Navy Personnel Research and Development Center

San Diego, California 92152-6800   AP-91-15   June 1991

AD-A240 145

Independent Research and Independent Exploratory Development
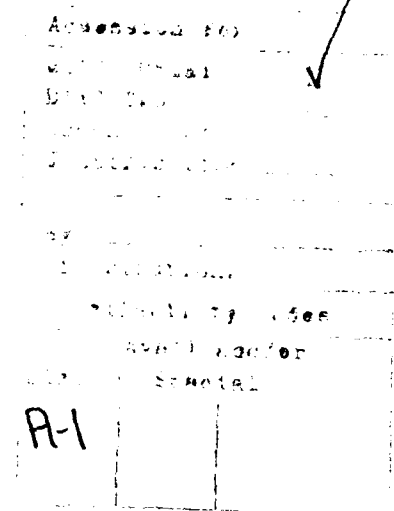
DTIC
ELECTE
SEP 1 0 1991
S   D
C

STOUT

FY90 ANNUAL REPORT

Administrative Information

NPRDC-AP-91-15

June 1991

Independent Research and Independent Exploratory Development Programs: FY90 Annual Report

William E. Montague
Carmen C. Scheifers
(Editors)

Reviewed by
R. C. Sorenson
Technical Director (Acting)

Approved and released by
T. F. Finley
Captain, U.S. Navy
Commanding Officer

Navv Personnel Research and Development Center
San Diego, California 92152-6800

91 9 9 008

Contents

Introduction

The Technical Director encourages scientists and engineers at the Navy Personnel Research and Development Center (NPRDC) to generate new and innovative proposals to promote scientific and technological growth in the organization and the development of knowledge and technology of interest to the Navy. Support for this is provided by discretionary funding furnished by the Independent Research (IR) and Independent Exploratory Development (IED) programs of the Office of Naval Research and the Office of Naval Technology. These programs support initial research and development of interest to the Navy with emphasis on the NPRDC mission areas of the acquisition, training, and effective utilization of personnel.

Funds are provided to the Technical Directors of Navy Laboratories to support innovative and promising research and development outside the procedures required under normal funding authorization. The funds are to encourage creative efforts important to mission accomplishment. They enable promising researchers to spend a portion of their time on examining the feasibility of self-generated new ideas and scientific advances. They can provide important and rapid test of promising new technology and can help fill gaps in the research and development program. This may involve preliminary work on speculative solutions too risky to be funded from existing programs.

The funds also serve as means to maintain and increase the necessary technology base skill levels and build in-house expertise in areas likely to become important in the future. These programs contribute to the scientific base for future improvements in the manpower, personnel, and training system technology and provide coupling to university and industrial research communities.

The FY90 IR/IED programs began with a call for proposals in June 1989. Technical reviews were provided by supervisors and scientific consultants and eight IR and three IED projects were funded. This report documents the results and accomplishments of these projects. Dr. William E. Montague administers the IR and IED programs, coordinating project selection, reporting, and reviewing to assure an innovative and productive program of science and technology.

Tables 1 and 2 list the projects active during FY90 and those supported in FY91. Two papers, one IR and one IED, chosen by the Technical Director as "Best Papers of 1990" are presented. Subsequent pages, which were written by the principal investigators of each project, contain brief reports of research progress during FY90. Appendix A lists the IR and IED projects that may have transitioned into other projects or into use by the Navy during the year. Appendix B itemizes the presentations and publications from IR and IED supported projects. Appendix C presents awards and honors related to the projects.

Table 1
Independent Research
Work Units for FY90 and FY91
(PE 0601152N)

Work Unit	Title	Principal Investigator	Internal Code	Telephone (619) 55 or DSN[a] 55	FY Funding (K) 90	FY Funding (K) 91
0601152N.R0001.01	Brain activity during visual recognition	Ryan-Jones, Lewis, Trejo, Hemmer	14	37710 37711 37711 37711	23.0	0.0
0601152N.R0001.02	Event-related potential correlat s of memory performance	Williams	15	37705	17.0	0.0
0601152N.R0001.03	Using diagrams for learning procedural tasks	Vogt, Gehring	15	37788	32.0	0.0
0601152N.R0001.04	Analysis of tutoring in technical training	Ellis, Montague	14	39273 37849	73.0	0.0
0601152N.R0001.06[b]	Experienced-based career development--II	Morrison	12	39256	15.5	0.0
0601152N.R0001.09	The role of feedback in computer-based training	Cowen	14	37698	50.0	50.0
0601152N.R0001.10	An exploratory examination of neural networks as an alternative to regression	Wilkins, Dickieson	13	37618	26.0	60.0
0601152N.R0001.11	Neural network modeling of skill acquisition	Dickieson, Gollub	15	37849	19.0	70.0
0601152N.R0001.12	Individual differences in brain electrical activity	Ryan-Jones, Lewis	14	37710	45.0	0.0
0601152N.R0001.13	Prediction models for dichotomous criteria	Sands, Wilkins	13	39266	38.5	80.0
0601152N.R0001.14	Neural network analysis physiological data	Ryan-Jones	14	37710	0.0	60.0
0601152N.R0001.15	Experienced-based career development--II	Morrison	12	39256	0.0	20.0

[a]Defense Switch Network
[b]Transitioned.

Table 2
Independent Exploratory Development
Work Units for FY90 and FY91
(PE 0602936N)

Work Unit	Title	Principal Investigator	Internal Code	Telephone (619) 55 or DSN 55	FY Funding (K) 90	FY Funding (K) 91
0602936N.RV36I27.08[a]	Developing measures of effectiveness for knowledge discovery systems	Sorensen	12	37782	44.0	0.0
0602936N.RV36I27.10	Decomposition methods	Thompson Krass	11	37925 37895	75.5	50.0
0602936N.RV36I27.11	Using a neural net approach manpower forecasting performance of a neural net versus conventional software	Huntley	11	38038	63.0	0.0
0602936N.RV36I27.12	Effects of administration method and anonymity/ identification of survey responses	Kewley	12	39251	0.0	60.0
0602936N.RV36I27.13	An examination of cognitive and motivational effects of employee involvement interventions	Sheposh	16	37947	0.0	85.0
0602936N.RV36I27.14	Special journal: What works	Montague	15	37849	0.0	22.0

[a]Transitioning.

Best Papers

Biography

Michael Cowen is a Personnel Research Psychologist in the Training Systems Department. Dr. Cowen has been with NPRDC since 1980 and has had research projects in personnel selection, career development, job turnover, least squares modeling, and computer-based training. He received his B.A. with honors in Experimental Psychology from the University of California, Santa Barbara in 1978 and a M.S. in Industrial/Organizational Psychology from San Diego State University in 1980. He recently received his Ph.D. in Applied Cognitive Psychology from the Claremont Graduate School (1991). Dr. Cowen is a member of the American Psychological Society, the American Educational Research Association, and the Association for the Development of Computer-based Instructional Systems.

Computer-based Training (CBT) for Operating a Digitally Controlled Device: The Role of Feedback

Michael Cowen

Current research provides little guidance as to when feedback should be provided and how to design feedback content in computer-based training (CBT). Experimental CBT on how to operate a digitally controlled device was administered individually to 80 Navy students. Feedback was provided either immediately following an error or delayed until the end of the button-pushing of the to-be-learned sequence. The content of the feedback was either a presentation of the corrective response or a "wrong" indication by the computer. All the CBT treatment groups outperformed a no-treatment control group. Among the treatment groups, those who received delayed feedback scored significantly higher on the post-test than those who received immediate feedback. Analysis of the post-test protocols suggest that the delayed feedback students acquired more conceptual knowledge of how to operate the complex device. Delaying the feedback in CBT may be an effective instructional technique, because it focuses the attention of the student.

Introduction

People often have difficulty operating many modern consumer products such as copy machines, exercise equipment, microwave ovens, videocassette recorders, and microcomputers. Likewise, military personnel often have trouble operating state-of-the-art radios, radars, airplanes, trucks, and weapons. The program entry panel (PEP) for these complex devices "often contain internal mechanisms that obscure the relationship between the user's input and the device's behavior" (Shrager & Klahr, 1986, p. 153). They tend to be designed without adequate consideration of the user interface. Although conceptual models of how the device works may be used in the engineering of the device (Norman, 1988), engineers often give little thought to how the user will make it work (e.g., pushing buttons, flipping switches).

Computer-based training (CBT) systems have been developed to help users overcome the learning difficulties associated with operating devices with this type of digital interface. For example, CBT simulations are used extensively at many industrial and military installations (Coleman, 1988). Computers are used to teach device operation, because they provide a safe environment for users to learn how to operate equipment without endangering themselves and others or harming the equipment. Also, CBT systems can be programmed to present instruction and simulate faults. Consequently, CBT is more flexible than training on the actual equipment.

Capabilities of CBT systems include the use of sound, animated graphics and video, and the presentation of information in a linear or interactive mode. Feedback to student responses is also an important capability of CBT. In order for CBT to be effective, it must use feedback strategies that enhance learning device procedures. Feedback is important in learning how to operate a complex device, because it provides information to the student about the correctness of the student's knowledge of the device procedures.

Meta-analyses have demonstrated that computerized instruction can improve student achievement (Bangert-Drowns, Kulik, & Kulik, 1985; Kulik & Kulik, 1985; Kulik, Kulik, & Bangert-Drowns, 1985; Niemiec & Walberg, 1987; Roblyer, 1990) and that the use of feedback in computer-based instruction is more effective than the use of no feedback (Schimmel, 1983). Schimmel located 15 studies that investigated both feedback and no feedback conditions in the administering of an adult lesson involving meaningful verbal material. Schimmel reported that those receiving feedback on average scored .47 standard deviations higher than those receiving no feedback. However, none of these analyses have provided insight as to when feedback should be provided during a CBT lesson and what should be the content of the feedback. An important question is how feedback can be most effectively used in a computerized lesson on how to operate an entry panel for a digitally controlled device. The purpose of this study was to investigate the relative effectiveness of different types of feedback in learning how to operate a PEP device during CBT.

Feedback to Corrective Responses and Incorrect Responses

The purpose of providing feedback to correct answers is to help the student perceive that his or her understanding of the subject matter is correct. Kulhavy and others (Kulhavy & Anderson, 1972; Kulhavy & Parsons, 1972; Kulhavy & Swenson, 1975) have found that corrective responses were repeated whether or not feedback was provided. These findings are consistent with the results of studies (Krumboltz & Kiesler, 1965; Oppenheim, 1964) that investigated feedback to corrective responses during verbal learning. These studies found that confirmation did not affect student achievement. Steinberg (1984) believed there is no need to provide extensive feedback after a corrective response: "It serves no useful purpose, and the students will not read it" (p. 86).

In contrast, feedback to incorrect responses is highly important to learning. Buss (Buss & Buss, 1956; Buss, Weiner, & Buss, 1954) found that subjects who were informed only of their errors performed better than subjects who received feedback only to corrective responses. Although feedback to corrective responses can be helpful in some settings, this investigation focused on the effects of providing feedback to student errors.

The effectiveness of feedback to student errors during CBT may be related to when it is provided and the suitability of its content. Feedback can be provided immediately after a student response or at some later point during the CBT. The feedback content can be either the presentation of a "you are wrong" statement or the presentation of the corrective response.

Cognitive Theories of Skill Acquisition

Two theories of cognition and skill acquisition, adaptive control of thought (ACT*) theory[1] (Anderson, 1983) and

---

[1]ACT* is Anderson's notation for the last version.

instructionless learning (Shrager & Klahr, 1986) have different implications for when and how much feedback should be provided to student errors. ACT* is a theory which describes the learning of procedures. ACT* is a theory of cognition and skill acquisition based on production systems. A production system is a hierarchical set of mental tasks consisting of condition-action pairs called productions. The condition-action pairs represent specific mental and physical actions that should occur if a particular state occurs in working memory. Anderson (1983) argued that new information encoded from our environment results in very general production actions. As a task is practiced, new, task-specific productions are created for each of the conditions of the now not-so-new information which are eventually collapsed back into a single new production.

Errors may occur when the student practices the task. ACT* suggests that feedback about erroneous production *application* should be provided immediately while the production is still in working memory. Moreover, ACT* suggests that the most im͡ ͡ rtant characteristic of feedback is its function of providing the correct action to working memory. Therefore, implications of ACT* are that (1) feedback that provides the corrective response to a student error should be more effective than feedback in the form of a "you are wrong" statement, and (2) immediate feedback should be more effective than delayed feedback.

Instructionless learning (Shrager & Klahr, 1986) is a theory of discovery learning. It suggests that users can learn how to work a complex device without the benefit of any written or verbal instructions:

> In order to figure out how a device works, our subjects must discover the possible actions of the device and the range of their own behaviors that is necessary to get the device to exhibit

such actions. In the course of learning these things, they may formulate and use a theory about why their actions cause the device to do whatever it does--a conceptual model. (p. 154)

Users figure out how to work a device by discovering device behaviors as a result of user actions. During instructionless learning, the students form hypotheses about the syntax and semantics of the device switches, buttons, and dials. For example, pushing a sequence of buttons may reflect a syntax hypothesis. The student hitting the "STOP" button and the student observing that the device has paused is consistent with formation of a semantic hypothesis.

Shrager and Klahr (1986) found that instructionless learning was "fraught with error" (p.179). Subjects in the Shrager and Klahr study learned how to operate a toy tank by constantly pressing keys and observing device behavior. These subjects formulated and refined hypotheses by designing mini-experiments. Ninety-three percent of these hypothesis contained two or more key presses. Feedback was important during instructionless learning because it assisted the students in confirming or rejecting hypotheses. If a hypothesis was wrong, students changed their formulation to account for the feedback. The protocols of the subjects revealed that the feedback from the device was most useful when it occurred after the completion of all the steps of the hypothesized sequence.

Accordingly, implications of instructionless learning follow. (1) Feedback in the form of a "you are wrong" statement as a response to a student error should be more effective than feedback that provides the corrective response because it encourages the generation of new hypotheses. (2) Delayed feedback should be more effective than immediate feedback because the delay in feedback allows the subject to complete the

10

hypothesized sequence of steps. The use of delayed and confirmatory feedback allow the student to formulate and refine hypotheses about how to work the device. As a result, the student will build a schema that contains information about the syntax and semantics of various button sequences.

Method

Subjects

The subjects in this study were 53 Navy students awaiting instruction for Sonar Technician (ST) "A" School at the Fleet Antisubmarine Warfare (ASW) Training Center Pacific and 27 Navy students awaiting instruction for Apprenticeship Training at the Recruit Training Command (RTC) in San Diego. The students enrolled at the ASW Training Center (henceforth, referred to as the ASW sample) generally score between the 65th and 99th percentile on the Armed Forces Qualification Test (Monzon & Foley, 1988). The mean enlisted rank for the ASW sample was E-3 (rank ranges from E-1 to E-9) and the mean age was about 21 years. The ASW subjects had been in the Navy approximately 11 months and at the ASW Training Center about 3 1/2 months.

Students enrolled at RTC Apprenticeship Training School (henceforth, referred to as the RTC sample) generally score between the 10th and 30th percentile on the Armed Forces Qualification Test (Monzon & Foley, 1988). The mean enlisted rank for the RTC sample was E-2 and the mean age was about 21 years. These subjects had been in the Navy approximately 7 1/2 months and at the RTC about 3 1/2 months.

Procedure

The study consisted of three phases. First the subjects were required to fill out a pre-questionnaire that contained consent forms. Next, they were administered the experimental CBT. Lastly, they were required to fill out a post-questionnaire. All phases were administered individually and were self-paced.

CBT

Each subject was seated in front of an Apple Macintosh Model SE/30 microcomputer. The microcomputer presented a CBT lesson on how to operate the Consolidated Area Telephone System (CATS). The objective of the lesson was to learn how to activate and employ eight features of the Navy's CATS using the AT&T Model 2500 Telephone (AT&T, 1987). The nucleus of the lesson consisted of a computerized graphic simulation of the Model 2500 Telephone (see Figure 1). The simulation was displayed on the nine-inch (diagonal) monochrome monitor of the SE/30. The subjects interacted with the simulation by "clicking" with a mouse on the computerized graphics, listening to simulated device sounds, and observing changes to the simulation. The remaining lesson materials consisted of frames of text and graphics that were linked to the simulation. The lesson had three parts: an introduction, a practice, and a performance test.

The introduction consisted of presenting frames of information on how to use the mouse, the objective of the CBT lesson, and the locations of the Model 2500 Telephone buttons (i.e., click areas). The introduction also included a sample practice item. The introduction did not present any information on how to work the phone. Practice consisted of activating one of the features of CATS by clicking with the mouse on 16 active buttons or click areas located on the computerized graphic representation of the CATS phone (see Figure 1). For example, if instructed to "Make a Call" on the CATS
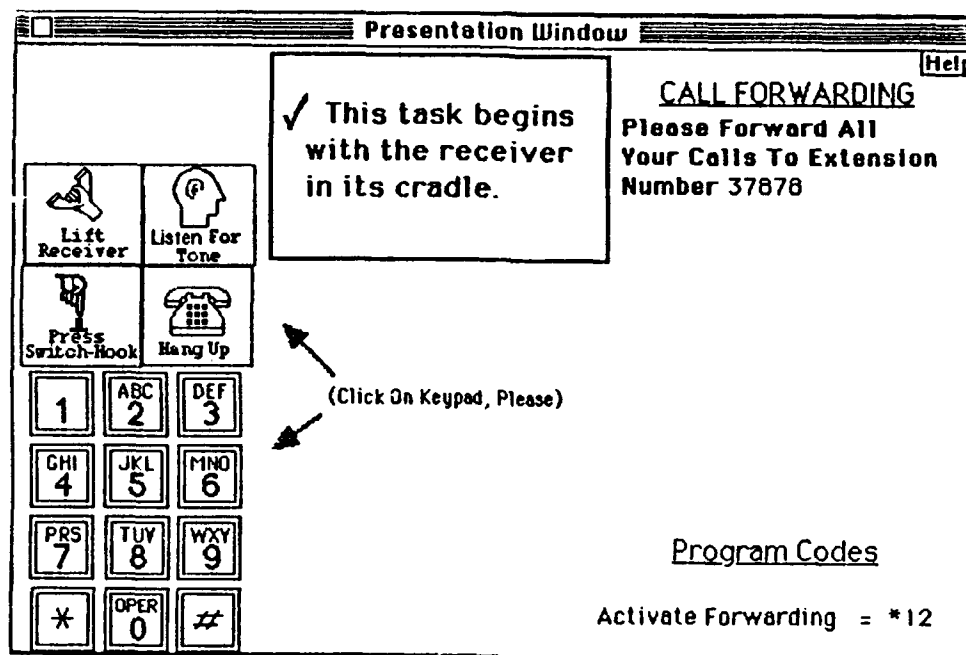
IR/IED FY90 Annual Report

11

Figure 1. Screen of AT&T model 2500 telephone and example of task.

phone, the subject would click on "Lift Receiver" button, then click on "Listen For Tone" button, and then click seven times on the number buttons. If an error was made during practice, the CBT system provided feedback.

Before the practice of each task, the subject was provided a general description of its purpose, but no information on how to do it. During practice, the information presented during the introduction could be accessed by clicking on a help button. The practice contained eight tasks (presented in the following order): Program Call Forwarding, Cancel Call Forwarding, Call Pickup, Program Abbreviated Dialing, Use Abbreviated Dialing, Call Transfer, Conference Calling, and Call Hold. The number of steps required to activate and employ each of these tasks is shown in Table 1.

The performance test consisted of performing the eight tasks contained in the

practice. There was only one performance item per task. Before each item, similar to the practice, the subject was provided a general description of the task's purpose but no information on how to do it. Each performance item required the subject to click the correct sequence and then click on the "done" button. No feedback was provided. The required steps for the performance tasks were the same as the practice tasks, but the extension numbers were different.

Post-questionnaire

The post-questionnaire consisted of questions about the subject's mood, understanding of CATS's procedures, and use of the experimental CBT. In brief, subjects were asked to list which features of CATS they felt comfortable using, and to describe the purpose of various CATS elements such as the "dial tone" and the "switch-hook." In addition, they were asked to rate the friendliness of the CBT and to comment on the

Table 1

Steps Required to Operate CATS Features

Feature	Number of Steps	Number of "Clicks"
Call Forwarding Program	7	13
Call Forwarding Cancel	5	7
Call Pickup	3	5
Abbreviated Dialing Program	11	20
Abbreviated Dialing Use	4	7
Call Transfer	4	8
Conference Calling	7	13
Call Hold	10	18

Note. Listed in CBT order of presentation.

feedback provided during CBT. They were also asked whether or not they would recommend using CBT to teach control panel procedures.

Experimental Design

Subjects were randomly assigned to the control group or to one of four treatment groups. The control subjects were administered the Pre-CBT Questionnaire, the CBT lesson introduction, the CBT performance test, and the Post-CBT Questionnaire. They were not provided with CBT practice. The treatment groups were provided all phases of the CBT. During CBT practice, these subjects received one of four types of feedback to errors: immediate confirmatory feedback, immediate corrective feedback, delayed confirmatory feedback, and delayed corrective feedback. No feedback was provided to corrective responses.

Immediate feedback was feedback provided the instant an error was made by the subject. In the immediate confirmatory feedback condition, feedback was a flashing box with the words "Incorrect... Try Again" displayed. In the immediate corrective response feedback condition, feedback was a presentation of a single correct operation. This presentation consisted of a picture of a hand pointing to the corrective response along with a flashing box with the words "Incorrect....The Hand Points to the Correct Response. (Click on the Hand.)" In these two conditions, the subjects were forced to repeat a step until it was performed correctly. After all the correct steps had been entered for that task, the subject received the message "You Have Correctly Entered the Steps For....". The subject then proceeded to the next task.

In the delayed feedback conditions, subjects performed the task by clicking sequentially on buttons and then clicking on the "done" button. Feedback was provided only after a click on the "done" button. The subjects could not move to the next task until the correct sequence of steps had been

entered without errors. In the delayed confirmatory feedback condition, feedback to student error(s) was a listing of the correctly entered steps. In this condition, subjects were shown two lists: (1) the correct steps entered in the correct order and (2) other correct steps that were not entered in the correct order. In the delayed corrective feedback condition, feedback to errors was a presentation of the entire correct sequence of steps for that task. When an error was committed, subjects were provided a demonstration that consisted of an animated hand pointing to the correct sequence of steps.

Analysis

Five subjects were excluded from the analysis for either failure to complete CBT or because of computer failures. Four of those subjects were from the delayed confirmatory conditions. One subject was from the delayed corrective response condition. The questionnaire and performance variables were analyzed using a one-way ANOVA comparing all five groups. A priori contrasts were performed for the main effects (i.e., immediate vs. delayed, confirmatory vs. corrective response, treatments vs. control). Post-comparisons among the five group means were performed using t tests. An additional measure of performance, adjusted post-test score, was generated by giving credit for responses that were nearly correct. For all test items, a response was re-scored as correct if the error involved a listening for a tone. In addition, responses to test items that apply Program Call Forwarding, Use Call Forwarding, and Program Abbreviate Dialing were re-scored as correct if the subject forgot to hang up the phone. These particular errors were re-scored as correct because if they occurred during the operation of the actual CATS equipment, these procedures might still work.

Results

Total Sample

Table 2 summarizes the biodemographic and self-rating variables from the pre-questionnaire for the total sample. No statistically significant differences among the experimental groups were found. Table 3 presents the means and standard deviations for post-test score by experimental group. These tables show a main effect for timing of feedback: the delayed groups significantly outperformed the immediate groups on the post-test score, $t(70) = 4.58, p < .01$, and on the adjusted post-test score, $t(70) = 3.65, p < .01$. Likewise, the delayed groups significantly outperformed the control group on the post-test score, $t(70) = 5.76, p < .01$, and on the adjusted post-test score, $t(70) = 7.94, p < .01$. The immediate feedback groups significantly outperformed the control group only when the post-test score was adjusted, $t(70) = 5.07, p < .01$. No effects were found for content of feedback.

Table 3 also presents the means and standard deviations for time to complete the post-test by experimental group. A main effect was again found for timing of feedback but not for content: The delayed groups took significantly less time to complete the post-test than the immediate groups, $t(70) = 3.52, p < .01$. Moreover, the delayed feedback groups used significantly less time than the control group, $t(70) = 2.22, p < .05$.

Table 4 presents the means and standard deviations for time to complete practice. In the confirmatory conditions, those in the immediate group used 19.4 compared to 42.1 minutes used by the delayed group. In the corrective response conditions, those in the immediate group used 16.4 minutes compared to 28.4 minutes used by the delayed group. The difference between the immediate and the delayed feedback groups

Table 2

Means of Biodemographic and Self-rating Variables

Variable	Confirmatory		Corrective Response		Control
	Immed	Delay	Immed	Delay	
Age	21.2	20.7	22.4	20.0	20.7
Rank	2.6	2.5	3.0	2.8	2.5
Months in the Navy	8.6	10.0	12.3	8.6	9.8
Months at Navy Training Center	4.4	2.6	3.7	3.4	3.0
Microcomputer experience	2.2	2.1	2.1	2.1	2.1
Video game experience	3.9	4.2	3.4	3.9	3.4
Mood	3.8	4.0	3.5	4.0	3.4
Times a day make a call	0.9	0.8	1.1	1.6	1.3
Times a day answer a call	2.1	1.8	1.9	4.8	4.3
Number of CATS features subject knows[a]	0.9	0.6	0.6	0.9	1.0
n	16	13	16	14	16

The header "Feedback Group" spans the Confirmatory, Corrective Response, and Control columns.

Note. Scores for those who completed the post-test (*N* = 75). Post-test was not completed for five subjects because of either computer failures or time limitations. Rank ranges from one to nine where one equals the military rank E-1 and nine equals the military rank E-9. Microcomputer and video game experience are on a five-point scale with one being equal to "none" and five being equal to "much." Mood is on a five-point scale with one being equal to "tired" and five being equal to "alert."

[a]Maximum score = 8.

IR/IED FY90 Annual Report

15

Table 3

Means and Standard Deviations of Post-test Performance

Variable	Confirmatory		Corrective Response		Control
	Immed	Delay	Immed	Delay	
Post-test score	1.4 (1.3)	3.1 (1.4)	1.4 (1.3)	2.9 (1.9)	0.6 (0.8)
Adjusted post-test score[a]	4.4 (1.9)	6.1 (1.6)	4.1 (2.2)	5.9 (2.0)	1.4 (1.3)
Post-test time	729 (172)	573 (127)	712 (223)	529 (129)	680 (226)
n	16	13	16	14	16

Note. Maximum score = 8. Standard deviations are given in parenthesis. Time in seconds.

[a]Errors involving "Listen for Tone" and hanging up receiver at end of task are ignored.

Table 4

Means and Standard Deviations of Practice Variables

Variable	Confirmatory		Corrective Response	
	Immed	Delay	Immed	Delay
Practice Time	1164 (231)	2528 (639)	982 (263)	1703 (443)
Practice Errors	44.9 (22.3)	151.5 (66.9)	25.8 (11.7)	122.2 (49.2)
n	16	13	16	14

Note. Time in seconds. Standard deviations are given in parenthesis.

was significant, $t(55) = 9.70, p < .01$. In addition, the delayed confirmatory group used significantly more time than the delayed corrective response group, $t(55) = 5.21, p < .01$. Table 4 also presents the means and standard deviations for the number of errors committed during practice. This table indicates that the delayed groups committed significantly more errors than the immediate groups, $t(55) = 9.36, p < .01$.

Generally, significant differences were not found for the open-ended questions or the self-rating items on the post-questionnaire. No significant differences were found for mood and the number of CATS features that caused subjects to feel mastery or frustration. However, significant differences were found for CBT satisfaction. The means and standard deviations for CBT satisfaction are shown in Table 5. This table indicates that the subjects in the confirmatory feedback groups enjoyed the training significantly more than the subjects in the corrective response feedback groups, $F(1,53) = 8.00, p < .01$. Yet, no

significant differences were found among the feedback groups with respect to rating how easy the CBT was to use (see Table 5).

The answers to questions about using CBT to teach how to operate a PEP device were generally positive. Approximately 93 percent recommended that CBT be used to teach control panel operation and 73 percent wrote positive comments (while only 5% wrote negative comments) to the open-ended "any comments" question. Again, no significant differences were found among the feedback groups for these measures.

Moderator Variables

Ability Level

As previously mentioned, the subjects were drawn from two populations of Navy students, those attending ASW training at the Fleet ASW Training Center Pacific and those attending apprenticeship training at the RTC. Since the ASW students generally have higher selection test scores than the RTC

Table 5

Means and Standard Deviations of CBT Satisfaction

Variable	Feedback Group			
	Confirmatory		Corrective Response	
	Immed	Delay	Immed	Delay
CBT Satisfaction	4.19 (0.65)	4.17 (0.72)	3.75 (0.77)	3.46 (0.88)
CBT Ease of Use	2.00 (1.15)	2.46 (1.45)	2.56 (1.09)	2.57 (0.94)
n	16	13	16	14

Note. Minimum rating score is one and maximum rating score is five. Standard deviations are given in parenthesis.

students, it was expected that the ASW students would perform better on the CBT tasks than the RTC students. A 2 x 2 x 2 ANOVA (sample by timing by content) was performed on the performance measures. Table 6 presents the means and standard deviations for post-test score, time to complete post-test, and practice time by sample. The ASW sample significantly outperformed the RTC sample on the post-test $F(1,51) = 16.18$, $p < .01$, used significantly less time to complete post-test, $F(1,51) = 7.64, p < .01$, and used significantly less time to complete practice, $F(1,51) = 8.39, p < .01$. However, in spite of this strong main effect for ability level (i.e., ASW sample vs. RTC sample), there were no significant interactions for sample by timing, sample by content, and sample by timing by content. Thus, ability level appears to have little impact on the relationship between type of feedback and CBT performance.

Practice Time

Since the subjects in the delayed conditions took longer to complete the practice section than the subjects in the immediate conditions, it is possible that the strong performance effects found for delayed feedback could be moderated by practice time. A 2x2 (timing by content) analysis of covariance (with practice time as the covariate) was performed on the adjusted post-test score. After accounting for the covariate, significant main effects were still found for the timing of feedback, $F(1,54) = 11.24, p < .01$, and no significant effects were found for content. The mean post-test scores and the mean post-test scores adjusted for practice time are summarized by timing of feedback in Figure 2. The pattern of results for the means adjusted by the covariate is very similar to those obtained with just mean post-test score.

Table 6

Means and Standard Deviations of Performance by Sample

Sample	Post-test Score		Post-test Time		Practice Time		n
	M	SD	M	SD	M	SD	
Immediate Conditions							
ASW students	1.7	1.3	655	161	972	206	22
RTC students	0.7	0.8	864	196	1293	235	10
Delayed Conditions							
ASW students	3.6	1.5	539	95	1996	665	19
RTC students	1.6	1.3	576	190	2347	696	8

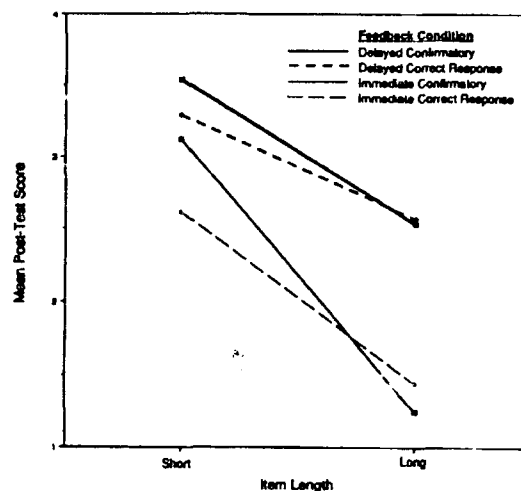*Note*. Maximum post-test score = 8. Time in seconds.
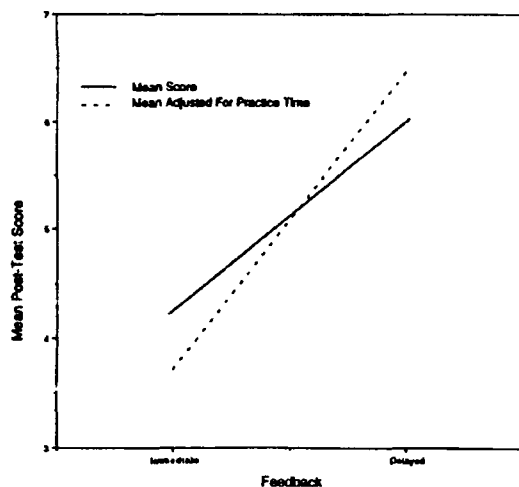
Figure 2. Mean post-test[a] score by feedback treatment

[a]Errors involving "Listen for Tone" and hanging up receiver at end of task are ignored.

Figure 3. Mean post-test[a] performance: long vs. short items

[a]Errors involving "Listen for Tone" and hanging up receiver at end of task are ignored.

Task Difficulty

The eight to-be-learned tasks presented by the CBT varied in difficulty. The length of the tasks ranged from 3 steps to 11 steps (see Table 1). The post-test, which was comprised of these eight tasks, was divided into two subscales: long items and short items. Long items had seven or more steps while short items had fewer than seven steps. Each of these subscales contained four items. A 2 x 2 x 2 ANOVA was performed. The between subject factors were feedback timing and feedback content, and the within subject factor was item length (long vs. short). Figure 3 presents the adjusted post-test mean scores for all conditions. Similar to the above results, the delayed feedback groups significantly outperformed the immediate feedback groups on both the long and short items, $F(1,55) = 22.36, p < .01$. However, a significant interaction was found between feedback timing and item length, $F(1,55) = 4.11$, $p < .05$. The advantage found for the delayed groups is much stronger for the long items than the short items.

Summary of Results

Main Effects

The results demonstrate that those provided corrective response feedback did not learn device procedures any better than those provided "you are wrong" feedback to student errors. This finding was consistent across the feedback content treatment groups. In contrast, significant differences were found for the timing of the CBT feedback. The results demonstrate that the CBT was significantly more effective when feedback to student errors was delayed until after the completion of a to-be-learned task. This effect was found when the content of the feedback was the corrective response and when the content of the feedback was an indication of error.

Covariates

Performance effects as a result of ability level, time to complete practice, and task difficulty were removed from the relationship between timing of feedback and

19

performance. For each ability level (i.e., ASW group vs. RTC group), CBT was found to be significantly more effective when feedback to student errors was delayed until after the completion of a to-be-learned task. Although the subjects in the delayed groups used significantly more time during the practice section of the CBT than the immediate groups, the above main effect was still significant after the effects of practice were removed. The correlation between practice time and performance was found to be negative and of small magnitude for both the immediate feedback subjects and the delayed feedback subjects.

The performance test was divided into long tasks and short tasks. The long tasks represented conceptually complex features that required activating more than one function. CBT was found to be more effective for delayed feedback than immediate feedback regardless of task length. However, the advantage of delayed feedback over immediate feedback was found to be much stronger for the longer tasks.

General Discussion

Instructionless learning predicts that subjects provided with confirmatory feedback will outperform those provided with corrective response feedback, and subjects provided delayed feedback will outperform those provided with immediate feedback. Little support was found for the former prediction. Performance was about the same whether feedback consisted of the corrective response or merely an indication that an error was made. However, these results indirectly support the use of confirmatory feedback over corrective response feedback. Both confirmatory feedback and corrective response feedback showed students the results of their actions and this implies that the identification of student errors is what is

central to learning procedures. This research suggests that adding the corrective response to feedback does not improve the learning of device sequences. In other words, these results provided little support for including the corrective response as part of the feedback provided to student errors.

Strong support was found for the latter prediction. CBT was found to be significantly more effective when feedback to student errors was delayed until after the completion of a to-be-learned task. The delayed feedback was found to be particularly effective in learning the features that were conceptually more complex. Delaying the feedback created an instructionless learning environment.

Shrager and Klahr (1986) found that during instructionless learning, subjects formed a conceptual model or a device schema of a PEP device. A highly developed device schema is one that contains conceptual knowledge. However, none of the data already presented provide any evidence that those in the delayed feedback groups acquired a more highly developed device schema.

Consequently, the data were reanalyzed for the presence of sequences of steps that were indicative of conceptual knowledge. The acquisition of conceptual knowledge can be inferred from steps that represent a single function. For example, the concept of programming is inferred when the "#" button follows the extension number in the Abbreviated Dialing Program task. The targeted sequences were enter program code "*17" (i.e., activate abbreviated list) followed by enter "19" (i.e, the list number) occurring in the Abbreviated Dialing Program and Abbreviated Dialing Use tasks, and enter "35656" (i.e., an extension number) followed by the "#" button occurring in the

Abbreviated Dialing Program task. The program code "*17" is conceptually related to a list number because together they represent a single programming function and the sequence is not interrupted by a confirmation tone. The most probable step following any program code is Listen For Tone which occurs 50 percent of the time for all tasks. The probability of a number following a program code is only 25 percent. The most probable steps following an extension number are Press Switch-Hook or Listen For Tone which occur about 67 percent of the time. The probability of the "#" button following an extension number is only about .16.

Table 7 presents the frequencies of corrective responses, high probability errors, and low probability errors for these conceptually related steps. Low probability errors are any errors not described above as the most probable. The table indicates that only 12.5 percent of the immediate feedback subjects remembered that step "19" followed step "*17" during the Abbreviated Dialing Program task. About 22 percent of the immediate feedback subjects correctly remembered this sequence during the Abbreviated Dialing Use task. Most of the subjects in this group followed the program code with a high probability error (i.e., Listen For Tone). In contrast, respectively over 59 percent and 70 percent of the delayed feedback subjects remembered the sequence correctly which is significantly higher than the frequency reported for the immediate feedback group, $\chi^2(2) = 14.90, p < .01, \chi^2(2) = 13.98, p < .01.$

Similar results were found for the other conceptual sequence occurring during the Abbreviated Dialing Program task. Significantly more subjects in the delayed feedback conditions remembered that the "#" button followed the extension number than in the immediate feedback

conditions, $\chi^2(2) = 12.27, p < .01.$ Again, the most likely response for the immediate feedback subjects was to commit a high probability error. Thus, subjects in the immediate conditions tended to provide the highly probable response even when it might be incorrect while subjects in the delayed conditions provided conceptually accurate responses even when it was not highly probable. These results suggest that subjects in the delayed feedback conditions acquired more conceptual knowledge about the PEP device than subjects in the immediate conditions. This analysis provides evidence that delaying the feedback to student errors during CBT practice was beneficial in the development of a usable device schema.

ACT* Theory

The predictions from ACT* theory were unsupported. ACT* suggested that feedback should be provided immediately after the occurrence of an error while a production was still in working memory. These results yielded no support for using immediate feedback. ACT* further suggested that the most important characteristic of feedback is its function of providing the correct action to working memory implying that feedback to a student error should provide the corrective response. These results provided little support for including the corrective response as part of the feedback provided to student errors. Corrective response feedback was not more beneficial in learning procedures than "you are wrong" feedback.

Current Designs of CBT Feedback

The results of the current study seem to run counter to the conventional wisdom on the role of feedback in computerized instruction promoted by many instructional design textbooks. Instructional designers generally believe that feedback in computerized instruction should be immediate. In their

Table 7

Frequencies of Corrective Responses and Errors for Conceptually Related Steps

Feedback Group	Correct	High Probability Error	Low Probability Error	n
Step *17 with Step 19 (From Abbreviated Dialing Program)				
Immediate	4	17	11	32
Delayed	16	5	6	27
Step *17 with Step 19 (From Abbreviated Dialing Use)				
Immediate	7	16	9	32
Delayed	19	5	3	27
Step 35656 with Step # (From Abbreviated Dialing Program)				
Immediate	12	13	7	32
Delayed	22	2	3	27

*Includes responses where first step is omitted.

Handbook of Computer-based Training, Dean and Whitlock (1983) stated that "once the learner has made his response, he should receive feedback on the adequacy of his performance" (p. 58). This belief is based on research in animal behavior (Skinner, 1968; Thorndike, 1911) and in programmed instruction (Bullock, 1978; Hartley, 1972; Taber, Glaser, & Schaefer, 1965). These researchers argue that immediate feedback serves to modify or maintain the learner's actions and that the delay of feedback may result in little or no learning.

However, in this study of CBT involving complex tasks, immediate feedback was found not to be better than delayed feedback. In fact, as the conceptual complexity of the CBT task increased, immediate feedback groups performed considerably worse than the delayed groups. It has been shown that immediate feedback does not seem to encourage the development of a usable device schema for complex tasks. Immediate feedback may be adequate for learning tasks where the to-be-learned sequences are short and activate one function. Yet, this form of instruction "does not result in meaningful learning, but only rote recall of information presented" (Jonassen, 1988; p. 151). Training approaches such as drill and practice that feature rote recall of information have been found not to be an effective form of instructional delivery (Kulik, Schwalb, & Kulik, 1982) and have been reported to be "dry, boring, and unpleasant" (Salisbury, 1988; p. 103).

Moreover, instructional designers also believe that content of feedback in CBT

should be the corrective response. When an answer is incorrect, feedback should not only inform learners that they are wrong but should also provide corrective information. (Steinberg, 1984; p. 86)

They maintain that computerized instruction "should do more than simply say your answer is incorrect" (Willis, 1987; p. 163). However, this study found no particular advantage in using corrective response feedback in learning how to operate a complex device and the data do not support the widely accepted belief that CBT should provide corrective feedback to student errors.

The Role of Attention

Delayed feedback created an effective learning environment for the subjects. It is conceivable that delaying the feedback was effective because it gained the learner's attention. Jacoby, Mazursky, Troutman, and Kuss (1984) found that students paid more attention to feedback which was delayed and detailed. Gaining attention has often been cited as important to learning (Gagne, 1985; Gagne & Briggs, 1979; Merrill, 1988), but little research has been done on how to gain students' attention. Gagne and Briggs argue that "gaining attention" is important to the learning process and they propose it as the first "event" when designing instructional systems.

The CBT developed here gained the student's attention by creating a puzzle-like environment. Subjects in all the conditions were asked to practice a task without first being shown how to do it. These environments tend to increase student interest in learning (Burton & Brown, 1982). Puzzles are attractive because they allow us to create order out of confusion and to discover the unknown.

The puzzle-like environment varied among the four CBT treatments. The immediate corrective response feedback condition only slightly resembled a puzzle. The CBT immediately provided the corrective response after an error. In contrast, the delayed confirmatory feedback condition was similar to a puzzle. Errors during CBT were not immediately acknowledged and the correct answers were never provided. Subjects were provided only clues on how to do the task. Subjects in this condition scored the highest on the post-test. Moreover, the subjects enjoyed the CBT more when the solutions to the tasks were not provided.

These results are consistent with the implication that delaying the feedback during the instruction gained the attention of the subjects. The delayed feedback subjects outperformed the immediate feedback subjects particularly on the more lengthy complex tasks. They also used more time during the learning phase of the CBT. It is suggested that the delayed feedback subjects solved these "puzzles" by creating and testing theories. These mental processes indeed require much focusing and concentration. Employing delayed feedback during CBT offers a challenging and exciting environment in which learners must actively respond.

Future Applications

The CBT developed in this study was found to be an effective form of teaching how to operate a digitally controlled device. This is consistent with the results found in meta-analyses of computerized instruction (Bangert-Drowns, Kulik, & Kulik, 1985; Kulik & Kulik, 1985; Kulik, Kulik, & Bangert-Drowns, 1985; Niemiec & Walberg, 1987; Roblyer, 1990). However, the instructional technologist must consider other factors besides effectiveness when designing

CBT. The CBT must also be affordable, time efficient, and user friendly. CBT should cost less than the actual equipment, use less training time than normal classroom instruction, and be easy to use. For instance, delayed confirmatory feedback during CBT may not be desired because it costs the most to develop. This strategy requires additional computer programming for the understanding of partially correct sequences. Delaying feedback during a CBT lesson may not be desirable because it might increase the length of the lesson beyond the time allowed for training. In addition, corrective response feedback must be employed carefully because it may bore some users.

These circumstances must be weighed in conjunction with performance effects when deciding what type of feedback to implement. If the to-be-learned device is simple, using feedback which consists of the corrective response or which is provided immediately will not significantly reduce the effectiveness of CBT. If CBT is being developed for a device as complex or more complex than CATS, then delaying the feedback is recommended. These findings, along with cost and feasibility data, should help the lesson developer design feedback that is sensitive to both the training requirements and the instructional setting.

References

AT&T Technologies Incorporated. (1987). *AT&T system 85 user's instructions.* Indianapolis: Author.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge: Harvard University Press.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. C. (1985). Effectiveness of computer-based education in secondary schools. *Journal of Computer-Based Instruction, 12,* 59-68.

Bullock, D. H. (1978). *Programmed instruction.* Englewood Cliffs, NJ: Educational Technologies Publications.

Burton, R. R., & Brown, J. S. (1982). An investigation of computer coaching for informal learning activities. In D. Sleeman & J. S. Brown (Ed.), *Intelligent tutoring systems* (pp. 79-98). London: Academic Press.

Buss, A. H., & Buss, E. (1956). The effect of verbal reinforcement combinations on conceptual learning. *Journal of Experimental Psychology, 52,* 283-287.

Buss, A. H., Weiner, M., & Buss, E. (1954). Stimulus generalization as a function of verbal reinforcement combinations. *Journal of Experimental Psychology, 48,* 433-436.

Coleman, D. W. (1988). *Naval computer-based instruction: Cost, implementation, and effectiveness issues* (Master's Thesis). Monterey: Naval Post-Graduate School.

Dean, D., & Whitlock, Q. (1983). *Handbook of computer-based training.* New York: Nichols.

Gagne, E. D. (1985). *The cognitive psychology of school learning.* Boston: Little-Brown.

Gagne, R. M., & Briggs, L. J. (1979). *Principles of instructional design* (2nd ed.). New York: Holt.

Hartley, J. (1972). *Strategies for programmed instruction: An educational technology.* New York: Crane, Russak, & Company.

Jacoby, J., Mazursky, D., Troutman, T., & Kuss, A. (1984). When feedback is ignored: Disutility of outcome feedback. *Journal of Applied Psychology, 69,* 531-545.

Jonassen, D. H. (1988). Integrating learning strategies into courseware to facilitate deeper processing. In D. H. Jonassen (Ed.), *Instructional designs for microcomputer courseware* (pp. 151-182). Hillsdale, NJ: Erlbaum.

Krumboltz, J. D., & Kiesler, C. A. (1965). The partial reinforcement paradigm and programmed instruction. *Journal of Programmed Instruction, 3,* 9-14.

Kulhavy, R. W., & Anderson, R. C. (1972). Delay-retention effect with multiple-choice tests. *Journal of Educational Psychology, 63,* 505-512.

Kulhavy, R. W., & Parsons, J. A. (1972). Learning-criterion error perseveration in text materials. *Journal of Educational Psychology, 63,* 81-86.

Kulhavy, R. W., & Swenson, I. (1975). Imagery instructions and the comprehension of text. *British Journal of Educational Psychology, 45,* 47-51.

Kulik, C. C., & Kulik, J. A. (1985). *Effectiveness of computer-based education in colleges.* Ann Arbor, MI: The University of Michigan, Center for Research in Learning and Teaching.

Kulik, C. C., Kulik, J. A., & Bangert-Drowns, R. L. (1985). Effectiveness of computer-based education in elementary schools. *Computers in Human Behavior, 1,* 59-74.

Kulik, C. C., Schwalb, B. J., & Kulik, J. A. (1982). Programmed instruction in

secondary education: A meta-analysis of evaluation findings. *Journal of Educational Research, 75,* 133-138.

Merrill, P. F. (1988). Applying component display theory to the design of courseware. In D. H. Jonassen (Ed.), *Instructional designs for microcomputer courseware* (pp. 61-96). Hillsdale, NJ: Erlbaum.

Monzon, R. I., & Foley, P. P. (1988). *Assessment of alternative Armed Forces Qualification Test (AFQT) composite definitions* (NPRDC-TN-88-53). San Diego: Navy Personnel Research and Development Center.

Niemiec, R., & Walberg, H. S. (1987). Comparative effects of computer-assisted instruction: A synthesis of reviews. *Journal of Educational Computing, 3,* 19-37.

Norman, D. A. (1988). *The psychology of everyday things.* New York: Basic Books.

Oppenheim, N. (1964). The relation between intelligence and difference patterns of feedback in a linear teaching program. *Columbia University. Dissertation Abstracts, 25,* 1753.

Roblyer, M. D. (1990). The impact of microcomputer-based instruction on teaching and learning: A review of recent research. *Educational Technology, 30,* 54-55.

Salisbury, D. F. (1988). Effective drill and practice strategies. In D. H. Jonassen (Ed.), *Instructional designs for microcomputer courseware* (pp. 103-124). Hillsdale, NJ: Erlbaum.

Schimmel, B. J. (1983). *A meta-analysis of feedback to learners in computerized and*

programmed instruction. Paper presented at the annual meeting of the American Educational Research Association, Montreal (ERIC Document 233 708).

Shrager, J., & Klahr, D. (1986). Instructionless learning about a complex device: The paradigm and observations. *International Journal of Man-Machine Studies, 25,* 153-189.

Skinner, B. F. (1968). *The technology of teaching.* New York: Appleton-Century-Crofts.

Steinberg, E. R. (1984). *Teaching computers to teach.* Hillsdale, NJ: Lawrence Erlbaum.

Taber, J. I., Glaser, R., & Schaefer, H. H. (1965). *Learning and programmed instruction.* Reading, MA: Addison-Wesley.

Thorndike, E. L. (1911). *Animal intelligence.* New York: Macmillan.

Willis, J. W. (1987). *Educational computing: A guide to practical applications.* Scottsdale: Gorsuch Scarisbrick.

See Appendix B-2.

Biography

Stephen W. Sorensen is an Operations Research Analyst in the Personnel Systems Department. His current research specialty is the application of artificial intelligence to problems in the management of training. He is principal investigator of a project that applies automatic discovery systems to the analysis of large personnel and training data bases.

Trained as a mathematician and physicist, he worked as a programmer at General Dynamics Corporation in Fort Worth, TX. He received his doctorate in Operations Research from the University of Texas at Austin in 1972. After 2 years as a post-doctoral fellow at Dalhousie University in Halifax, Nova Scotia, he joined NPRDC in 1974 and worked in manpower and personnel planning until 1978. From 1978 to 1983 he consulted for several companies in San Diego. He rejoined NPRDC in 1983 to work on statistical process control systems. Since 1986, he has specialized in applications of operations research to training management and implemented a planning system for Navy Enlisted Classification "C" Schools. Dr. Sorensen is a member of The Institute of Management Sciences and he has published research papers in decision analysis and machine learning.

Effectiveness of Knowledge Discovery Systems

Stephen W. Sorensen

A test was made to determine the effectiveness of a knowledge discovery system on Navy manpower, personnel, and training data. An experimental version of TETRAD II was used on a previously unstudied dataset to develop measurement models and structural models. Comparisons were made between developing the models in this way and using factor analysis to develop measurement models. EQS [1] was used as a final evaluation for measurement models as well as for structural models. We found TETRAD II useful in building sets of measurement models and in suggesting sets of possible structural models. These sets of models can be very useful for suggesting directions of further research and for eliminating areas that may not be fruitful. The test showed that a knowledge discovery system, acting without intervention by researchers, produced usable models.

Introduction

The Navy has many large data bases, each containing hundreds of data items. Often the user wants to form models from the data to improve Navy policies. The problem with applying traditional software analysis packages to very large data bases is an implicit assumption that the analyst will exhaust all models relevant to the posed problem. Unfortunately, the analysis process is very labor intensive and only a few of the models are likely to be examined. NPRDC is building a knowledge discovery system that generates meaningful models from data. NPRDC is also examining specialized discovery systems that have been built by other scientists (Callahan & Sorensen, in press). The technology will be useful not only for manpower, personnel, and training (MPT) data bases but also for any large Navy data base.

Previous work on knowledge discovery systems has never measured the effectiveness of these systems. Sometimes a researcher's result will be reanalyzed using the discovery system and reasons will be given for believing that the new analysis is better. But this approach reveals little about the strengths and weaknesses of the discovery systems. What is needed are measures of effectiveness for discovery systems that show the relative strengths of the systems and point the way for improvements.

The goal of this independent exploratory development project was to develop a measure of effectiveness for discovery systems. Since NPRDC's discovery system is not completed yet, we intended to develop a measure for another discovery system and later apply it to NPRDC's system. A 1989 independent research project successfully applied the TETRAD II program to

[1]EQS, SCALES, PARTIAL, PROMAX, and TETRAD II are names of computer programs.

discovering causal structure in non-experimental data in large Navy MPT data bases (Spirtes, Glymour, Scheines, & Sorensen, 1990). NPRDC has the TETRAD II program available and used it in the study of effectiveness.

Methods

NPRDC maintains data bases on Naval Academy admission and performance, NROTC data, officer career data, and many other issues. These data bases have been used to study many questions on Navy personnel, careers, and training. The researchers who maintain the data bases were asked to give examples of their work with the data. The intent was to use TETRAD II to analyze similar problems. Two immediate problems arose. First, many of the analyses turned out to be descriptive, consisting of means, standard deviations, and frequency tables. Complex models were not as common as expected--perhaps because of the difficulty of constructing them. Second, TETRAD II turned out to be less user friendly than expected and it became clear that a major effort would be required to apply it.

The decision was made to scale back the project's scope and analyze one problem in depth. Rather than quantitative measures of effectiveness, we decided to make a qualitative judgment of the effectiveness of the system on the single problem. As a secondary goal, we wanted to prove the concept of discovery systems: We wanted the computer to build a model from beginning to end with no human intervention.

TETRAD II was designed as an aid in elaborating or respecifying structural equation models that had been developed through traditional research and statistical methods (Glymour, Scheines, Spirtes, & Kelly, 1987). TETRAD II's elaborator performs well in simulations, but the elaborator assumes that a

body of knowledge exists about the problem being studied and that this body of knowledge has been summarized in a tentative structural equation model. The elaborator then finds alternate sets of structural equation models that satisfy the correlation constraints of the data and the model.

We used two experimental modules of TETRAD II to build structural equation models in a situation having no body of knowledge and no tentative structural equation model. We found TETRAD II helpful in discovering measurement models and in building sets of structural equation models.

For this investigation, we used a large dataset with which we were unfamiliar; we developed measurement models for latent variables and then built structural models detecting the causal relations among such latent variables. Measurement models were developed using factor analyses and TETRAD II to cluster variables and then using TETRAD II to select cluster subsets that form pure measurement models. Structural models were built using TETRAD II. We evaluated all the measurement models and structural models using EQS (Bentler, 1985).

The newly developed and very experimental SCALES module of TETRAD II was used to form pure measurement models. A pure measurement model is one in which the correlation between the indicator variables is solely due to the common effects of a single latent variable. Given a group of possible indicators of one latent variable, SCALES searches for five-variable subsets that are pure measurement models for that latent variable. High correlations are expected among the variables in this group; variables with zero correlations are eliminated.

All tetrad equations among the indicators of a single latent variable are implied by

29

a pure measurement model. SCALES judges the pureness of a measurement model by evaluating the 15 possible tetrad equations from each fivesome. On the assumption that each equation holds in the population, SCALES calculates the probability that the residual from the tetrad equation is as large as or larger than, the one observed. The minimum probability of the 15 and the average of the 15 probabilities are printed out. Along with these two probabilities, SCALES also prints out the proportion of tetrad equations that can only be explained by a latent variable. A tetrad equation can only be explained by a ...tent variable if there does not exist another measured variable, say m, such that all partial correlations, including pairs in the tetrad equation, partialled on m are zero. Thus, SCALES prints out the above three assessments for every possible fivesome. The same assessments are then calculated and printed for all possible foursomes.

Figure 1 contains a flow diagram of our procedure for developing a measurement model. We first clustered variables as indicators of some latent variable. We tried two methods to arrive at clusters. The first method was a relative-ignorance equivalent to substantive knowledge. We merely grouped questions that addressed the same issue. The second method was factor analysis. We chose PROMAX rotations because these rotations made the most sense compared with other rotations. The biggest problem with the factor analysis measurement models was that after the fourth factor there were only two or three questions in each factor. These factors could not be evaluated by SCALES, because SCALES requires at least five variables. Nor could these factors be evaluated by EQS, because with only two or three questions the models are under-identified. As discussed below, a third technique is available for evaluating measurement models; a TETRAD II module called PARTIAL. However, we had
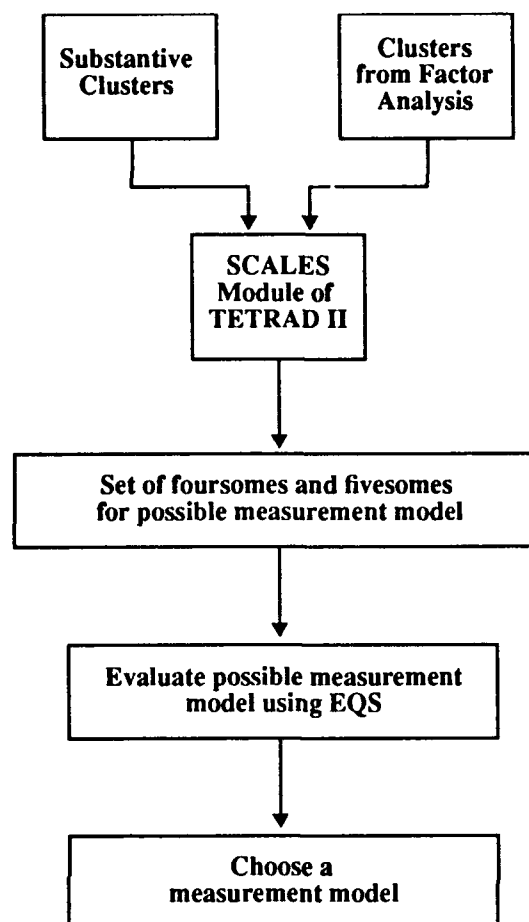
Figure 1. Flow chart for developing a measurement model.

too many variables to use PARTIAL. As a result, those factors had to be dropped from further analysis. All factor analyses were performed using a statistical analysis system on an IBM 4381.

After giving SCALES a cluster of variables and receiving sets of foursomes and fivesomes, we ranked these sets according to the sum of the three assessments discussed above. The top 20 sets (many latent variables did not have that many) were then evaluated by EQS. SCALES automatically generates EQS code. For each latent variable, we selected as the measurement model the foursome or fivesome with the highest p-value for

the Chi-square goodness-of-fit test from EQS.

After a set of scales had been developed, we used TETRAD II to develop a set of structural models, or directed graphs, among the latent variables. We first estimated the correlations among the latent variables with EQS and then treated them as measures for which path models were constructed with another experimental module of TETRAD II, PARTIAL.

The PARTIAL module of TETRAD II eliminates causal connections between two variables by using partial correlations as statistical tests for conditional independence. If two variables are independent conditional on some other set of variables, then the two variables are not directly causally connected. Thus, PARTIAL produces a set of possible structural models. We evaluated all these structural models for goodness of fit with EQS.

The complete details of our analyses are available in another paper (Callahan & Sorensen, submitted).

Results

The data we used consisted of responses to a questionnaire known as the Youth Attitude Survey (YATS). Each year this questionnaire is given to about 10,000 young persons between the ages of 18 and 25. These individuals are tracked to see if they enlist in the military. The purposes of the questionnaire are to evaluate the effectiveness of recruitment advertising and to look for any enlistee identifying characteristics. We used the responses from the 1985 questionnaire because 5 years had elapsed during which respondents could have enlisted.

Many of the questions in YATS can be answered yes or no. Because of this, for

TETRAD II, we derived further variables that approximate continuous measures. For example, a series of seven questions ask if the respondent has seen advertisements for a given branch of the military. We summed the number of positive responses deriving a variable for the number of military branch advertisements the respondent had seen.

We were primarily interested in identifying any differences between those who enlisted and those who did not. Rather than include this binary variable in any model, we divided the responses into two datasets, military and civilian. Each dataset was analyzed separately. There were 9,086 civilian respondents (i.e., respondents with no enlistment to date) and 854 individuals who had enlisted.

Most of our efforts were directed at the measurement models. This is reasonable, because we do not know which latent variables, if any, were in the mind of the questionnaire designer. We had 10 to 15 possible latent variables, two methods of generating measurement models, and two datasets.

We first grouped variables along common-sense lines. Table 1 contains the questions that we believed addressed "how likely a respondent felt he or she was to enlist in some branch of the military in the future" (abbreviated to Likely Military). TETRAD II's SCALES module on this group produced a list of possible combinations of variables in sets of four or five. As discussed above, we automatically ordered the list according to the sum of the three assessment probabilities calculated by TETRAD II. We tried other algorithms for selecting the measurement models to evaluate: The sum always captured models with at least one large probability. It is desirable to have all three of these statistics equal to one.

For the top 20 subgroups, we had TETRAD II generate EQS input files to

31

Table 1

Cluster of Questions for the "Likely Military" Latent Variable

Variable	Question
Q505	How likely is it that you will be serving in the National Guard?
Q507	How likely is it that you will be serving in the Reserves?
Q509	How likely is it that you will be serving on active duty in the Coast Guard?
Q510	How likely is it that you will be serving on active duty in the Army?
Q511	How likely is it that you will be serving on active duty in the Air Force?
Q512	How likely is it that you will be serving on active duty in the Marine Corps?
Q513	How likely is it that you will be serving on active duty in the Navy?
Q514	How likely is it that you will be going to college?
Q515	How likely is it that you will be going to vocational or technical school?

evaluate each measurement model. Table 2 contains a part of this list for the Likely Military latent variable with the civilian respondents. The second through fourth columns contain the TETRAD II assessment probabilities. The rightmost column contains the p-level from the goodness-of-fit Chi-square test from EQS.

The seven tentative measurement models in Table 2 were the best three models of fivesomes and the best four models of four-somes. Here "best" means the highest values for the sum of the three TETRAD II probabilities. Notice that the goodness-of-fit statistics agree in general with the TETRAD II statistics. In almost all cases, when looking at the model with the highest sum of the three TETRAD II assessments, it also had the highest p-value from EQS.

Table 3 contains a listing of the clusters of questions for each latent variable defined by common sense; capital letters define abbreviations used in figures. The same clusters were used for civilian respondents and for military respondents. When we applied PROMAX to the data, the first factors were the same for civilian and military respondents

and agreed well with the Likely Military clustering. Of course, we chose the PROMAX rotations because they produced clusters that made sense to us. Many factors from PROMAX did not contain enough variables for evaluation with the TETRAD II SCALES module (at least 5).

Occasionally TETRAD II was unable to find any subset of indicators that could form a pure measurement model for some latent variable. This occurred a few times in the common-sense clusters, but was a recurring problem with the factor analysis groupings. When this occurred with a common-sense cluster, we enlarged the cluster until TETRAD II was successful in finding subsets that could form a pure measurement model. Thus, some variables occurred in more than one cluster. However, we were careful never to include one variable as an indicator for more than one latent variable. In fact, the best fitting measurement models did not contain any overlapping variables. With the factor analyses, we tried unsuccessfully to combine factors.

To build a structural equation model, we selected the measurement models with the

Table 2

SCALES Selections for "Likely Military" Latent Variable for Civilian Respondents

Variables	Average Prob.	Minimum Prob.	Prob. Needs Latent	EQS P (Chi SQ)
Q507 Q509 Q510 Q511 Q515	0.633	0.000	0.87	0.0048
Q507 Q509 Q511 Q512 Q515	0.340	0.011	1.00	0.0305
Q507 Q510 Q511 Q513 Q515	0.377	0.000	0.80	<.001
Q507 Q509 Q511 Q515	0.835	0.750	1.00	0.9485
Q507 Q510 Q511 Q515	0.704	0.570	1.00	0.8443
Q509 Q510 Q511 Q515	0.688	0.532	1.00	0.8208
Q507 Q509 Q510 Q515	0.679	0.537	1.00	0.8469

Table 3

Common-sense Clusters for Latent Variables

Latent Variable	Variables
Demographics	Q402 Q403 Q404 SUM405 Q407 Q410A Q411 Q415 Q693 Q694 Q713F Q713M Q714 Q715
CURent JOB	Q416 Q424 Q425 Q430 Q431 Q436 Q427
LIKELY military	Q505 Q507 Q509 Q510- Q511 Q512 Q513 Q514 Q515
FUTure MILitary plans	Q503 Q522 Q523 Q524 Q554 SUM644 Q698 Q700 Q701 SUM702 CPYATS RSVING SALARY YEARLY
FRIENDS support military	Q682 SUM 684 Q687 Q690 Q691 Q692
Military advertising	Q503 Q504 Q517 SUM438 Q523 Q524 SUM601 SUM602 SUM617 SUM644 Q698 SUM702 CIVJOB CPYATS MILJOB

largest p-values for the goodness-of-fit Chi-square. We used EQS to estimate the correlation structure among the latent variables and then used this correlation matrix in TETRAD II to find tentative structural models. In all cases, the PARTIAL module provided 3 to 20 models. Most of these models differed only in the causal direction between one or more pairs of latent variables. Some latent variables were never causally connected to others in the structural models that PARTIAL suggested. For example,

neither the Demographic nor the Military Advertising latent variables were included in any suggested structural models.

Lastly, we used EQS to evaluate each of the TETRAD II suggested structural models. None of the structural models using factor analysis measurement models fit well: The highest p-value for the Chi-square goodness-of-fit was .0201 and most p-values were <.001. Similarly, the structural models for military enlistees using substantive

groupings for measurement models all had p-values <.001.

However, for the civilians using substantive groupings for the measurement models, the process found consistent and revealing models. Starting with seven latent variables (see Table 3), TETRAD II suggested 20 possible structural models among five latent variables, dropping two latent variables (Demographics and Military Advertising).

Four structural models had Chi-square goodness-of-fit p-values of 0.77. The number of observations for this model was 7625. This is lower than the number of civilian respondents because of missing data. Figure 2 contains a directed graph of one of the models. The latent variables have been abbreviated using only the capital letters in Table 3. All four models contained the edges drawn with heavy lines with coefficients as indicated. The other three models had one of the three edges drawn with a light line

pointed in the other direction. Two models with two of the three light-lined edges in the opposite directions had p-values near 0.4. Structural models with any of the heavy-lined edges pointing in the opposite direction had very low p-values, $< =.05$. Half of the models had an edge from FUTMIL to LIKELY. We were unable to get EQS to converge for any of these models except one that had a p-value <.001.

Many of the heavy-lined edges make sense. An individual's current job impacts on that individual's thoughts about his or her future job and how the individuals feel about possible future non-military jobs is causally connected to how they feel about a future in the military. The negative coefficient for that edge means that the more positive individuals are about a future non-military job the less they see a future in the military. Similarly, if friends or relatives are positive about the military, or if the friends or relatives think positively about a future military career, then
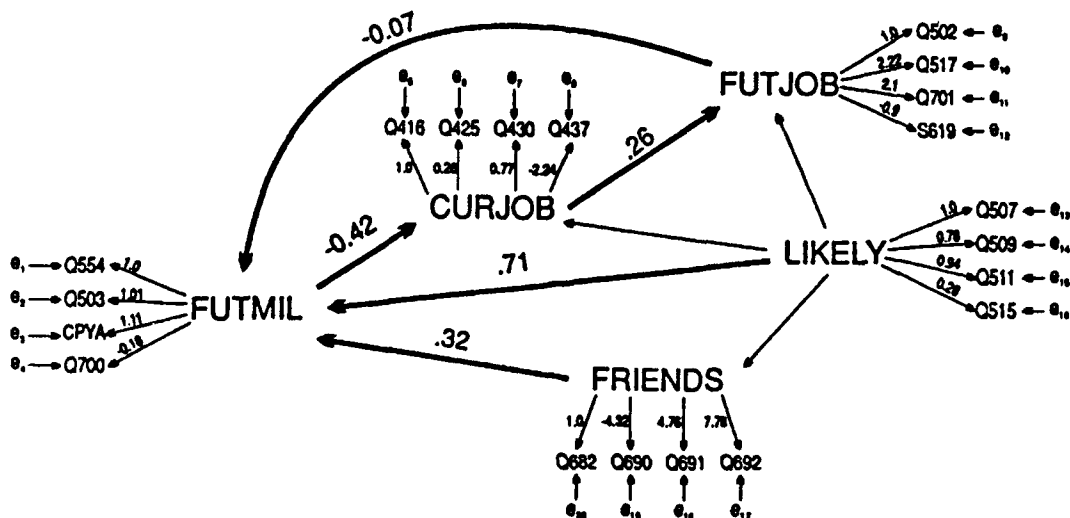

Figure 2. Civilian structural model, n=7625, p=.77.

34

the respondents are more likely to see a future for themselves in the military.

Conclusions

Building a structural equation model is a difficult process. However, researchers must start somewhere. We found TETRAD II, although still very experimental, a useful method for uncovering and evaluating tentative measurement models and for building and eliminating structural equation models on the latent variables. We began with a new data base and were able to generate structural equation models that fit remarkably well. If we had been in a research environment, we would have had a number of hypothetical models for future study.

TETRAD II is very helpful in indicating what does not work. Variables that never appear in a foursome or fivesome are not useful as an indicator for that latent variable. Also, if a latent variable does not appear in any suggested structural equation model, that latent variable is not useful.

We felt more comfortable using TETRAD II on common-sense clusters of data than using factor analyses. We found ourselves rotating until we found a set of factors that made sense to us. Even then, there were a few factors whose combination of questions made no sense and we questioned what the latent variable could possible be. Many factors contained two or three variables, which made evaluation of the measurement model with SCALES or EQS impossible.

We were impressed with the quality of the model that TETRAD II produced (Figure 2). We had hoped for a similar structural model for military, but perhaps the data does not support such a model. Could a human researcher have found a model as good as that found by the computer? We don't know. We

know the computer found the model and that the model is excellent. Although we had to write computer programs to link TETRAD II and EQS, the computer made its own decisions as it searched for the model. The computer acted solely from the evaluation functions that we provided. The value of knowledge discovery systems on Navy MPT data is clearly established.

References

Bentler, P. (1985). *Theory and implementation of EQS: A structural equations program*. Manual and documentation for a computer program available from BMDP Statistical Software, Inc., 1440 Sepulveda Blvd., #316, Los Angeles, CA 90025.

Callahan, J., & Sorensen, S. (in press). Rule induction for group decisions with statistical data--an example. *Journal of the Operational Research Society*.

Callahan, J., & Sorensen, S. (submitted). Using TETRAD II as an exploratory tool. *Social Science Computer Reviews*.

Glymour, C., Scheines R., Spirtes, P., & Kelly, K. (1987). *Discovering causal structure: Artificial intelligence, philosophy of science, and statistical modeling*. San Diego: Academic Press.

Spirtes, P., Glymour, C., Scheines, R., & Sorensen, S. (1990). *Using artificial intelligence to aid in the development of causal models*. Final report to Manpower, Personnel, and Training R&D Program, Office of Naval Research, and to the Navy Personnel Research and Development Center (Contract No. N00014-89-J-1964).

See Appendix B-3.

Progress Reports

Biography

David L. Ryan-Jones is a Personnel Research Psychologist in the Neurosciences Division. He received a B.S. (1973) and M.S. (1975) in Experimental Psychology from the University of Texas, El Paso, and a Ph.D. (1988) from Florida State University in Experimental Psychology (Psychobiology). Dr. Ryan-Jones is a member of the American Psychological Association, Human Factors Society, and the Association for Research in Vision and Ophthalmology. His research interests are in the electrophysiological correlates of perceptual and cognitive information processing.

Gregory W. Lewis was born, raised, and educated in ▬▬▬▬▬▬▬. During his graduate work at Washington State University, he had extensive training in vision electrophysiology and neurophysiology. His doctoral dissertation was in the area of vision biometry using ophthalmic ultrasonography. From 1970 to 1974, Dr. Lewis fulfilled his military obligation as an Army officer in the U.S. Army Medical Research Laboratory, Fort Knox, Kentucky. He has been with NPRDC since 1974. He pioneered neurosciences research at NPRDC and currently heads the Neurosciences Division within the Training Systems Department. His research interests include the psychophysiology of individual differences, digital processing of biological signals, and physiological correlates of brain and behavior.

Individual Variations in Event-related Brain Potentials

David L. Ryan-Jones
Gregory W. Lewis

Recent research has shown that cortical event-related potentials (ERPs) provide an accurate record of the spatial and temporal distribution of processing within the brain during a cognitive task. One problem which currently limits the usefulness of this technique is that the signal generated by external stimuli is small compared to the background activity of the brain. The purpose of this study was to estimate the minimum number of trials necessary to derive a representative average ERP. The results showed that while there were large individual differences in signal-to-noise ratio, post-stimulus variability, and shape of the ERP waveform, as few as 10 trials produced an ERP that was very similar to one derived from a much larger number of epochs.

Background and Problem

One important use of the analysis of cortical event-related potentials (ERPs) is the prediction of future performance from the patterns of potentials elicited during a predictor task. One problem, which currently limits the usefulness of this technique, is the signals generated in the cerebral cortex of the brain by external stimuli are small compared to the background activity of the brain. One technique that is most commonly used to improve the event-related potential signal strength is to average 100+ samples of brain activity. However, there is no practical guidance available to help estimate the number of epochs that need to be averaged to derive an ERP which is representative of the actual response elicited by the stimulus. One common strategy is to collect and average the largest number of trials that can be obtained during the course of an experiment. This may not be the best strategy to use in an applied setting if subject availability is limited or task

demands prevent extended use of the subject in each experimental condition.

Purpose and Findings

The purpose of this study was to determine the degree to which individual variation in brain activity in a representative sample of military subjects affects the minimum number of trials necessary to derive a representative ERP. ERP data were collected from 130 Marines during a bimodal discrimination task. Each subject was presented with a randomized sequence of visual and auditory stimuli consisting of a checkerboard pattern alone or a checkerboard pattern presented with a tone. The task of the subject was to press one of two keys indicating whether or not the checkerboard pattern was accompanied by the tone. The data showed that there were important individual variations in signal-to-noise ratio, post-stimulus variability, and waveform shape as a function of the number of trials in the average and the

40

time on task. In spite of these individual differences, several different analyses (component analysis, spectral analysis, variability analysis, and correlational analysis) consistently suggested that an average consisting of as few as 10-20 epochs produces an ERP that is very similar to one derived from 90 or more epochs. The conclusion of this study is that an average of only 10-20 epochs may capture most of the information that is available in an average of several hundred epochs.

References

See Appendix B-2.

Biography

John A. Ellis is a Senior Research Psychologist in the Training Technology Department. He has been involved in research and development programs dealing with quality control of instructional development, criterion reference testing, cognition and instruction, computer-based training, and instructional systems design. He received his doctorate in Psychology from the University of Illinois, Champaign-Urbana in 1976. He has authored over 100 technical reports and professional publications. His current research interests are in instructional development and techniques for enhancing the retention of procedural tasks. Dr. Ellis is a member of the American Educational Research Association and the Steering Committee of the Military Testing Association. He is currently a consulting editor for the *Journal of Educational Psychology* and is a peer review advisor for the Office for Educational Research and Improvement of the U.S. Department of Education.

William E. Montague is a Senior Scientist in the Training Technology Department. His research specialty is cognition and learning. For several years he directed projects developing improvements of instructional design methods and using computers for training. Trained as an experimental psychologist at the University of Virginia, he did research in human factors for the Navy Electronics Laboratory, taught Psychology and Educational Psychology at the University of Illinois, and moved to NPRDC in 1972 as a project leader. He is an active member of several professional organizations including: American Educational Research Association, Cognitive Science Society, Psychonomic Society, American Psychological Association, Human Factors Society, and Military Testing Association. He has authored or co-authored over 100 professional and technical papers, and has co-edited three books concerned with instructional psychology. He is currently a consulting editor for the *Journal of Educational Psychology*, the *Journal of Applied Psychology*, the *Human Factors Journal*, and is a peer review advisor for the Office for Educational Research and Improvement of the U.S. Department of Education.

Tutoring/On-the-job Technical Training

John A. Ellis
William E. Montague

In addition to the more than 7,000 formal courses taught in Navy schools, there is a considerable amount of training conducted on-the-job in ship- and shore-based commands. Much of this training occurs informally. Although the Navy has courses and programs that prepare petty officers to be leaders (e.g., Leadership Management Education and Training or LMET), the only course that addresses shipboard training emphasizes classroom training aboard ship and is given to very few petty officers per year (less than 600). This project will provide information for designing and developing a formal program for teaching senior Navy petty officers to be effective on-the-job trainers/tutors. Data Collection aboard ship has been completed and the data are being analyzed for project documentation.

Background and Problem

In peace time, the Navy is heavily involved in training. This is especially true for new job incumbents and for those in jobs that change frequently or are difficult to master (e.g., the tasks are complex, there are infrequent opportunities for practice, etc.) Much of this training occurs informally in one-on-one or one-on-two or three situations, with a senior petty officer (e.g., E-6, E-7) working with/teaching seaman and seaman apprentice personnel on/about shipboard tasks. These senior petty officers are in effect tutors and are responsible for bringing "A" school (and non "A" school) graduates from a novice status to a journeyman. This involves preparing them to take and pass advancement exams, meet personnel qualification standards (PQS) and practical factor requirements, and perform their jobs. Although the Navy has courses and programs that prepare petty officers to be leaders (e.g., Leadership Management Education and Training or LMET), there is only one short (5 day) course that addresses training aboard ship and very few petty officers have the opportunity to take it.

Objective

The objective of this project is to do the basic research required to provide information for designing and developing a formal program for teaching senior Navy petty officers to be effective on-the-job trainers/tutors.

Approach

The project consists of four phases: (1) analysis of tutoring, (2) developing a data collection methodology, (3) data collection on shipboard personnel, and (4) analysis and recommendations.

Phase 1 involved an analysis of tutoring to determine the factors involved in tutoring and the characteristics of good tutors. Several researchers are currently investigating these

issues (Fox, 1988a, 1988b, 1988c; Gordon, 1988) with tutors in college subjects. Phase 1 extended this work to technical training.

In Phase 2, a data collection methodology was developed for assessing tutorial skills and knowledge. The development process involved field observations and resulted in a paper-and-pencil survey administered to fleet enlisted supervisory personnel.

The questionnaire is divided into four sections. The first concerns information about personnel the respondant supervises (e.g., number, rating, "A" school graduates). The second section asks questions about the respondant's Navy training background and his opinion of the training his/her personnel have received. The third section deals with shipboard tutoring issues, and the fourth asks for demographic information on the respondant.

In Phase 3, data were collected by mailing the questionnaire to petty officers aboard ship. The questionnaire was analyzed in Phase 4 and recommendations was made for a formal training program in tutoring for senior petty officers and for modifications in instructor training to enhance tutoring skills.

Progress

Phase 1, 2, and 3 have been completed.

Plans

Phase 4 will be completed in FY91.

References

Fox, B. A. (1988a). *Cognitive and interactional aspects of correction in tutoring*. (NPRDC-TR-88-2). Boulder, CO: University of Colorado, Institute of Cognitive Science.

Fox, B. A. (1988b). *Interaction as a diagnostic resource in tutoring*. (NPRDC-TR-88-3). Boulder, CO: University of Colorado, Institute of Cognitive Science.

Fox, B. A. (1988c). *Repair as a factor in interface design*. (NPRDC-TR-88-4). Boulder, CO: University of Colorado, Institute of Cognitive Science.

Gordon, E. E. (1988). *Description of a field tutoring project and IIP curriculum*. Paper presented at the 1988 annual meeting of the American Educational Research Association, New Orleans, LA.

Biography

Jan Dickieson directs the Instructional Technology Division within the Training Technology Department. Projects in her division deal with incorporating current and emerging technologies into the automation of training development and delivery. Ms. Dickieson received a B.S. in Computer Science from the University of Nebraska in 1972 and an MBA in Management from San Diego State University in 1978. Her present interests include areas of automated curriculum design and development, applications of hyper-media to instructional delivery, and explorations into the efficacy of applying artificial neural network technology to issues in training. Ms. Dickieson is a member of the American Educational Research Association.

Lew Gollub is a [part-time] Personnel Research Psychologist in the Training Technology Department and a Professor of Psychology at the University of Maryland at College Park. His research has contributed to the quantitative analysis of the effects of reinforcement schedules on learned behavior, conditioned reinforcement, and the effects of psycho-active drugs on behavior. Recently, he has begun analyzing performance acquisition using neural network modeling techniques. Dr. Gollub received a B.A. in Psychology and Mathematics from the University of Pennsylvania, and a Ph.D. in Psychology from Harvard University. He joined NPRDC as a U.S. Navy ASEE Summer Faculty Fellow in 1989 and 1990. He is a Fellow of the American Psychological Association and the American Association for the Advancement of Science. He is also a member of the Association for Behavior Analysis and the International Neural Network Society. He has served as associate editor of the *Journal of the Experimental Analysis of Behavior* and reviewing editor for numerous other journals. He is the author of over 40 contributions published in scientific journals and books as well as a number of technical reports.

Artificial Neural Networks and Training

Jan Dickieson
Lew Gollub

The development of effective training methods requires predicting how the trainee will respond to the training procedure. At present, only qualitative models and intuitions can guide the training developer. A quantitative model of human behavior, as it changes through training, would facilitate the development of optimal training methods and materials by providing a platform for rapidly pretesting procedures prior to a more costly field test. This project will use neural network analysis techniques to develop a model of the acquisition of some aspects of a Navy training task (Air Intercept Controller). This model will then serve as the basis for predicting the effects of changes in training conditions.

Rationale

The primary goal of this research project is to develop an artificial neural network (ANN) to model human performance in a Navy relevant task. The model can then be used to predict the effects of changes in training conditions with the long-term goal of developing a model that can predict the effects of novel combinations of training parameters. Target training situations are high-performance tasks that require rapid response to complex, demanding environmental situations, such as Air Intercept Control (AIC), Air Traffic Control, or Air Combat Maneuvering (ACM).

Artificial neural networks are a recently developed method for modeling complex systems. A great variety of processes have already been modeled with this approach, including perceptual systems for detecting sound navigation and ranging (SONAR) profiles, entrance criteria for U.S. Naval Academy applicants, and financial characteristics of good and bad loan applicants. The diversity of applications indicates the flexibility of the approach.

In essence, an ANN is a collection of highly interconnected processing elements similar to the neurons of the nervous system. A specific ANN model consists of a certain structure of interconnected elements and rules governing how the model changes. The influence of input information (similar to sensory neurons) on interconnecting ("hidden") elements and on output ("motor") elements is adjusted according to learning rules.

Although there are a number of individual findings that describe the acquisition of high-performance skills, no successful unified model of learning has been developed for this important class of activities. Since these situations involve changes in response with repeated presentation of learning scenarios, ANNs offer a plausible modeling approach because they permit a broad range of model structures and learning rules.

Relatively few studies have been reported in which ANNs have been used to model human or other complex behavior (Fix, 1990). However, successful attempts have modeled ACM decision making (Crowe, 1990; McMahon, 1990) as well as simple discrimination learning in laboratory animals (Commons, Grossberg, Staddon, in press). Thus, it is plausible to apply ANNs to the training environment.

Approach

The Navy relevant skill that will be studied is the AIC situation. In this task, a trainee examines the Navy Tactical Data System (NTDS) display screen, issuing commands to a controlled aircraft to intercept (or avoid) a target aircraft. This task involves a series of perceptual and fine motor response requirements as well as complex decision-making. This task demands a variety of skills whose acquisition can be studied under different training conditions.

The research plan involves three phases, which will be done in sequence. The first phase will be the development of a simulation of the AIC task on personal computers (PC), so that studies can be performed on a variety of computer equipment. The second phase involves collecting training data on simple tasks in which stimulus and response conditions can be recorded in detail by the simulation program. The third phase is selecting and developing a neural network model involving stimulus and time (training) inputs for subject behaviors.

Progress

We will modify the NTDS/PC system software that we have obtained from Walter Schneider of University of Pittsburgh. Considerable recoding must be done to permit us to perform the acquisition studies and detailed recording that we plan. A student contract programmer has been hired and has begun the task of analyzing and recoding the program. This phase of the research should be completed during the third quarter of FY91.

We will continue to review the published literature for applications of ANNs to learning and performance.

References

Commons, M. L., Grossberg, S., & Staddon, J. E. R. (Eds.) (in press). Quantitative analysis of behavior, *Volume XII: Neural network models of conditioning and action.* Hillsdale: Lawrence Erlbaum.

Crowe, M. X. (1990). *The application of artificial neural systems to the training of air combat decision-making skills.* Paper presented at the 12th Interservice/Industry Training Systems Conference, Orlando, FL.

Fix, E. (1990). *Modeling human performance with neural networks.* Proceedings JINNS Meeting, San Diego, I-127-I-252.

McMahon, D. C. (1990). *A neural network trained to select aircraft maneuvers during air combat: A comparison of network and rule-based performance.* Proceedings JINNS Meeting, San Diego, I-107-I-112.

Biography

Charles Wilkins is a Personnel Research Psychologist in the Testing Systems Department. He earned a Ph.D. in Quantitative Psychology at the Johns Hopkins University in 1989. He began working at NPRDC in July 1989 and is involved in psychometric research for the Computerized Adaptive Testing-Armed Services Vocational Aptitude Battery (CAT-ASVAB) Program. He recently began working on the application of Artificial Neural Network technology to personnel selection problems.

Jan Dickieson directs the Instructional Technology Division within the Training Technology Department. Projects in her division deal with incorporating current and emerging technologies into the automation of training development and delivery. Ms. Dickieson received a B.S. in Computer Science from the University of Nebraska in 1972 and an MBA in Management from San Diego State University in 1978. Her present interests include areas of automated curriculum design and development, applications of hyper-media to instructional delivery and explorations into the efficacy of applying artificial neural network technology to issues in training. Ms. Dickieson is a member of the American Educational Research Association.

An Exploratory Examination of Neural Networks as an Alternative to Regression

Charles Wilkins
Jan Dickieson

Accurate prediction of the future performance of personnel is a vital piece of information that the Navy uses to make appropriate decisions. These predictions are most commonly made using some type of linear regression-based method. This means that not all available information (e.g., nonlinear relationships) will be used in making the prediction. Neural network technology, on the other hand, allows prediction models to be created that do take into account nonlinear as well as linear relationships, This study was designed primarily to explore the feasibility of applying neural network technology to an important Navy problem, specifically, predicting premature voluntary attrition from the U.S. Naval Academy. Data will be used from various classes of the Naval Academy and premature attrition will be forecasted from various predictors, using both models. Then, the relative predictive efficacy of the two approaches will be compared.

Background and Problem

To a large extent, the Navy must make significant decisions about a variety of personnel issues based upon the prediction of future performance. To the degree that these predictions are suboptimal, the results of these decisions will have an adverse impact. The current technique used to predict performance is multiple linear regression. However, to the degree that the relationships include unspecified nonlinear components, the data being studied cannot be accounted for by linear regression-based procedures. Nonlinear regression techniques exist; however, most of these require a priori specification of a model of the nonlinear relationship. Unless the form of the nonlinearity is well understood, it is difficult to choose such a nonlinear model.

In recent years, new techniques have been developed in which computers "learn" the relationships between a set of variables without needing a priori information about the relationships. These techniques, called neural networks (NNs), have been used to model both human performance (e.g. cognition, vision, speech, etc.) and to improve computer performance of skilled tasks (e.g., pattern recognition, robotics, etc.) (Rumelhart & McClelland, 1988).

Far less research has been conducted on the efficacy of using neural networks for prediction. The formal properties of NNs are complex and are still being studied by mathematicians. Despite this complexity, some theoretical properties have been expounded. Theoretically, a neural network can be designed which can model any linear or nonlinear continuous function (Hecht-Neilsen, 1987; Hornik, Stinchcombe, & White, 1989). However, neural networks are iterative techniques and many questions

remain about how to best train the network and what the optimal configuration of the network should be. There are many unresolved questions about the predictive efficacy of neural networks in practice.

Objective

The objective of this work is to explore, empirically, the advantages and drawbacks of using neural network technology as an alternative to linear regression techniques for various prediction problems of concern to the Navy. In particular, this study expects to show that, if used properly, neural networks have the capability to predict performance as well or better than linear regression. The study also hopes to learn something about what factors lead to optimal use of neural networks. This study will be conducted in the context of rank-ordering applicants to the U.S. Naval Academy.

Approach

Naval Academy data will be used to compare linear regression and neural networks. Currently, candidates to the Naval Academy are rank-ordered, using linear regression procedures. The seven variables used for prediction are SAT-Verbal, SAT-Quantitative, high school rank-in-class, recommendations from high school officials, extracurricular activity score, technical interest score, and career interest score (Wahrenbrock & Neumann, 1989). The criterion variables of primary interest are academic performance, military performance, type of major (technical or non-technical), and voluntary resignation.

This study will focus on two of these criteria, academic performance and voluntary resignation. Two different data sets will be used. In one case, data from the class of 1984 will be used to predict performance in the class of 1988 (the most recently available

data) and in the other case, the class of 1988 will be split into evaluation and cross-validation samples.

There are three phases of the project. In the first phase, the regression equations were calculated and used on the cross-validation samples. These will serve as the baseline against which the neural nets will be compared. The second phase is the neural network phase. This consists of creating and using 216 different network configurations. Once this is accomplished, phase three will consist of comparing the regression with the neural network results to learn about their relative performance.

Progress to Date

Phase 1, the regression portion of the study has been completed. Phase 2 is underway. All of the necessary networks have been created and are being run. Table 1 shows the cross-validation coefficients for the regression equations, as well as the best performance from neural networks to date.

Table 1

Cross-validation Coefficients for Linear Regression and Neural Networks

Cross Validation Criterion Dataset	Regression	Neural Net
AQPR 60/40	.5337	.5526
AQPR 84/88	.5814	.5806
VOLR 60/40	.0475	.0781
VOLR 84/88	.0561	.0858

Plans

The initial findings will be presented at the International Joint Conference of Neural Networks in Seattle in July 1991. A center report is planned. Areas for further research are being developed.

References

Hecht-Nielsen, R. (1987). *Kolmogorov's mapping neural network existence theorem*. In IEEE First International Conference on Neural Networks, Vol III, pp. 11-14.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feed forward networks as universal approximators, *Neural Networks*, 2, 359-366.

Rumelhart, D. E., & McClelland, J. L. (1988). *Parallel distributed processing*, MIT Press, Cambridge, Mass.

Wahrenbrock, A. L., & Neumann, I. (1989). *United States Naval Academy longitudinal officer data base documentation*. San Diego: Navy Personnel Research and Development Center.

Biography

William A. Sands is the Director of the Testing Systems Department. This department develops and evaluates systems for personnel testing, performance measurement, and person-job matching. Mr. Sands earned his B.S. in Social Sciences (1966) and his M.A. in Counseling and Testing Psychology (1967) from The American University, Washington, DC.

He is the Officer-in-Charge, U.S. Department of Defense Joint-service Computerized Adaptive Testing-Armed Services Vocational Aptitude Battery (CAT-ASVAB) Program; U.S. Navy Representative on the Technical Cooperation Program, Technical Panel UTP-3 (Military Human Resource Issues); consultant to NATO Panel 8, Research Study Group 15 (Computer-based Assessment of Military Personnel); and NPRDC representative on the Military Testing Association (MTA) Steering Committee.

He has authored over 40 journal articles, professional papers, and technical reports on psychological testing, personnel selection and classification, vocational guidance, and computer-based personnel systems.

Charles Wilkins is a Personnel Research Psychologist in the Testing Systems Department. He earned a Ph.D. in Quantitative Psychology at the Johns Hopkins University in 1989. He began working at NPRDC in July 1989 and is involved in psychometric research for the Computerized Adaptive Testing-Armed Services Vocational Aptitude Battery (CAT-ASVAB) Program. He recently began working on the application of Artificial Neural Network technology to personnel selection problems.

IR/IED FY90 Annual Report

Prediction Models for Dichotomous Criteria

William A. Sands
Charles Wilkins

Many important criteria in military personnel research are dichotomous or dichotomized (e.g., successful completion of first-term obligated service vs. premature attrition). Frequently, prediction of dichotomous criteria is done using Ordinary Least-Squares (OLS) linear regression techniques. This study was designed to develop, evaluate, and compare alternative prediction models for forecasting dichotomized criteria, using Artificial Neural Network (ANN) technology. Computer-simulated data distributions were created and used to evaluate the cross-validated predictive efficacy of OLS and ANN under a variety of personnel selection decision/outcome situations. Classification accuracy, defined as the proportion of correct selection/rejection decisions, will be the basis for comparing the different approaches. It is believed that ANN models will lead to equal or better predictive effectiveness than OLS models under various personnel selection decision conditions.

Background and Problem

Many of the important criteria for military personnel prediction problems are dichotomous (or continuous variables that have been dichotomized). An excellent example of these dichotomous criteria is successful completion of first tour of obligated service vs. premature attrition. The efficacy of prediction models for forecasting these dichotomous criteria is a paramount issue in personnel selection research.

Objectives

The primary objective of this research is to compare and contrast the predictive utility of the Ordinary Least Squares (OLS) linear regression model with an ANN approach for personnel selection. Use of ANN models for a variety of research problems has been receiving increasing attention (Hecht-Nielsen, 1990; Khanna, 1990; McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986). Ideally, this comparison should provide results that will allow generalization to a diverse set of personnel selection situations. This goal suggested the use of computer-simulated data. These data have the advantage of being "well-behaved" (in a statistical sense). Use of any real empirical data set runs the risk of limiting (perhaps severely) the extent to which the results may be generalized. A secondary objective is to develop expertise in the design and application of Artificial Neural Networks (ANNs) to personnel research.

Approach

The first step was to specify the dimensions of the study. These included the following:

1. Function forms of the data distributions (linear and curvilinear).

2. Sample sizes (100, 500, and 5000).

3. Errors--the degree to which the data points in the distribution deviate randomly from the ideal function form. This is measured by the standard deviation of the points around the function form. In the linear case, this error can be transformed to a validity coefficient (the correlation between the predictor and criterion).

4. Base rates--the proportion of successes before introducing a new selection instrument.

5. Selection ratios--the proportion of persons selected, using the new selection instrument.

6. Sample splits--the allocation of simulated persons ("simulees") in each total sample into two subsets: the development sample (used to develop a prediction model) and the evaluation sample (used to evaluate the prediction model).

The error of the distributions will be chosen so as to yield desired validities in the linear case (.05, .25, .50, .75, and .90). The errors corresponding to these target correlations were used to generate total bivariate distributions for three sample sizes (100, 500, and 5000) for each function form (linear and curvilinear). Then, the simulees in these total distributions were allocated to development or evaluation samples, with the following alternative splits (20-80%, 50-50%, and 60-40%).

A binary vector representing actual success vs. failure was created for each development sample. This distribution of simulees was rank-ordered and then separated into two groups (successes and failures). This procedure was followed separately for each alternative base rate (.05, .25, .50, and .95).

OLS regression equations were determined for each development sample. These equations were used to predict criterion scores for each simulee in the associated evaluation sample. These evaluation sample simulees were then rank-ordered by the predicted criterion score. Alternative selection ratios (.05, .25, .75, and .90) were imposed, dividing the evaluation sample into those who would be selected and those who would be rejected at each specified selection ratio.

At this point, we have the actual criterion status (success vs. failure) and selection vs. rejection status for each simulee. This allows us to form four decision-outcome combinations: (1) correct acceptances, (2) erroneous acceptances, (3) correct rejections, and (4) erroneous rejections. This information will be combined into the total number of correct decisions and the total number of erroneous decisions. The proportion of correct decisions ("hit rate") will be employed as the measure of effectiveness for comparing the OLS approach to the ANN approach, under each combination of conditions (function form, sample size, degree of error from the ideal function form, base rate, selection ratio, and sample split). These comparisons will enable us to identify those personnel selection situations wherein ANN models are likely to improve prediction over OLS models.

Progress and Plans

The dimensions of the study have been specified and the procedures for conducting the study have been delineated. The computer programs necessary for simulating, manipulating, and evaluating the bivariate score distributions have been written and tested.

The total bivariate distributions have been computer-generated for each function form and for each alternative degree of error.

Each total sample has been divided into a development and an evaluation sample. OLS regression equations have been computed on each development sample and then applied to the associated evaluation sample. The effectiveness of each OLS model has been determined.

Plans include the specification of the type of ANN model to be employed and the parameters for the model (architecture, transfer function, learning rate, momentum, etc.).

The ANN model will be trained on each development sample and then tested on the associated evaluation sample. Predictions will be made for the evaluation sample, based upon the ANN model, and then assessed in terms of the "hit rate" (proportion of correct decisions). Finally, the OLS and ANN approaches will be compared.

Results of the study will be documented in an NPRDC technical report.

References

Hecht-Nielsen, R. (1990). *Neurocomputing*. Reading, MA: Addison-Wesley Publishing Co.

Khanna, T. (1990). *Foundations of neural networks*. Reading, MA: Addison-Wesley Publishing Co.

McClelland, J. L., & Rumelhart, D. E. (Eds.) (1986). *Parallel distributed processing: Explorations in the microstructure of cognition-psychological and biological models*. (Vol. 2) Cambridge, MA: The MIT Press.

Rumelhart, D. E., & McClelland, J. L. (Eds.) (1986). *Parallel distributed processing: Explorations in the microstructure of cognition-foundations*. (Vol. 1) Cambridge, MA: The MIT Press.

Biography

Robert F. Morrison is a Senior Scientist in the Personnel Systems Department who has worked at NPRDC since 1976. He was born in ███████████████████. He acquired a B.S. (1952) in General Science and an M.S. (1956) in Applied Psychology from Iowa State University. In 1962, he received a Ph.D. in Industrial Psychology from Purdue University. Dr. Morrison worked in human resource management, emphasizing career development for Mobil Corporation, the Mead Corporation, and Martin-Baltimore. He has headed the personnel research activity for Sun Company and his own consulting firm as well as teaching in the School of Management Studies at the University of Toronto. His areas of research interest are management identification and the career development of managers, professionals, and scientists. His professional publications include: books, book chapters, and professional scientific articles. Dr. Morrison has won the James McKeen Cattell Award for Research Design.

Experienced-based Career Development

Robert F. Morrison

Human resource specialists need to be able to systematically design patterns of assignment that lead to the development of effective performance in positions many years after the career development process is begun. This research program will identify the steps and the time it takes an individual to master a single assignment and use this as a component in a life-span model of experiential learning. This effort provides an initial description (model) of the factors that influence how long it takes individuals to develop expertise within a specific assignment. When the research is completed, the Navy will have an algorithm to add learning time and performance level factors to the present methodology used to establish tour length. Currently, manpower and permanent-change-of-station (PCS) cost factors are the major factors considered.

Background and Problems

One assumption of formal education and training programs is that all efficient learning must occur in a structured (classroom-like) environment. Preliminary research on managerial positions challenges this assumption and indicates that the majority of learning occurs as a result of work experience (Brousseau, 1984; Campbell;[1] Fleischman & Mumford, 1989; Hall, 1991; Hall & Fukami, 1979; Kanarick;[2] McCall, Lombardo, & Morrison 1988; Morgan, Hall, & Martier, 1979; Vineberg & Taylor, 1972). While a model and propositions covering an entire career have been proposed (Morrison & Hock, 1986), the attributes of its components were not defined in detail. This definition is imperative to the adequate explication of the career development process.

---

[1]Campbell, S. (personal communication, 4 January 1985).

[2]Kanarick, A. (personal communication, 3 April 1984).

Since the Navy moved from pursuing its primary warfare mission in the mid-seventies to a peacetime status, the demands for its personnel, especially officers, to perform effectively in a wide variety of tasks and situations have increased markedly. For example, a program to encourage unrestricted line (URL) officers in the development of a secondary skill (subspecialty) foundered in the 1970s, yet culminated in the introduction of a material professional community in 1986 in response to Congress. In 1981, the Surface Warfare Commanders Conference focused on junior officers to increase their technical skills. In 1985, Congress imposed the requirements that all officers must serve in joint assignments in order to be eligible for promotion to flag (O-7).

This plethora of demands has forced policymakers to shorten billet and command tours until they are frequently less than 18 months. Such policies have been designed using manpower flow models without considering their effect on the officers' performance and career development. The fleet's

personnel readiness and the effectiveness with which support activities perform are affected directly by the opportunity that officers have to develop the capability to learn the requisite knowledge and skill of each billet and to develop them beyond the level of mastery. Tour lengths that are too short do not provide the opportunity to develop, while ones that are too long make inefficient use of the officer force and may lower the officers motivation to perform at a high level or learn new tasks/jobs.

Objective

The broad objective of this research is to develop a generic model describing the factors that influence how long it takes an individual to develop an expert-level of skill in performing work. The specific objective is to develop, test, and modify a preliminary model of the learning that occurs while the incumbent is in a mid-level leadership position on a surface ship.

General Approach

A literature search was used to identify: (1) the steps that an individual goes through in learning how to perform a job to the point of mastery, (2) the parameters that contribute to the level of performance at entry, and (3) the factors influencing what is learned and how quickly it is learned. This information was used to form an initial model of the experiential learning process.

The preliminary model was tested qualitatively via repeated interviews with 26 surface warfare department heads and 8 executive officers. Using these data, the initial model was revised (see Figure 1) and the results prepared for publication (Morrison & Brantner, 1989). Research was designed to test a situationally specific model of experiential learning on the population of surface warfare department heads. Therefore, only measures on the constructs that were considered most relevant to this population of officers were included.

With the support of the Office of the Chief of Naval Operations (OP-13) and Surface Forces, Atlantic Fleet and Pacific Fleet, the research was initiated by requesting commanding officer and department head data from 322 surface ships. These data emphasized perceptions of the individual differences, context, and job factors presented in Figure 1 and the extent to which the officers had learned their jobs regardless of how long they had been in the positions.

Results

The data from 292 officers representing 167 ships and 4 department head positions were available for use in the causal path analyses. As expected, a non-linear (quadratic) function of the number of months on the job accounted for the most variance (28%) in the level of learning achieved by the department heads. Individual models were developed for each of the four factors, individual differences, context, environment, and job/job characteristics. By themselves, the measures within the job/job characteristics factor increased the variance accounted for to 42 percent, 14 percent more than that accounted for by months on the job alone. In turn, measures for the individual differences, context, and environment factors increased the variance accounted for in the criterion to 39, 38, and 30 percent, respectively. When all significant variables from the four individual models were included in the development of the complete model, 50 percent of the variance in the dependent variable was explained, 79 percent more than that accounted for by just months on the job. The direct explanatory variables were as follows:
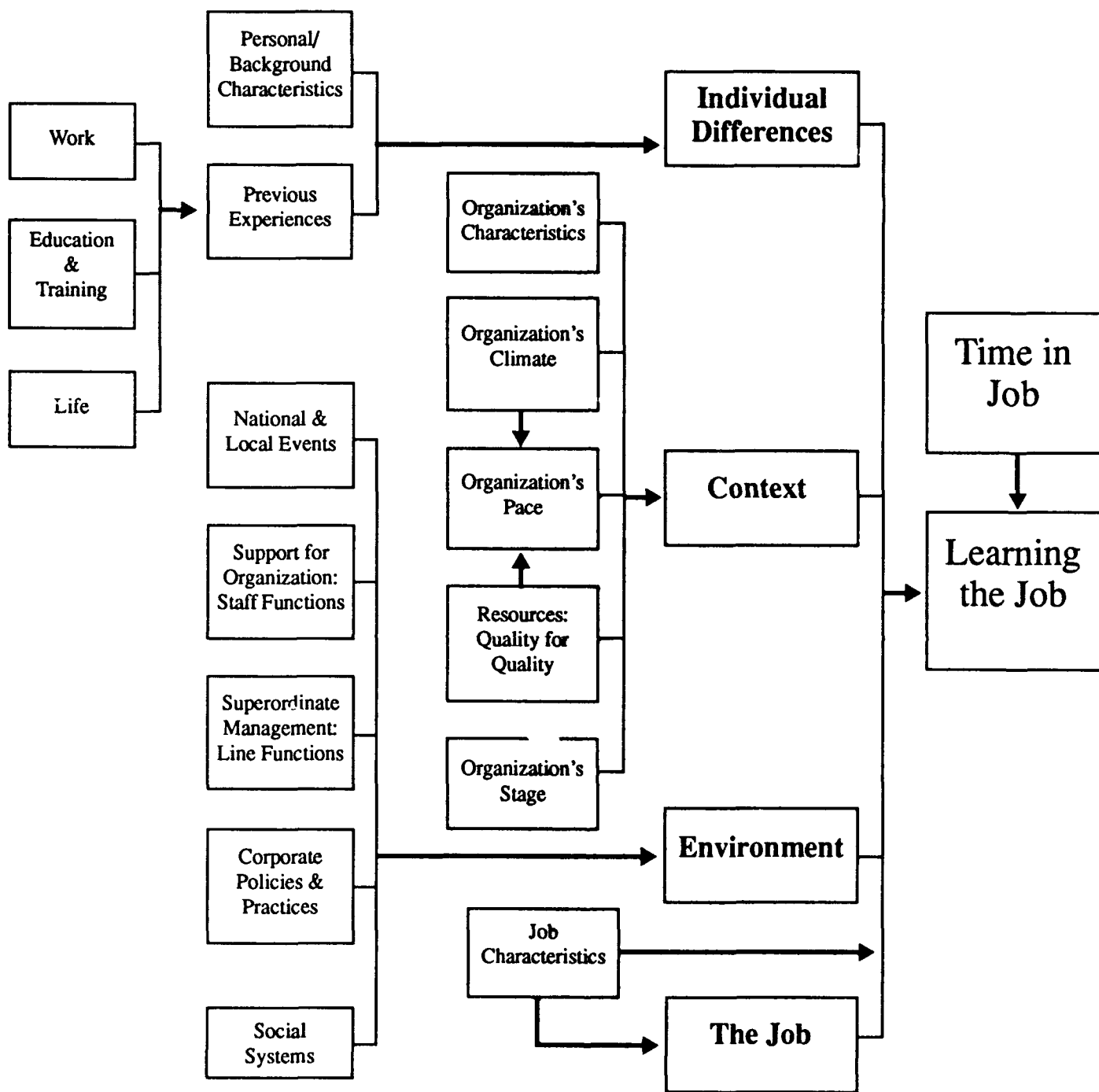
Figure 1. Job learning model.

1. Individual Differences

    a. Completing the first half of a split tour (immediate, similar experience).

    b. Perceiving that the department head assignment and prior experience were suitably matched (self efficacy).

    c. Feeling that leadership training did not aid in learning the job.

2. Job Characteristics

    a. Feeling that the department's mission-oriented goals were clear.

    b. Feeling that the job was boring.

3. Context

    a. Currently not in the local operations or repair portions of the ship's cycle.

    b. Entering the job while deployed or not during work-up.

    c. Having reasonable time to do the job and work on professional development.

    d. Having high quality enlisted personnel in the department.

4. Environment

    a. Feeling that marital status did not increase the time taken to learn the job.

Discussion and Conclusions

Descriptions of the secondary variables will be provided at the 59th Military Operations Research Society Symposium, June 1991. However, there were two major influences on the primary variables noted previously. One was the leadership provided by the commanding officer and executive officer in establishing clear mission-oriented command goals and reasonable time to do the

work. The other consisted of the prior experiences that prepared the officer for the present position. Carefully planned assignments that follow a hierarchy of learning appears to significantly reduce the time that it takes to learn the department head position in surface warfare.

Whether the relationships found in this research will hold when replicated in other organizations (Navy commands) or at different levels in the organization hierarchy is unknown, but theory would appear to support it (Morrison & Hoch, 1986). Replication is required if a model that generalizes across organizations and levels is to be developed. In addition, two major factors were not adequately measured in this research because there was not adequate time for the participants to provide the appropriate data. One of these factors was the environment and the other was the personal characteristics portion of individual differences. When the learning curves for component roles of the job were inspected, it appeared that the environment and the assignment and transition processes were related to a drop in learning the nontechnical aspects of the job. In two independent groups, decrements in the learning curves occurred at two points in the assignment, 6 to 8 months prior to the end and during the last 2 months. Other research has shown that personal characteristics, such as mental ability and self- confidence, should also be included in further investigations required to develop further this model of factors that influence the speed with which the job is learned (Fleischman & Mumford, 1989).

Expected Benefit

The Navy spends millions of dollars annually on PCS moves that are based primarily on manning and PCS cost considerations. By adding individual career development and performance factors to the decision process, the readiness of the fleet and the effectiveness

60

with which PCS dollars are used will improve. Since training proved to be a significant factor in the model, the model provides a basis for evaluating training programs.

References

Brousseau, K. R. (1984). Job-person dynamics and career development. *Research in Personnel and Human Resources Management, 2,* 125-154.

Fleischman, E. A., & Mumford, M. D. (1989). Individual attributes and training performance. In I. L. Goldstein (Ed.), *Training and development in organizations* (pp.183-255). San Francisco: Jossey Bass.

Hall, D. T. (1991). Twenty questions: Research needed to advance the field of careers. In R. F. Morrison & J. Adams (Eds.), *Contemporary issues in career development,* Hillsdale, N. J.: Lawrence Erlbaum.

Hall, D. T., & Fukami, C. V. (1979). Organization design and adult learning. *Research in Organizational Behavior, 2,* 125-167.

McCall, M. W., Lombardo, M. M., & Morrison, A. M. (1988). *The lessons of experience.* Lexington, MA: Lexington Books.

Morgan, M., Hall, D. T., & Martier, A. (1979). Career development strategies in industry: Where are we and where should we be? *Personnel, 56*(2), 13-31.

Morrison, R. F., & Brantner, T. M. (1989). *An experience-based learning model: A pilot study.* Unpublished manuscript. San Diego: Navy Personnel Research and Development Center.

Morrison, R. F., & Hoch, R. R. (1986). Career building: Learning from cumulative work experience. In D. T. Hall (Ed.), *Career development in organizations,* 236-273. San Francisco, CA.

Vineberg, R., & Taylor, E. N. (1972). *Performance in four Army jobs by men at different aptitude (AFQT) levels: 3. The relationship of AFQT and job experience to job performance* (HumRRO Tech. Rep. 72-22). Alexandria, VA: Human Resources Research Organization.

See Appendix B-2.

Biography

Jerry L. Vogt was a Research Psychologist in the Training Technology Department, where he directed several projects dealing with incorporating current and emerging technologies into the automation of training materials development. He is currently Professor of Psychology, St. Norbert College, Northfield, MN. He received a B.A. in Psychology from Pomona College in 1970 and a Ph.D. in Psychobiology from the State University of New York at Stony Brook in 1975. Prior to joining NPRDC in 1985, he had been involved in university teaching at St. John's University in Minnesota and in basic research at Stanford University as a National Science Foundation postdoctoral fellow. He also worked briefly for Courseware, Inc. in San Diego. His present interests include the areas of computer tools for writing, automated curriculum design and development, computer-based instruction, and the role of graphics in understanding instructional text. He has a number of publications, including two book chapters, a dozen journal articles, and several technical papers. He is a member of Human Factors Society, American Educational Research Association, and Association of Computing Machinery.

Photo Unavailable

Robert E. Gehring is a [part-time] Personnel Research Psychologist in the Training Technology Department and an Associate Professor in the Department of Psychology at the University of Southern Indiana. He teaches courses in experimental psychology including: Learning and Memory, Cognitive Psychology, Behavior Modification, and Motivation and Behavior. He received a B.A. in Psychology from Yale University in 1960 and a Ph.D. in Psychology from the University of Colorado in 1973. He served in the U.S. Army as a Personnel Psychology Specialist in testing from 1960 to 1962. Until 1968, he was training consultant to Shell Oil, Lever Brothers, and Sherwin-Williams Corporations. In 1987, he began a project funded by a grant from the Indiana Corporation for Science and Technology to develop and test videotapes to teach illiterate adults basic reading skills. He is currently interested in the superiority of pictorial to verbal learning and memory, the use of graphics and audiovisuals to enhance long-term retention, and children's first use of silent verbal thinking to improve memory performance. The latter research at the University of Oxford, England, 1990-1991. He has a number of publications including basic and applied research and two textbooks. He is a member of the American, Eastern, Midwestern, and Southwest Indiana Psychological Association and the American Society for Training and Development.

Using Diagrams for Learning Procedural Tasks[1]

Jerry L. Vogt
Robert E. Gehring

Although there has been a considerable amount of recent interest in the role of illustrations for understanding instructional material, little is known about the organization of diagrams or how they interact with text to aid learning. The goal of this research is to investigate how different types of diagrams assist students in understanding instructions so they can better perform procedural tasks. This work involves a series of experiments looking at the effects of different diagrams upon the learning and performance of different tasks. Undergraduate students and Navy recruits were tested on the retention of engines information obtained from actual self-study training material and from revised materials with simplified diagrams. Computer graphics enabled a gas-turbine-engine diagram to be simplified and separated into revised materials that yielded higher face validity and better immediate recognition memory. Navy subjects did better than undergraduates and benefitted more from revisions.

Background

A considerable amount of recent research focuses on the role of illustrations for understanding instructional material (Willows & Houghton, 1987a, 1987b). Beyond the general finding that text-relevant pictures enhance comprehension of the textual material (Winn, 1987; Levin, Anglin, & Carney, 1987), little is known about the microstructure of diagrams (Mayer, 1989) or how they interact with text to aid learning. The goal of this research is to investigate how different types of diagrams assist students in understanding instructional text so they can perform different types of tasks.

The instructional usefulness of diagrams can be viewed from several perspectives: adjuncts that can assist in processing at an enhanced level, aids for relearning procedural skills, and tools enabling people to form mental models of device operation. Any one of these factors may help explain why a diagram alone--compared to text-diagram condition--produces good performance in simple procedural assembly tasks (Konoske & Ellis, 1987; Schorr & Glock, 1983). Diagrams are probably more relevant to a more complex task, such as inferring structure or troubleshooting an operational problem, where the functional understanding provided by the graphical representation helps build the appropriate

IR/IED FY90 Annual Report

mental model (Kieras & Bovair, 1984). Moreover, a diagram showing the functional operation of a device should be more effective for teaching complex procedural task than a diagram showing parts of the device.

Technical Objective and Problem

The technical objective of this project is to investigate the efficiency of diagrams as adjuncts to learning procedural tasks. A review of diagram use in Navy training materials indicated a need for diagram guidelines relating to learning procedural tasks.

General Approach

This research performed experiments examining effects of different diagrams upon the learning and performance of different tasks. Before any studies were done, preliminary data were gathered on the types of diagrams used in Navy training materials and their relationship to the accompanying explanatory text. The purposes of this survey work were to (1) choose an appropriate domain area; (2) devise a classification scheme for diagrams, probably a representational vs. functional set of categories based on the work of Levin (Levin, Anglin, & Carney, 1987); (3) determine the prevalence of the types of diagrams (preliminary observations indicate that representational diagrams are about twice as numerous as functional diagrams); and (4) investigate the congruence of those graphic representations with the textual material.

Progress

Experiments were done to evaluate the effectiveness of different diagrams learning and retention from Navy Rate Training Manuals (RTMs). These are explanatory texts that contain many technical diagrams in order to help communicate principles of science and engineering that underlay technology. A section of a RTM explaining the gas turbine engine was chosen for study. Apparent problems with the presentation of technical information via text and diagrams were examined and rewritten with modified illustration. The effectiveness of the traditional RTM material was compared to the effectiveness of the revision. The main experimental questions pertained to the effects of manual changes on learner attitudes and retention. Would diagram simplifications lead to better immediate and delayed retention and would learners prefer to learn from simpler than from more complicated diagrams? Navy and undergraduate subjects were randomly divided into groups learning the different versions of the material.

We found that functional diagrams produced better performance than representational diagrams for initial learning of concepts and facts about the gas turbine engine. Furthermore, the Navy subjects outperformed the college undergraduates, even though none were experienced with gas turbines.

These results show that substantial improvements can be made in retention from training manuals through better organization and simplification of technical diagrams. Diagram simplification appears to be the key single factor. The results were reported in a paper presented at the annual meeting of the Psychonomics Society in November 1990.

A follow-up survey was done asking college students to compare the original and revised materials on several dimensions. Results showed that subjects reported that the revised materials would be easier to learn, that they were better organized, neater, and more appealing. Most subjects reported that the original materials were more confusing and poorly depicted airflow through the turbine.

Plans

The researchers are no longer directly associated with NPRDC. Results from the studies are being prepared for submission to a professional journal for publication.

Heuristic generalization of the results suggest that developers of training and reference diagrams need to carefully depict function in order to maximize effectiveness.

References

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science, 8,* 255-273.

Konoske, P., & Ellis, J. (1987). *Cognitive factors in learning and retention of procedural tasks* (NPRDC-TR-87-14). San Diego: Navy Personnel Research and Development Center.

Levin, J. L., Anglin, G. J., & Carney, R. N. (1987). *On empirically validating functions of pictures in prose.* In D. M. Willows & H. A. Houghton (1987a), *The psychology of illustration, Vol. 1, Basic research,* 51-85, New York: Springer Verlag.

Mayer, R. D. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology,* 240-246.

Schorr, F. L., & Glock, M. D. (1983). *Comprehending procedural instructions: The influence of comprehension monitoring strategies and instructional materials* (Tech. Rep. 83-10). Ithaca, NY: Cornell University Department of Education.

Willows, D. M., & Houghton, H. A. (1987a). *The psychology of illustration, Vol. 1, Basic research.* New York: Springer Verlag.

Willows, D. M., & Houghton, H. A. (1987b). *The psychology of illustration, Vol. 2, Instructional Issues.* New York: Springer Verlag.

Winn, B. (1987). Charts, graphs, and diagrams in educational materials. In D. M. Willows & H. A. Houghton (1987a), *The psychology of illustration, Vol. 1, Basic research* (1987a), 152-198, New York: Springer Verlag.

See Appendix B-2.

Biography

David G. Huntley is a Mathematical Statistician in the Manpower Systems Department. He earned a dual degree with distinction in French literature and mathematics, as well as an M.S. in Operations Research, from Stanford University. He is a member of Phi Beta Kappa. His research interests include applications of neural nets to forecasting, stochastic processes, and optimization.

Matthew F. Keblis is an Operations Research Analyst in the Manpower Systems Department. He received his B.A. in Economics from the University of Chicago in 1986. He earned his M.S. in Operations Research from the Illinois Institute of Technology in 1989. His current research interest is force management forecasting models. He is presently a member of the Operations Research Society of America.

Applications of Neural Net Technology to the Forecasting of Time Series

David G. Huntley
Matthew F. Keblis

The purpose of this research is to discover whether a neural net approach provides more accurate forecasts than conventional techniques such as Box-Jenkins and Autoregressive Integrated Moving Average Models (ARIMA). The "turning point" problem in forecasting is defined as the inability of conventional forecasting methods to anticipate fluctuations in the data; they are excellent at predicting trends, but when the trends change, there is a delay in recognizing the change. The goal of this research is to examine whether neural nets are more sensitive to fluctuations in the data.

Background and Problem

Navy personnel managers face conflicting objectives. It is their responsibility that manpower requirements are met in terms of quantity and quality. It is also their responsibility that the military personnel budget does not experience cost overruns. In fiscal year 1990, the budget amounted to $18,276,297,000, over half of which was allocated for the basic pay of enlisted personnel. The amount obligated for basic pay is directly proportional to the number of enlisted members. To estimate the budget accurately, Navy managers need an accurate forecast of the force structure matrix (the number of enlisted members by length-of-service (LOS) and paygrade).

Navy managers operate the Navy Personnel Pay Predictor, Enlisted (NAPPE) model, in part, to project the force structure matrix. The projected force matrix is then used with a table of expected statutory pay rates to generate an estimate of enlisted base pay. In NAPPE, ARIMA models are used to forecast the needed matrix. These models have proved satisfactory, although they have experienced problems with turning points. This shortcoming has stimulated interest in alternative methods to forecasting military personnel. Work by Lapedes and Farber (1987), among others, has established the potential benefit of using artificial neural networks in prediction. This in turn has created interest in applying neural network models to a variety of time series problems.

Objective

This effort was to compare the accuracy of the existing NAPPE forecasting method with an alternative neural network approach and to recommend a best model for forecasting each LOS contained within the force structure matrix.

Approach

A single hidden layer feed forward network was selected for this application. This model has been shown by White, Stinchcombe, and Hornik (1989) to be a universal approximator; it can provide an

accurate approximation to any function provided that the hidden layer has a sufficient number of hidden units. The generality of this model allowed the consideration of neural networks with the following architectures: an input layer with between 1 and 10 units, a single hidden layer with between 1 and 4 hidden units, and an output layer that contained 1 unit.

Within this framework fifty different models were fit to each LOS time series. These fifty models included simple linear networks (autoregressive models) and nonlinear autoregressive (NLAR) models. During training the simple linear models use the original LOS time series observations as targets, and use between 1 and 10 lags of the original time series as inputs to the network. There is no use of a hidden layer. The NLAR models involve the use of between 1 and 4 hidden units. During training the targets for these models are the errors that resulted from fitting the simple linear models to the original time series. The inputs to a hidden unit are the lags of the original time series (where the number of lags used corresponds to the number of lags used in generating the series of errors now being used as a target).

The Schwarz Information Criterion (SIC) was used to find the best model, from among the fifty fitted to a time series, for use in forecasting that LOS time series. It is defined as follows: $SIC=\ln(MSE)+p*\ln(n)/n$, where MSE is the mean square error for the model under consideration, p is the number of parameters in the model, and n is the number of observations in the training set. The SIC allows the selection of a model that balances increasing network complexity against decreasing mean square error. The use of an overparameterized model can result in poor forecasts. Use of the SIC helps to prevent this.

Results

End of quarter force structure matrices, spanning the period 3/57 to 6/90, were used to derive loss rate and population time series for each LOS. Starting with the 3/77 quarter and rolling forward through time, one, four, eight, and twelve quarters ahead, forecasts were generated by the models within NAPPE and by the 'best' neural network models for each LOS time series. A mean square error criterion was adopted to assess the results. They are summarized in Table 1. For example, when forecasting one quarter into the future, the neural network models outperformed the NAPPE models in approximately 53% of the LOS's. The table shows that the neural network models outperformed the NAPPE models the majority of the time, but care must be taken in interpreting the results. For the majority of the LOS's, simple linear models (no hidden units) were selected as the best neural network models.

Table 1

Forecasting Performance of Neural Network Models

Forecast Horizon	Percentage Better
One quarter	53.3
Four quarters	76.6
Eight quarters	75.8
Twelve quarters	71.4

Plans

Research continues in this area with the examination of more complex neural network models. Further information and details concerning this effort will appear in a report entitled *Forecasting military manpower with neural networks*.

Expected Benefit

The impact of slight improvements in forecast accuracy can be substantial. This is true because even small percent errors in a forecast of the Navy's 18 billion dollar military personnel budget can result in very large cost overruns.

References

Lapedes, A., & Farber, R. (1987). *Nonlinear signal processing using neural networks: Prediction and system modeling.* (Tech. Rep., LA-UR-87-2662). Los Alamos National Lab.

White, H., Stinchcombe, M., & Hornik, K. (1989). Multilayer feedfoward networks are universal approximators. *Neural Network, 2,* 359-366.

See Appendix B-2.

Biography

Theodore J. Thompson was born ███████████ ████████████. He earned a B.A., cum laude, in Mathematics and an M.S. in Statistics from the State University of New York at Buffalo. Ted was employed as a statistician for 3 years by David F. Herring Incorporated prior to coming to NPRDC in June 1981. As an operations research analyst in the Manpower Systems Department, Ted has applied his skills in statistics, operations research, and computer science to solve problems in Navy enlisted personnel distribution. He also served as Science Advisor to Chief of Naval Personnel for 6 months during 1990. He is presently a member of the American Statistical Association and the Biometric Society.

Iosif A. Krass is an Operations Research Analyst in the Manpower Systems Department. His research specialties are large-scale optimization and dynamic control systems. For the last 5 years, he has been developing and applying advanced techniques in operations research, computer science, and control theory to solve complex problems in Navy enlisted personnel assignment and rotation. Educated and trained in the Soviet Union as an operations research analyst, he started his research work at the Institute of Mathematics, Siberian Branch of Academy of Science, Novosibirsk, USSR (1967-78). After immigrating to the U.S., he was an Associate Professor at Southern Illinois and Kansas Universities (1979-81). He also worked as a systems and programming consultant for Control Data Corporation in Connecticut (1981-85). He is a member of the Operations Research Society of America. He has over 20 years of research experience and he has authored or co-authored numerous books, journal articles, and professional papers.

Decomposition Method for Solving a Large-scale Model to Maximize Navy Unit Personnel Readiness

Theodore J. Thompson
Iosif A. Krass

Every Navy combat unit is required to report the status of personnel readiness for the unit. The goal of the personnel unit readiness report is to ensure that combat ships and squadrons have sufficient personnel with specific skills to operate. However, personnel unit readiness criteria are not directly considered when personnel are assigned. The problem of maximizing personnel readiness of the U.S. Navy fleet can be formulated as a large-scale, mixed integer problem. The size of the readiness problem is approximately 50,000 variables and 30,000 constraints. We have developed a heuristic algorithm, which provides a starting solution, and a decomposition method capable of solving this problem on moderate sized mainframe computers (e.g., NPRDC's IBM 4381).

Background

The problem of maximizing personnel readiness of the U.S. Navy fleet can be formulated as a large-scale, mixed integer problem. The size of the readiness problem is approximately 50,000 variables and 30,000 constraints. Commercial linear programming packages cannot solve this size problem, within reasonable time and space requirements, even using a supercomputer. A possible exception is the AT&T KORBX machine, which is an integrated hardware and software optimization system. We are developing a decomposition method capable of solving this problem on moderate sized mainframe computers (e.g., NPRDC's IBM 4381).

The readiness problem can be formulated using existing data files. The demand side of the problem can be defined from manning files. These files provide the shortage or excess of personnel within pay grade, skill, and mission area for readiness activities.

The supply side of the problem is defined from Enlisted Projection System files. These files contain projected number of people by skill, pay grade, and composite available to fill job openings. Composite defines sea or shore eligible.

Progress

A mathematical model was formulated which presents personnel unit readiness as an optimization on a network with side constraints. Based on this mathematical model, the problem was decomposed into a large-scale network (about 22,000 nodes and 50,000 arcs) with a relatively small number (about 4,000) of side constraints. The model generator was written in Fortran and debugged. We attempted to solve this

problem with Netsid but were unsuccessful. Netsid is a state-of-the-art "network with side constraints" code that we obtained from Professor Kennington of Southern Methodist University.

We then developed a heuristic algorithm to solve the readiness optimization problem on our IBM 4341. This heuristic solution was then used as the starting point for Netsid. This was also unsuccessful. However, we now have what we believe is a good heuristic algorithm for solving the problem. Combining our heuristic with a decomposition into subproblems, which can be solved optimally, may provide a good solution strategy.

Additionally, this will provide information on the effectiveness of our heuristic.

Our work on this problem has generated interest in the academic community. Professor Stavros Zenias from Warton School, University of Pennsylvania said that he has a code that can solve this size problem. We provided the model specifications to him. Their initial attempts to solve the optimization problem failed. We will continue to collaborate with him.

References

See Appendix B-2.

APPENDIX A

PROJECT TRANSITIONS

Transitions

Independent Research

Experienced-based career development (0601152N.R000.06) transitioned into exploratory development 6.2 research.

Independent Exploratory Development

Measures of effectiveness (0602936N.RV36I27.08) will transition into the 6.2 project *Training management expert systems.*

APPENDIX B

PRESENTATIONS AND PUBLICATIONS

Presentations

Independent Research

Federico, P-A. (September 1991). *Recognition measurement: Computer-based and paper-based methods.* Paper accepted for presentation at the annual meeting of the Human Factors Society, San Francisco, CA.

Federico, P-A. (November 1991). *Protocol analysis procedures and complex cognitive performance.* Abstract submitted for presentation at the annual meeting of the Psychonomic Society, San Francisco, CA.

Lewis, G. W., & Ryan-Jones, D. L. (October 1991). *Event-related potential averages: How many epochs are needed?* Paper submitted for presentation at the 31st annual meeting of the Society of Psychophysiological Research, Chicago, IL.

Morrison, R. F., & Brantner, T. M. (June 1991). *What affects how quickly a new job is learned?* Paper to be presented at the annual meeting of the Military Operations Research Society, West Point, NY.

Morrison, R. F.[1] (August 1991). *What affects how quickly a new job is learned?* Paper to be presented at the annual meeting of the Academy of Management, Miami Beach, FL.

Williams, D. (1989). *ERPs and the suffix effect.* Paper presented at the annual meeting of the Pscyhonomics Society, Atlanta, GA.

Independent Exploratory Development

Huntley, D. G. (April 1990). *Neural nets: An approach to the forecasting of time series.* Paper presented at the Conference on Advanced Computing for the Social Sciences, Williamsburg, VA.

Krass, I. A. (May 1990). *Shadow method for convex programming with application.* Paper presented at the meeting of the Operations Research Society of America and the Institute of Management Sciences, Las Vegas, NV.

Krass, I. A. (October 1990). *Optimization technique for multiple objective assignment programs.* Paper presented at the 10th Office of Naval Technology Symposium, Washington, DC.

Krass, I.A., & Thompson, T. J. (October 1991). *Decomposition method for solving a large-scale model to maximize Navy unit personnel readiness.* Paper to be presented at the meeting of the Operations Research Society of America and the Institute of Management Sciences, Anaheim, CA.

Gehring, R. E., & Toglia, M. P. (November 1990). *Effects of technical diagram simplification on retention from training manuals.* Paper presented at the annual meeting of the Psychonomics Society, New Orleans, LA.

[1]See Appendix C-1, Awards and Honors.

Publications

Independent Research

Callahan, J., & Sorensen, S. (1991). Rule induction for group decisions with statistical data, an example. *Journal of the Operational Research Society, 42*(3), 227-234.

Federico, P-A. (in press). Measuring recognition performance using computer-based and paper-based methods. *Behavior Research Methods, Instruments, & Computers.*

Federico, P-A. (1991). Student cognitive attributes and performance in a computer-managed instructional setting. In R. F. Dillon & J. W. Pellegrino (Eds.), *Instruction: Theoretical and applied perspectives.* New York, NY: Praeger Publishers.

Federico, P-A. (in press). Assessing semantic knowledge using computer-based and paper-based media. *Computers in Human Behavior.*

Huntley, D. G. (1991). Neural nets: An approach to the forecasting of time series *The Social Science Computer Review, 9*(1), 27-38.

Morrison, R. F., & Brantner, T. M. (August 1991). What affects how quickly a new job is learned? *Proceedings of the Academy of Management Annual Meeting.*

Ryan-Jones, D. L., & Lewis, G. W. (1991). Application of neural network methods to prediction of job performance. *Proceedings of the 2nd Workshop on Neural Networks: Academic/Industrial/NASA/Defense, Auburn University, AL,* (731-735). Madison, WI: Omnipress.

Vogt, J., Gross, S., & Winnifeld, M. (in preparation). *Use of graphics in Navy rate training manuals.* San Diego: Navy Personnel Research and Development Center.

Wilkins, C., & Dickieson, J. (July 1991). Using neural network technology for predicting military attrition. *Proceedings of the International Joint Conference on Neural Networks-91, Seattle, WA.*

Independent Exploratory Development

Boyle, P. (August 1989). A comparison of Stein and Lindley rules. *Proceedings of the Business and Economic Statistics Section 1989 of the American Statistical Association.*

Cowen, M. (in process). Computer-based training (CBT) for operating a digitally controlled device: The role of feedback. *Educational Technology Research and Development Journal.*

Krass, I. A., & Thompson, T. J. (submitting). Decomposition method for solving a large-scale model to maximize Navy unit personnel readiness. *Computer and Mathematical Applications.*

Krass, I. A. (October 1989). Shadow method for convex programming with application for Navy credit sea/shore rotation problem. *Computer and Mathematics with Applications, 16*(12), 1065-1077.

Sorensen, S., & Callahan, J. D. (submitted). Using TETRAD II as an automated exploratory tool, *Social Science Computer Reviews.*

Sorensen, S., Sprites, P., Glymour, C., & Scheines, R. (1991). *Using artificial intelligence to aid in the development of causal models.* Technical report to the MP&T R&D OCNR Program.

Sorensen, S., Sprites, P., Glymour, C., & Scheines, R. (in process). Using artificial intelligence to aid in the development of causal models. Portions of the work appear in the book, *Causality, Statistics, and Search.*

APPENDIX C

AWARDS AND HONORS

Awards and Honors

Independent Research

A paper describing the *Experienced-based Learning II* (0601152N.R0001.15) research received a "best paper" award when submitted to the Careers Division of the Academy of Management for presentation at its annual meeting in August 1991. About 1 in 20 papers submitted receive such an award. Dr. Robert Morrison's paper, *What Affects How Quickly a New Job is Learned?* will be published in the proceedings of the August meeting.

Distribution List

Distribution:

Assistant Secretary of Defense

Deputy Under Secretary of Defense for Research and Engineering (Research and Advanced Technology)

Assistant for Training and Personnel Technology (OUSD) (A)/R&T(E&LS)

Assistant Secretary of the Navy (Manpower and Reserve Affairs) (OASN) (M&RA)

Deputy Assistant Secretary of the Navy (Manpower)

Chief of Naval Personnel

Assistant for Manpower, Personnel, and Training (OP-987H)

Commander, Naval Sea Systems Command

Commander, Naval Air Systems Command

Annapolis Technical Information Center, D. W. Taylor Naval Ship R&D Center

Commanding Officer, Naval Coastal Systems Center

Commanding Officer, Naval Underwater Systems Center

Commander, Naval Ocean Systems Center

Commander, Naval Air Development Center

Officer in Charge, Naval Surface Warfare Center Detachment, White Oak Laboratory

Commander, Naval Weapons Center, Technical Library (Code 343)

Life Sciences Directorate (OCNR-30T)

Office of Naval Research (OCNR-10)

Office of Naval Research, Detachment London

Office of Naval Technology (OCNR-20T)

Chief of Naval Education and Training (Code 00)

Chief, Bureau of Medicine and Surgery (MED 25)

Commanding Officer, Naval Aerospace Medical Research Laboratory

Commandant of the Marine Corps

Commandant of the Marine Corps (RD)

Commander, U.S. ARI, Behavioral and Social Sciences (PERI-POT-I)

Department of the Air Force, DET5, Armstrong Laboratory Directorate, Brooks Air Force Base, TX

Program Manager, Life Sciences Directorate, Bolling Air Force Base (AFOSR/NRL)

Commander, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base (AFHRL/LRS-TDC)

Department of the Air Force (AFSC), Air Force Human Resources Laboratory, Lowry Air Force Base (AFHRL/ID)

Director of Training, Office of Civilian Personnel Management

Superintendent, Naval Postgraduate School

Director of Research, U.S. Naval Academy

Executive Director, U.S. Naval Institute

Institute for Defense Analyses, Science and Technology Division

Center for Naval Analyses, Acquisition Unit
Defense Technical Information Center (DTIC)

Copy to:

Life Sciences Directorate (OCNR-30T), (OCNR-0722)

IR/IED FY90 Annual Report

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE June 1991	3. REPORT TYPE AND DATE COVERED Administrative Publication--Oct 89-Oct 90

4. TITLE AND SUBTITLE
Independent Research and Independent Exploratory Development Programs:
FY90 Annual Report

5. FUNDING NUMBERS
Program Elements 0601152N
and 0602936N

6. AUTHOR(S)
Montague, William E. and Scheifers, Carmen C. (Editors)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Navy Personnel Research and Development Center
San Diego, California 92152-6800

8. PERFORMING ORGANIZATION REPORT NUMBER
NPRDC-AP-91-15

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Technology/Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The FY90 Independent Research/Independent Exploratory Development (IR/IED) programs began with a call for proposals in June 1989. Technical reviews were provided by supervisors and scientific consultants and eight IR and three IED projects were funded. This report documents the results and accomplishments of these projects. It lists the projects active during FY90 and those supported in FY91. Two papers, one IR and one IED, chosen by the Technical Director as "Best Papers of 1989" are presented. Subsequent pages contain brief reports of research progress during FY90 written by the principal investigators of each project. This report lists the IR and IED projects that may have transitioned into other projects or into use by the Navy during the year. It itemizes the presentations and publications from IR and IED supported projects and presents awards and honors related to the projects.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UNLIMITED

NSN 7540-01-280-5500

NPRDC

NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
SAN DIEGO

Seapower Through People