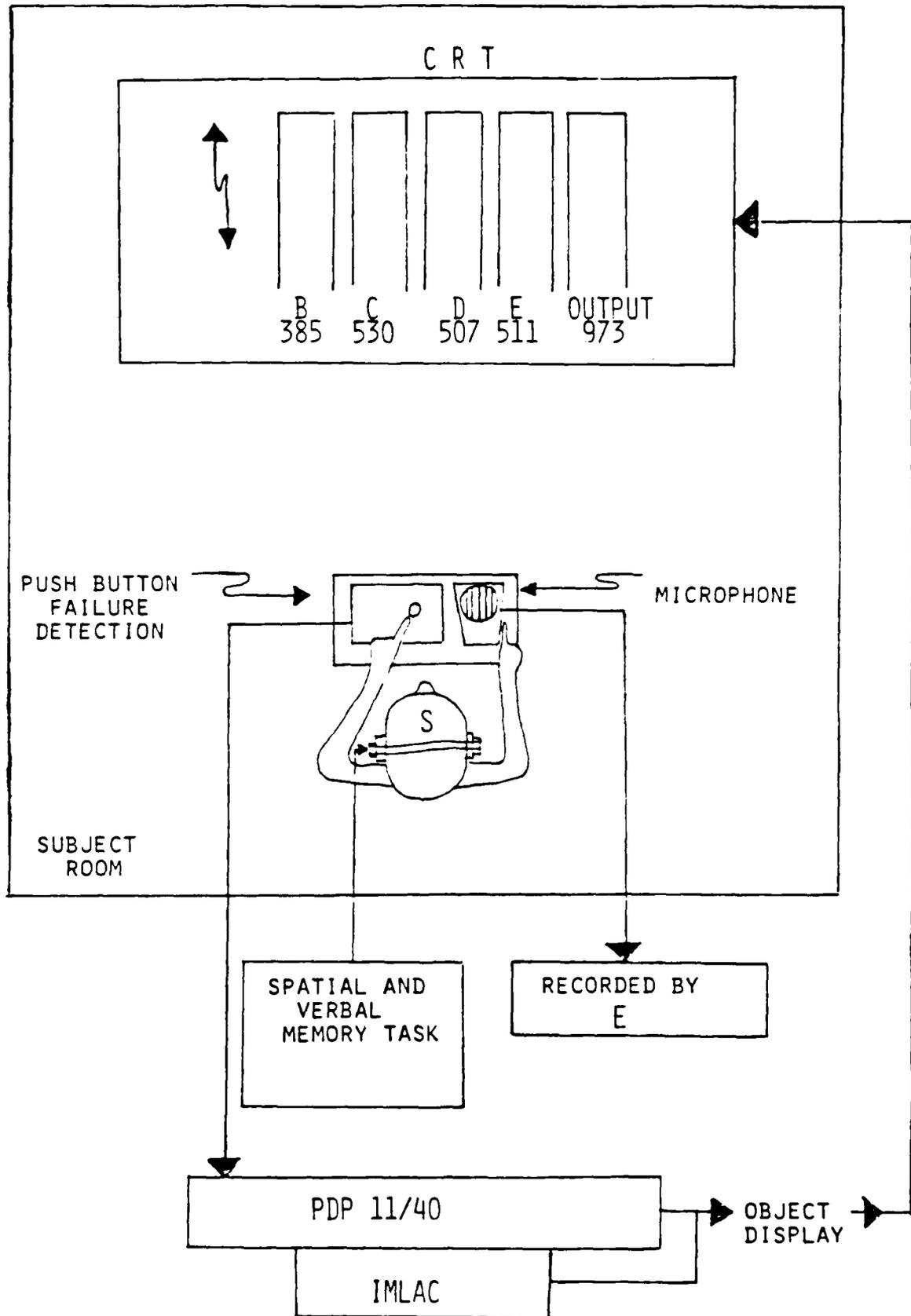
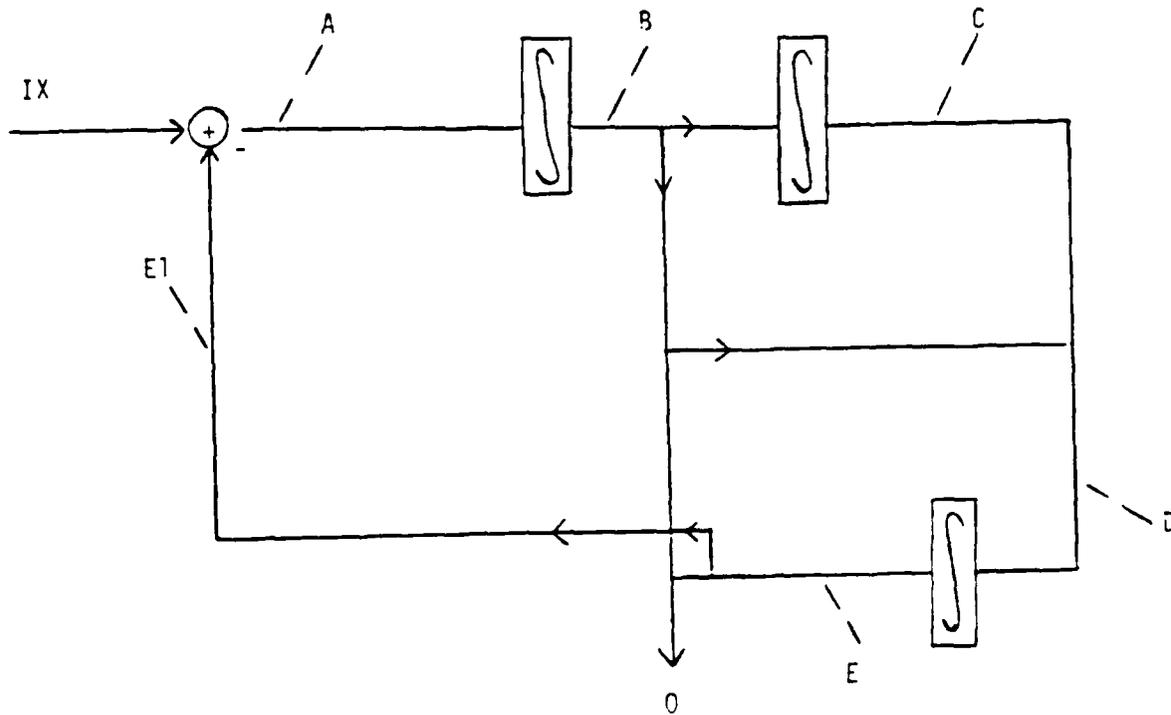


Figure 3. Schematic representation of experiment setup and display.

A Fortran program was written to compute and display a system in operation. Figure 4 gives a mathematical diagram of the system and includes the equations for the variables in the system. The simulation was of a complex negative feedback, stable system "driven" by a slowly changing random input signal IX. When failures occurred, they were manifest as gradual ramp changes in gain values from normal values to terminal steady state values. The ramps were ten seconds in length. From the subject's point of view, this produced gradual changes in the amplitude and frequency response of different variables, and in the degree of coupling or covariation between pairs of variables within the system. After each failure was initiated the subject was given ten seconds to detect it. If a response was not made within this interval, the sequence was scored as a MISS and the system was automatically reset to normal operation. If the failure was detected, the sequence was scored as a HIT and the system, again, automatically reset itself at the end of the ten second interval. Failures in the system occurred at random intervals, averaging four per three minute trial. Figure 4 includes those changes in the equations which produced the two types of failures.

Spatial memory task. The spatial memory task was derived from Brooks' (1967) matrix task. The subject was asked to imagine a four x four matrix and heard, through headphones, directions to place the number "1" in a particular square. The subject was then directed every three seconds to place a consecutive number in an adjacent square (up, down, left, or right). On average, eight digits were placed in the matrix, with a range of seven to ten. After the last number was placed in the matrix the subject was cued with a number to which she responded with its location through a microphone. The subject was given four seconds to respond after which the next matrix



$$A = (IX - E \times S7)/2$$

$$B1 = A + B1$$

$$*B = B1 \times S1$$

$$C1 = C1 + B$$

$$C2 = C1 \times S2$$

$$D2 = B \times S4 + C2 \times S4$$

$$D1 = B \times S3 + C2 \times S4$$

$$E1 = E1 + D2$$

$$E1 = E1 \times S5$$

$$*O = E1 + B \times S6/2$$

$$*C = C2/5$$

$$*D = D1/4$$

$$*E = E1$$

$$S1 = .7$$

$$S2 = .5$$

$$S3 = .5$$

$$S4 = .5$$

$$S5 = .3$$

$$S6 = .6$$

$$S7 = .5$$

* = Variable on display

Failure 1: $S3 = .5 \rightarrow 2$ through increments of $.09/100$ msec

Failure 2: $S5 = .3 \rightarrow 6$ through increments of $.01/100$ msec

Figure 4. Mathematical diagram of dynamic system used in failure detection task. Shown at the bottom are the equations used to compute the values of the system's variables and those changes in the equations which produced the two types of failures.

problem began. Three matrices were presented in a three minute trial. The percentage of correct responses were recorded by hand.

Verbal memory task. In this task the subject listened through headphones to a list of abstract words, each word presented at three second intervals. The lists averaged four words in length, with a range of three to six. After the list was presented, the subject was cued with a word from the list and orally responded with the word which came before it. The subject was given four seconds to respond and then the next list was presented. Seven word lists were presented in a three minute trial. The percent of correct responses were recorded by hand.

The verbal memory task in this experiment used 104 nouns taken from a list of 925 nouns developed by Paivio, Yuille, and Madigan (1969). The nouns chosen were those rated with an imagery value less than three on a scale of one to six, in which one meant "arouses an image not at all." This ensured that the words were not easily visualized and the task was verbal in nature. The task put demands on verbal working memory and briefly unloaded during the response period. An exemplary list follows:

Subject Hours: "amount, essence, ingratitude, answer."

Cue Report: "ingratitude."

Subject's Response: "essence."

A pilot study was run to determine average error rates for different levels of difficulty for each of the two secondary tasks. Those levels which yielded equivalent error rates that were between .2 and .5 were used. It was assumed that error rates within this range were indicants of a task that was neither too difficult nor too easy. Four male volunteers ran in the pilot.

Procedure

Each subject performed for two hours on each of four days. The first two days were used for practice, an amount, as determined by pilot work, that was necessary to stabilize performance. Instructions were first given on the failure detection task (see Appendix A). The subject was given equal amounts of verbal and graphic explanations of the system so as to avoid a bias toward developing one type of internal representation over the other. In the thirty 3-minute trials given for practicing this task, a subject could see one of three types of trials: 1) a normal operations trial in which no failures occurred; 2) an announced alarm trial during which the word "ALARM" appeared when a failure occurred and remained for the duration of the failure (this provided a graphic demonstration of the nature of the failure and the difference between normal- and failed-system response); 3) a failure detection trial in which a failure occurred but the word "ALARM" did not appear. The subject was instructed to detect failures as quickly as possible but to avoid making false alarms.

After completing the practice for the failure detection task, the subject was given written and oral instructions on the two memory tasks. The subject practiced each task until performance stabilized (i.e., remained approximately the same for three trials) at or above a sixty percent correct response rate. If the subject could not reach that criterion then further practice was provided until the percentage of correct responses remained the same for three trials. On average, for the verbal group, this required four trials for the verbal task and nine for the spatial. For the spatial group, this training required an average of five trials for the verbal task and seven trials for the spatial task.

On the third day, the subject first practiced each of the three tasks individually. Then the subject practiced performing the failure detection task concurrently with each of the memory tasks. The subject was instructed to keep performance on both tasks in the dual task condition as nearly as possible at the same level as when performing each task alone. Two practice trials were allowed for each of the dual task conditions. The subject also began the fourth day with two practice trials in each of the dual task conditions. Practice performance for each dual task condition was averaged and used as a baseline for judging maintenance or improvement in dual task performance on that day. A bonus system was implemented to provide subjects an incentive for maintaining/improving dual task performance.

Following practice on the third and fourth day, the subject performed three experimental blocks of trials. Each block consisted of each of the three single tasks conditions, followed by the two dual tasks. During the dual task trials, both instruction and a financial bonus system were used to induce the subjects to allocate attention equally between both primary and secondary tasks. The bonus system rewarded subjects for performance on all tasks that was close to, or better than their single task level. Ordering of trials within the single and dual task conditions was randomized. Prior to presentation of a trial, the subject was informed of the task to be performed and was instructed over an intercom when the trial was to begin.

RESULTS

Data from both the primary and secondary task were subjected to repeated measures analyses of variance. For the secondary tasks, only accuracy (% correct) was employed, whereas separate analyses were conducted on primary task latency, correct detection, and false alarm rate. Rather

than describing each ANOVA in turn, the effects will be discussed in terms of certain fundamental underlying hypotheses.

It was necessary first to insure that the two ability groups that had been selected on the basis of paper-and-pencil tests continued to be differentiated on performance of the dynamic tasks employed in the main experiment. Simultaneously, it was essential for the hypotheses under consideration to guarantee that the two memory tasks were in fact spatial and verbal in their underlying processing demands. Both effects could be simultaneously validated by the presence of a strong "crossover" interaction between task type and ability group.

As indicated in Figure 5, which plots the accuracy of each secondary task as a function of ability groups, such a crossover interaction was clearly and unequivocally obtained ($F(1,16) = 4.26, p = .05$). Both tasks and ability groups showed roughly the same level of overall performance (both main effects showed $F < 1$), but each group performed best with the task compatible with their own abilities.

Having confirmed then that the secondary tasks were in fact "spatial" and "verbal" (as operationally defined by group ability differences), it was then possible to examine their influences on the primary task. Table 2 shows the measure of primary task failure detection accuracy for the single and the two dual task conditions. These detection data were submitted to a 3 (condition) x 2 (group) repeated measure ANOVA. From the table it is apparent that there was a dual task decrement--detection accuracy was lower with the secondary tasks present than when they were absent. Secondly, this decrement appears to be greater with the spatial than with the verbal task for both groups. These effects were substantiated statistically by the main

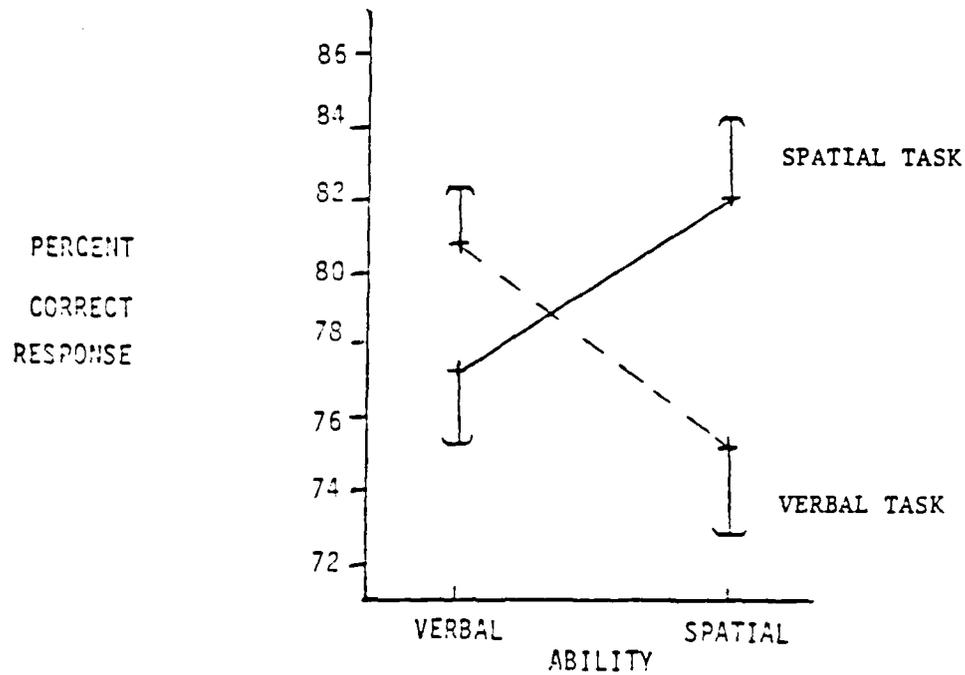


Figure 5. Performance accuracy on secondary memory task as a function of task type and ability group.

| | <u>Secondary Task Condition</u> | | |
|---------|---------------------------------|--------|---------|
| | None | Verbal | Spatial |
| Verbal | 96 | 95 | 91 |
| Spatial | 97 | 95 | 94 |

Table 2. % correct detection on primary task.

effect of task ($F(2,32) = 4.19, p < .05$), and by the subsequent HSD planned comparisons test between the means which indicated a reliable loss of performance accompanying the introduction of the spatial side task ($HSD(32) = 4.02, p < .05$), but not the verbal side task. There were no significant effects nor interactions involving group in the primary task data. That is, both groups showing the same (statistically) pattern of interference effects on the primary task. Neither of the other primary task dependent variables, detection latency, nor false alarm rate showed statistically reliable effects of any of the independent variables.

The secondary task data as a source of information regarding dual task interference are shown in Figure 6. For both ability groups, the data show a pattern of interference supporting the conclusion drawn from the primary task data, that the failure detection task requires spatial resources. Under single task conditions, the spatial tasks are performed as well or better than the verbal tasks, while under dual task conditions the pattern reverses. Stated in other terms, the spatial task shows a substantial dual task decrement, while the verbal task shows almost none. This effect is substantiated by the significant task x condition interaction ($F(1,16) = 6.39, p < .05$). The most interesting aspect of the data however is that the pattern of task load by task type interaction is identical for both ability groups (left and right panel), and the 3-way interaction of group x task x load failed to approach significance ($F(1,16) = 0.32$). The main difference between the two groups is the improved performance of the spatial group on the spatial task, an interaction already noted above.

Finally, it should be noted that there were reliable effects of practice (across the six trials in each condition) on primary task response latency ($F(5,80) = 2.78, p < .05$) and secondary task performance ($F(5,80) =$

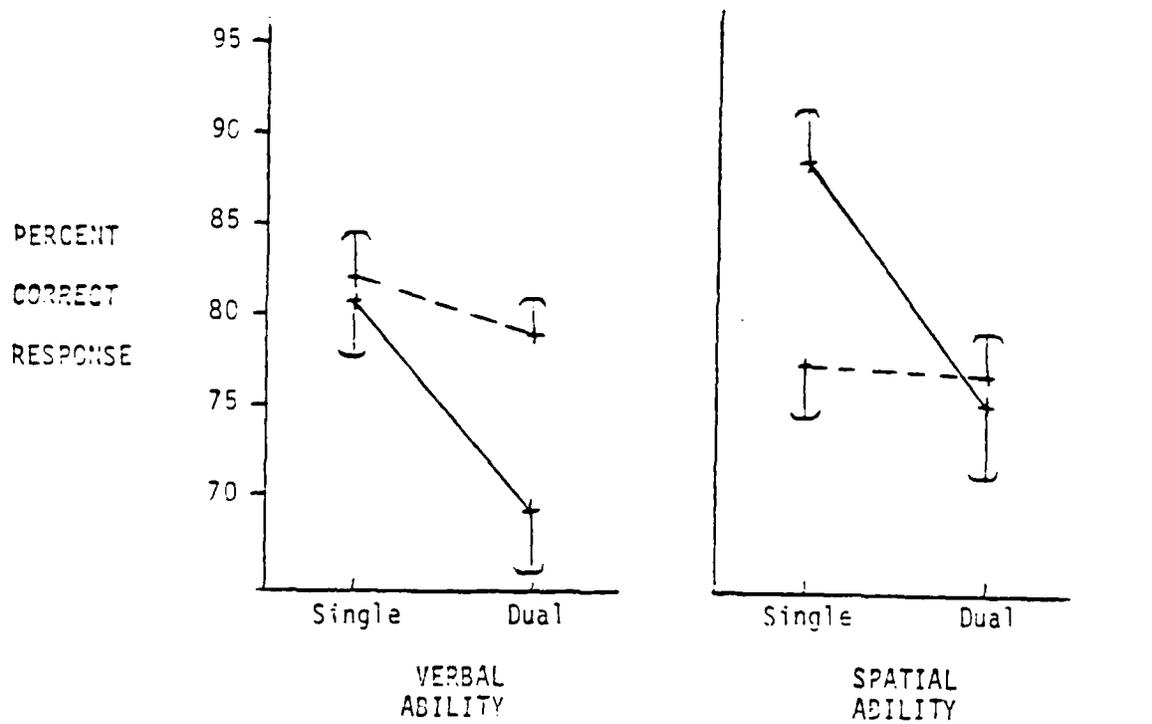


Figure 6. Secondary task performance data as a function of task type, task load (single-dual) and ability groups.

3.18, $p < .05$). The latter measure also showed a significant task load (single-dual) by practice interaction ($F(5,80) = 3.45, p < .05$). This interaction reflected a large increase in dual task performance from trial 1 to trial 2, which was not manifest in performance of the single task. Neither task type nor ability entered into any reliable interactions with practice.

DISCUSSION

The pattern of the dual task interference data confirm that the demands of the primary task are spatial for both ability groups. Therefore the code of the internal model appears to be task-driven rather than ability-driven. Both groups showed substantial interference on both the primary and the spatial secondary task, while both also showed nearly perfect time-sharing with the verbal secondary task. The results presented in Figure 5 also clearly indicated that the ability and task type manipulations were successful in the direction intended. The interesting question then is how the two ability groups differed in terms of their time-sharing of verbal and spatial material. The answer to this question can be considered in the framework of the fixed and varied capacity models described in Figure 1.

The data in Figure 6 indicate that the fixed capacity model must be rejected. This model predicts that the spatial ability group uses its fixed quantity of spatial resources more efficiently than does the verbal group. Were this the case, then dividing resources between tasks should produce a lesser decrease in performance on the primary and/or the secondary task for the spatial group. According to the data presented in Table 2 and Figure 6, this was not the case.

The varied capacity model then seems to be more plausible. According to this model the spatial group differs from the verbal simply by having

more total capacity available; for example, a "larger" visual-spatial scratchpad, or one from which information is lost at a slower rate. This interpretation assumes that all spatial resources are deployed by both groups when performing the spatial secondary task. Hence, secondary task performance will be better for the spatial group, who can place more "units" of resources in service (e.g., 4 units for the spatial group versus 3 for the verbal group). It assumes secondly that an equivalent amount of resources (e.g., 3 units) are deployed to the single task failure detection task by each group (this is the maximum of 3 units for the verbal group and one unit less than maximum for the spatial group). Hence, sharing resources between the two tasks will leave the same "shortfall" between demanded and available resources in dual task conditions (7 demanded - 4 available = 3, for the spatial group; 6 - 3 = 3 for the verbal group). Since this shortfall predicts the magnitude of the dual task decrement, such an accounting can readily explain the equivalence in timesharing performance between groups shown in Figure 6 and Table 2.

This accounting of the results is less than fully satisfying however in one respect. If the spatial group has more spatial resources available, why don't they deploy more of them to improve performance on the primary task, relative to the verbal group? One possibility is that such improvements were in fact present, but they were distributed across the three performance measures of the primary task (hits, false alarms, and latency) in such a way that none of the measures alone showed reliable group differences. In fact, all three measures did "favor" the spatial group in single as well as dual task performance. Single task response latency was slightly faster for the spatial group (3.08 vs. 3.12 sec). Accuracy was higher (97% vs. 96%) and false alarm rate was lower (17% vs. 22%). However, none of these variables

examined in isolation reached significance. Nevertheless such data could suggest that the primary task was indeed performed better by the spatial group, as a consequence of their investment of all of the resources toward its performance (i.e., 4 units).

A second possible accounting of the results is that the pattern of task interference was attributable to factors that were quite independent from the source of group differences. According to this view, the spatial and verbal groups may indeed have differed in terms of their working memory capacities; however; the primary task was considerably more visual than spatial in its demands and the two groups did not differ in terms of their basic visual capacity. This distinction between spatial and visual resources is one that has been nicely demonstrated by Baddeley and Lieberman (1980). Hence, the decrement that was obtained in dual task conditions was simply the consequence of visual interference resulting perhaps from a greater degree of unwanted (and unnecessary) visual scanning that was induced when the spatial side task was performed. The fact that both side tasks used auditory presentation would not preclude the possibility of induced scanning. This scanning would of course lead to a loss in performance because of the heavy visual demands of the primary task. At the present time, these two interpretations cannot be discriminated.

In conclusion, the results stand in contrast with those of Landeweerd (1979) who found ability differences in performance on process control monitoring kinds of tasks. Such differences as noted were not observed here. Although the number of differences between the two paradigms do not require one to conclude that they are in conflict. The results do appear to support the interpretation that the processes used to perform the failure detection task used here were visual-spatial and not verbal phonetic. As

noted above, whether the "spatial" or the "visual" is the more important component of demands cannot be easily determined. The former refers to characteristics of an internal model described by the capacity-based first hypothesis proposed above; the latter to the peripheral scanning interference of the second hypothesis. The data do suggest the feasibility of imposing concurrent tasks of verbal phonetic nature on the process monitor, with little fear of interference. Hence, the data are compatible with the emerging distinction between verbal and spatial resources as a useful dichotomy for describing and predicting dual task performance (Wickens, 1984b). The precise way in which this performance is influenced by ability differences clearly require further research.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Amir Mane for his efforts in programming the failure detection task, and Dr. David Dupree for providing the test data from which the sample of subjects were chosen.

REFERENCES

- Baddeley, A.D., & Hitch, G. (1974). In G. Bower (Ed.), Recent advances in learning and motivation (Vol. 8). New York: Academic Press.
- Baddeley, A.D., & Lieberman, K. (1980). Spatial working memory. In: R.S. Nickerson (Ed.), Attention and performance VIII. Hillsdale, NJ: Erlbaum.
- Bainbridge, L. (1974). Analysis of verbal protocols from a process control task. In E. Edwards & F. Lees (Eds.), The human operator in process control. New York: Halsted Press.
- Bainbridge, L. (1981, in press). Mathematical equations or processing routines? In J. Rasmussen & W. Rouse (Eds.), Human failure detection and fault diagnosis. New York: Plenum Press.
- Brigham, F., & Laios, L. (1975). Operator performance in the control of a laboratory process plant. Ergonomics, 18, 53-66.
- Brooks, L.R. (1967). The suppression of visualization in reading. Quart. Jrnl. of Exptl. Psych., 19, 289-299.
- Brooks, L.R. (1968). Spatial and verbal components in the act of recall. Cand. Jrnl. of Psych., 22, 349-368.
- Carpenter, P., & Just, M. (1975). Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review, 82, 45-73.
- Clark, H., & Chase, W.G. (1972). On the process of comparing sentences against pictures. Cognitive Psychology, 3, 472-517.
- Cooper, L. (1980). Recent themes in visual information processing: A selective overview. In R. Nickerson (Ed.), Attention and performance VIII. New York: Academic Press.

- Crossman, E.R.F.W., & Cooke, J. (1962). Manual control of slow-response systems. Paper presented at the International Congress on Human Factors in Electronics, Long Beach, California.
- Eberts, R., & Schneider, W. (1985). Internalizing the system dynamics for a second-order system. Human Factors, 27, 371-394.
- French, J., Ekstrom, R., & Price, L. (1963). Manual for kit of reference tests for cognitive factors. Princeton, NJ: Educational Testing Service.
- Friedman, A., Polson, M.C., Dafoe, C.G., & Gaskill, S. (1982). Dividing attention within and between hemispheres: Testing a multiple resources approach to limited-capacity information processing. Journal of Experimental Psychology: Human Perception and Performance, 8, 625-650.
- Genter, D., & Stevens, A. (1983). Mental Models. Hillsdale, NJ: Erlbaum.
- Gill, R., Wickens, C.D., Reid, R., & Donchin, E. (1982). The internal model: A mean of analyzing manual control in dynamic systems. Proceedings, 1982 Annual Meeting of the IEEE Conference on Systems, Man, & Cybernetics, Seattle, WA.
- Hunt, E., Frost, N., & Lunneborg, C. (1973). Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (Vol. 7). New York: Academic Press.
- Hunt, E., & Lansman, M. (1982). Individual differences in attention. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1, pp. 207-254). Hillsdale, NJ: Lawrence Erlbaum.
- Jagacinski, R., & Miller, R. (1978). Describing the human operator's internal model of a dynamic system. Human Factors, 20, 425-433.

- Klee, H., & Eysenck, M. (1973). Comprehension of abstract and concrete sentences. Journal of Verbal Learning and Verbal Behavior, 12, 522-529.
- Kragt, H., & Landeweerd, J.A. (1974). Mental skills in process control. In E. Edwards and F.P. Lees (Eds.), The Human Operator in Process Control. London: Taylor & Francis.
- Landeweerd, J.A. (1979). Internal representation of a process fault diagnosis and fault correction. Ergonomics, 22, 1343-1351.
- MacLeod, C., Hunt, E., & Mathews, N. (1978). Individual differences in the verification of sentence-picture relationships. Journal of verbal learning and verbal behavior, 17, 493-507.
- Moray, N. (1980). Human information processing and supervisory control. Cambridge: MIT, Man-Machine Systems Laboratory.
- Nelson, M.J., & Denny, E.C. (1960). The Nelson-Denny reading test. Boston: Houghton Mifflin.
- Norman, D.A., & Bobrow (1975). On resources and data limited processes.
- Paivio, A. (1971). Imagery and verbal processes. Chicago: Holt, Rinehart, & Winston, Inc.
- Paivio, A., Yuille, Y., & Madigan, S. (1969). Concreteness imagery and meaningfulness values for 925 common nouns. Journal of Experimental Psychology Monograph Supplement, 79, 458-463.
- Polson, M.C., Wickens, C.D., Klapp, S., & Colle, H. (in press). Human interactive processes. In M. Chignall & P. Hancock (Eds.), Dynamic interfaces. N. Holland.
- Posner, M., Boies, S., Eichelman, W., & Taylor, R. (1969). Retention of visual and name codes of single letters. Journal of Experimental Psychology Monograph, 79, 1-16.

- Rasmussen, J. (1979). Reflections on the concept of operator workload. In N. Moray (Ed.), Mental workload. New York: Plenum Press.
- Rasmussen, J. (1981). Models of mental strategies in process plan diagnosis. In J. Rasmussen & W.B. Rouse (Eds.), Human Detection and Diagnosis of System Failures. New York: Plenum Press.
- Rouse, W., & Morris, N. (1985). Mental models: Inside the black box. Georgia Institute of Technology Center for Man Machine Studies Technical Report.
- Sheridan, T. (1981). Understanding human error and aiding human diagnostic behavior in nuclear power plants. In J. Rasmussen & W.B. Rouse (Eds.), Human detection and diagnosis of system failures. New York: Plenum Press.
- Umbers, I.G. (1979). Models of the process operator, Intl. Jrnl. of Man-Machine Studies, 11, 263-284.
- Wickens, C.D. (1984a). Engineering Psychology and Human Performance. Columbus, OH: Charles Merrill.
- Wickens, C.D. (1984b). Processing resources in attention. In R. Parasuraman & R. Davies (Eds.), Varieties of Attention. New York: Academic Press.

APPENDIX · FAILURE DETECTION TASK INSTRUCTIONS

In the failure detection task, you will assume the role of a nuclear power plant operator, that is you will monitor and detect failures of 5 components of a highly simplified nuclear power plant. the 5 variables are 1) the temperature of water in the reactor core, 2) the depth of the rods in the reactor core, 3) the pressure of water in the cooling system, 4) the temperature of the steam in the steam generator, and 5) the pressure of the steam driving the turbine.

The system works as follows. Water flows into the nuclear reactor core. This water comes from an outside source and is also part of the cooling loop, thus the temperature of this water in the reactor core fluctuates. The radioactive rods are lowered into the core and this heats the water. This extremely hot, radioactive water flows out of the core, under pressure. This water then flows around pipes, it heats up the contents of the pipes and, in so doing, reduces its own temperature. This water is then channelled back to the core where it is also mixed with cool water from the outside source. Inside the pipes, described above, is nonradioactive water. As said before, this water is heated up and turned to steam. This steam is then channelled, under pressure, to the turbine which it drives. Rotation of the turbine generates electrical power.

As you can understand, all the variables can influence the others.

The 5 variables will be represented on a visual display as bars which vary in length. You will observe the system in normal states of operation, during which each variable effects the others in a normal, expected manner. Periodically however "failures" will occur such as

a leak in a pipe, a clogged pipe, or a pressure relief valve that is stuck open. These failures will change the relative influence that certain variables have on others.

Your goal is to try to understand the overall pattern of relations between the normal variables so that you can detect, as soon as possible, the departure from normal and indicate your detection by pressing a button. If you do not detect a failure, 6 seconds after its onset the system will reset itself. Try to detect failures as quickly as possible but also be certain that there is a failure. That is, in your effort to quickly detect failures, don't make any false alarms.

Any questions?

During this and the next session of the experiment, you will practice doing each task alone.

Now you will practice doing the spatial task and doing the verbal task concurrently with the failure detection task. Try to keep your performance on both tasks in this dual task condition as nearly as possible at the same level as when performing each task alone.

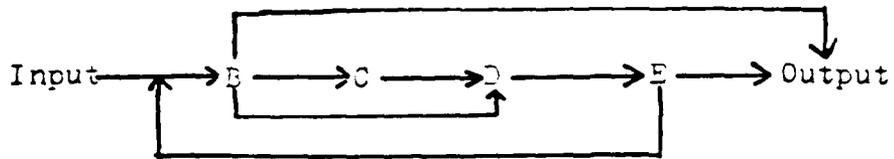
Any questions?

Now that you have had sufficient opportunity to practice these tasks, we enter the second part of this experiment. In this part you will have the opportunity to earn bonuses for your dual task performance. Your failure detection reaction time on each trial will be compared to your average reaction time from practice, as will your performance on the memory tasks. If they are better than

the practice session's, then you will receive a 10 cent bonus for that trial. If you make one false alarm during the trial, you receive only 5 cents. If you make more than 1 there will be no bonus for that trial.

This session will proceed as follows. In order to reacquaint yourself with the failure detection task, you will first watch the system in its normal state, then in an announced failure condition and then an unannounced one (in which you will make a failure detection). Then there will be 3 sets of trials, each will include performing each of the 3 tasks alone and then the 2 dual task conditions.

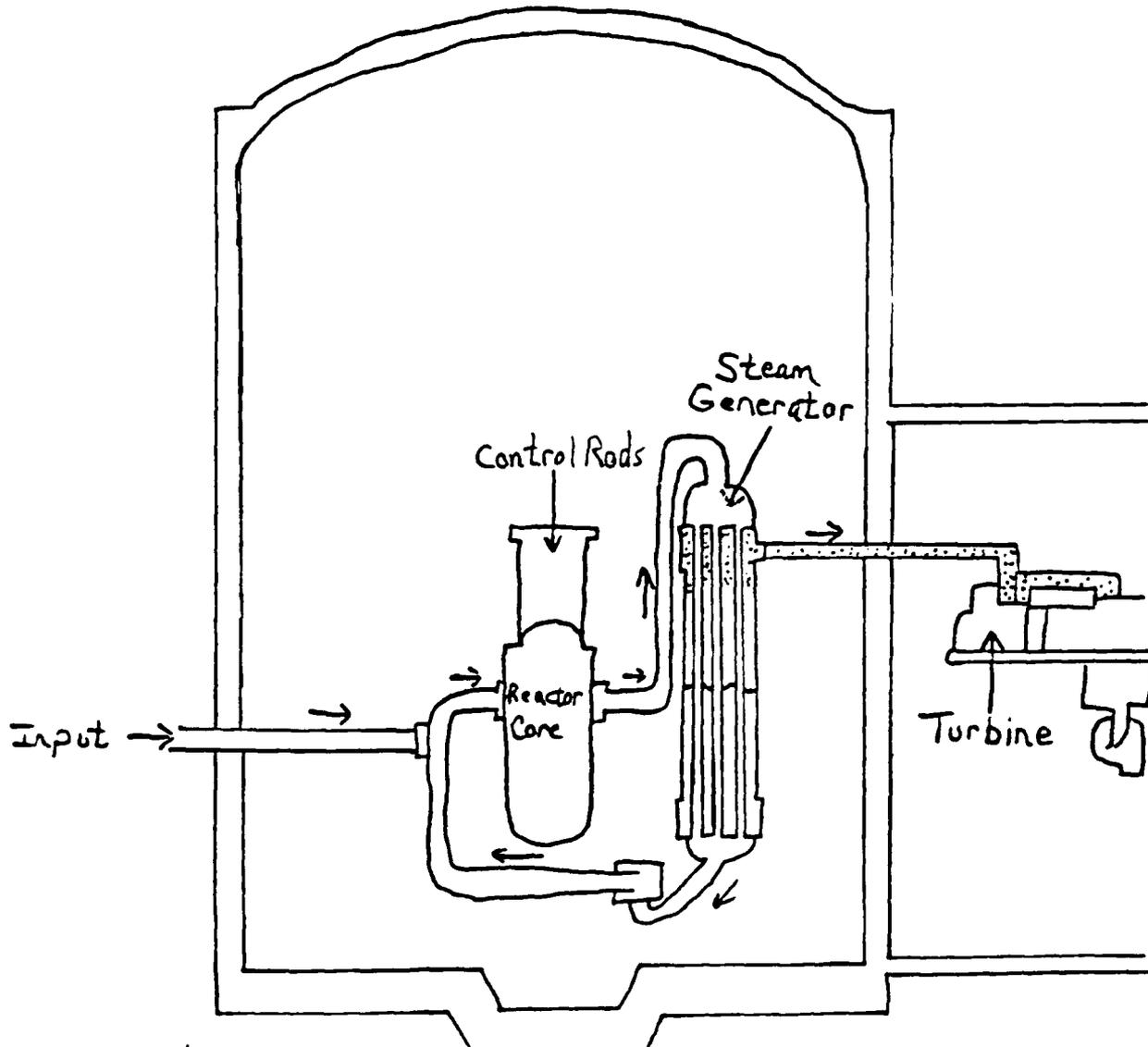
Any questions?



Primary Cooling System {

- Input: water from a cooling tank
- B: temperature of water in reactor
- C: depth of rods
- D: pressure of water in cooling system
- E: temperature of steam in steam generator

 Output: steam pressure driving the turbine



Examples of Failures

Since all variables in the system are related, one failed variable will probably influence most of the system. Some variables will be influenced more than others.

Some examples:

---If the rods get stuck in one position in the core, then the changes in pressure in the cooling system will be caused only by the temperature changes of the water.

---If a leak springs in the cooling loop, hot water will flow through out the containment building--around pipes it previously did not and thereby exert greater influence over (heat) other parts of the system.

---Normally, a relief valve opens to allow excess pressure an outlet. If a relief valve in the cooling system gets stuck closed, then pressure within the system will covary more closely with the water temperature and rod depth than under normal conditions.