

RICE UNIVERSITY

**Evaluating Organizational Response to a Cognitive Problem:
A Human Factors Approach**

by

Elizabeth May Serig

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Kenneth R. Laughery, Sr., Chair
Herbert S. Autrey Professor, Psychology

Michael D. Byrne
Assistant Professor of Psychology

James R. Pomerantz
Professor, Director of Neuroscience
Psychology

H. Albert Napier
Professor, Director of Center for
Management of Information Technology
Jones Graduate School of Management

Houston, Texas

May, 2001

ABSTRACT

Evaluating Organizational Response to a Cognitive Problem: A Human Factors

Approach

by

Elizabeth May Serig

The commission of error is often perceived as the result of such internal attributes as negligence, laziness, carelessness, and inattention. In organizational settings, such a perception often leads to the administration of punitive actions against the responsible individual. Recent research on error, however, has moved thinking from a “conventional wisdom” perspective of human error to a systems perspective. According to this systems perspective, humans are remarkably reliable “stand-alone” systems, and errors tend to arise primarily when humans interact with technological systems. Errors can be triggered by technology and its environment, as a result of the way these factors interact and challenge human limitations. Byrne and Bovair (1997) found that the commission of a particular type of error, postcompletion error, is related to a high working memory load imposed by external forces or task complexity. Two experiments were designed to assess the effects of typical organizational responses to error on the commission of postcompletion errors over time. Because organizations tend to assume that errors are under the control of the individual, methods such as reprimands and re-instruction are often administered to “motivate” individuals to not commit errors. Similarly, praise is often administered to encourage the continuation of appropriate behavior. A systems perspective, however, would argue that a troublesome task should be redesigned to accommodate the limitations of the human cognitive system under certain circumstances.

The results of the experiments reported here indicated that, over time, simple tasks were learned so well that people made few errors, and therefore, responses to error appeared to have little effect on the commission of error. It was found, however, that when a task was redesigned, participants were much quicker at executing a critical redesigned task step than participants who were reprimanded, received re-instruction, or were praised for their performance. This indicates that the cost of low-error performance for these participants came at the cost of increased time to complete the critical step, further indicating that these participants had to consciously expend effort to not commit the error.

ACKNOWLEDGMENTS

For sharing their time and expertise, I would like to express my appreciation to the members of the committee: Dr. Kenneth R. Laughery, Sr., Dr. Michael D. Byrne, Dr. James R. Pomerantz, and Dr. H. Albert Napier. Their comments, constructive criticisms, and support have proven invaluable throughout this process. In particular, I would like to thank Dr. Byrne for developing the software application for the Bridge Officer Qualification Tasks. Additionally, I would like to thank those others without whom this research never would have been completed: Brian Webster for helping with the software application development, Bryan Blauvelt and Peter Centgraf for running participants, Monica Freedman for developing the reverse digit-span task, all the guys in the CHIL lab for pitching in when I needed it, the Office of Performance Improvement at The University of Texas M. D. Anderson Cancer Center for giving me a chance and letting me abuse office supplies, and finally, Ashley Sides for being moral support and for helping me with that most important but frustrating of all tasks - pinning down the details of experimental design. I would like to dedicate my dissertation to my family (and yes, that means you too, Bob); the best darn “autistic” committee around. Your input, support, patience in the face of incessant babbling, data mangling, and most importantly, kicks in the butt have made all the difference. Finally, where would I be without my “boys” - Jake, Tucker and Tyler - who let me know when it was time to work, time to take a break, and time to break out the Friskies.

TABLE OF CONTENTS

INTRODUCTION AND OVERVIEW	1
HUMAN ERROR	3
WHAT ARE ERRORS?	3
TWO PERSPECTIVES ON ERROR	4
SYSTEMS THEORIES OF ERROR	6
POSTCOMPLETION ERROR	12
GENERAL THEORY OF POSTCOMPLETION ERROR	13
COMPUTATIONAL MODEL OF POSTCOMPLETION ERROR	15
ORGANIZATIONAL RESPONSES TO ERROR	17
MODEL OF ORGANIZATIONAL DISCIPLINE	18
ORGANIZATIONAL DISCIPLINE IN HEALTHCARE	23
THE UNDERLYING ASSUMPTION OF ORGANIZATIONAL RESPONSES TO ERROR	28
EXPERIMENTS	32
EXPERIMENT 1	42
METHOD	42
PARTICIPANTS	42
MATERIALS	42
DESIGN	42
PROCEDURE	45
RESULTS	47
TRAINING	47
TESTING	49
DISCUSSION	61
EXPERIMENT 2	62
METHOD	64
PARTICIPANTS	64
MATERIALS	64
DESIGN	64
PROCEDURE	66
RESULTS	70
TRAINING	71

TESTING	73
OVERALL TASK PERFORMANCE	78
PERFORMANCE AT “TURN OFF TRACKING”	81
PERFORMANCE AT “POWER CONNECTED 1” AND “SCANNER ON”	88
PERFORMANCE AT “POWER CONNECTED 2” AND “SCANNER OFF”	93
LETTER RECALL TASK	96
WORKING MEMORY CAPACITY	98
DISCUSSION	98
GENERAL DISCUSSION	104
GENERAL MODEL OF TASK PERFORMANCE OVER TIME	106
COGNITIVE COMPONENTS OF TASK PERFORMANCE	106
SKILL ACQUISITION IN LEARNING	106
WORKING MEMORY	108
ERROR	110
INTEGRATED MODEL OF OVERALL TASK PERFORMANCE OVER TIME	112
CONCLUSION	119
REFERENCES	123
APPENDIX A OPERATIONS OFFICER QUALIFYING EXAMS: CADET MANUAL	126
APPENDIX B TACTICAL TASK TRAINING MANUAL: POSTCOMPLETION VERSION	129
APPENDIX C TACTICAL TASK TRAINING MANUAL: CONTROL VERSION	142
APPENDIX D TRANSPORTER TASK TRAINING MANUAL	155
APPENDIX E CONN TASK TRAINING MANUAL	167

APPENDIX F	
OPERATIONS TASK TRAINING MANUAL	174
 APPENDIX G	
EXIT QUESTIONNAIRE	181
 APPENDIX G-1	
EXIT QUESTIONNAIRE FOR CONTROL AND	
POSTCOMPLETION CONDITIONS	182
 APPENDIX G-2	
EXIT QUESTIONNAIRE FOR REPRIMAND CONDITION	185
 APPENDIX G-3	
EXIT QUESTIONNAIRE FOR RE-INSTRUCTION CONDITION	189
 APPENDIX G-4	
EXIT QUESTIONNAIRE FOR PRAISE CONDITION	193
 APPENDIX G-5	
EXIT QUESTIONNAIRE FOR REDESIGN CONDITION	197

LIST OF TABLES

TABLE 1.	CORRECTIVE ACTIONS TAKEN BY NRC LICENSEES IN RESPONSE TO RADIOPHARMACEUTICAL MISADMINISTRATIONS	25
TABLE 2.	EXPERIMENT 2: INTERVENTION TEXT FOR THE REPRIMAND, RE-INSTRUCTION, PRAISE, AND REDESIGN CONDITIONS	69
TABLE 3.	EXPERIMENT 2: ANALYSIS OF VARIANCE FOR NUMBER OF TRAINING TRIALS BY BRIDGE TASK STATION AS A FUNCTION OF CONDITION	72
TABLE 4.	EXPERIMENT 2: PLANNED COMPARISONS FOR THE INTERACTION OF TEST PERIOD AND CONDITION ON ERROR FREQUENCY ON THE “TURN OFF TRACKING” STEP IN THE TACTICAL TASK	84
TABLE 5.	EXPERIMENT 2: PLANNED COMPARISONS FOR THE INTERACTION OF TEST PERIOD AND CONDITION ON THE PROPORTION OF POSTCOMPLETION ERROR TO TOTAL ERROR AT THE “TURN OFF TRACKING” STEP IN THE TACTICAL TASK	85
TABLE 6.	EXPERIMENT 2: PLANNED COMPARISONS FOR THE INTERACTION OF TEST PERIOD AND CONDITION ON STEP COMPLETION TIME AT THE “TURN OFF TRACKING” STEP IN THE TACTICAL TASK	87

LIST OF FIGURES

FIGURE 1.	EVIDENCE IN THE POPULAR PRESS FOR THE BELIEF THAT THE OCCURRENCE OF ERROR IS UNDER THE CONTROL OF THE INDIVIDUAL AND THAT PUNITIVE METHODS “MOTIVATE” EMPLOYEES TO ERROR-FREE BEHAVIOR (YOUNG & LEBRUN, 2001)	29
FIGURE 2a.	TACTICAL TASK STRUCTURE - CONTROL VERSION	35
FIGURE 2b.	TACTICAL TASK STRUCTURE - POSTCOMPLETION VERSION ..	36
FIGURE 3.	TACTICAL TASK SCREEN AND STEP SUMMARY FOR BOTH THE CONTROL AND POSTCOMPLETION VERSIONS	37
FIGURE 4.	TRANSPORTER TASK STRUCTURE	38
FIGURE 5.	TRANSPORTER TASK SCREEN AND TASK SUMMARY	39
FIGURE 6.	CONN TASK STRUCTURE AND CONN TASK SCREEN WITH STEP SUMMARY	40
FIGURE 7.	OPERATIONS TASK STRUCTURE AND OPERATIONS TASK SCREEN WITH STEP SUMMARY	41
FIGURE 8.	EXPERIMENT 1: MEAN NUMBER OF TRAINING TRIALS NECESSARY AS A FUNCTION OF CONDITION	48
FIGURE 9.	EXPERIMENT 1: FREQUENCY OF ERROR AT THE POSTCOMPLETION STEP (“TURN OFF TRACKING”) IN THE TACTICAL TASK AS A FUNCTION OF CONDITION	51
FIGURE 10.	EXPERIMENT 1: PROPORTION OF ERROR AT THE POSTCOMPLETION STEP (“TURN OFF TRACKING”) TO TOTAL ERROR IN THE TACTICAL TASK AS A FUNCTION OF CONDITION	52
FIGURE 11.	EXPERIMENT 1: MEAN TASK COMPLETION TIME ON THE TACTICAL TASK AS A FUNCTION OF CONDITION	55
FIGURE 12.	EXPERIMENT 1: MEAN STEP COMPLETION TIME AT THE “TURN OFF TRACKING” STEP IN THE TACTICAL TASK AS FUNCTION OF CONDITION	56

FIGURE 13.	EXPERIMENT 1: MEAN PERCENT OF TRIALS CORRECT ON THE LETTER RECALL TASK AS A FUNCTION OF CONDITION	59
FIGURE 14.	EXPERIMENT 1: MEAN RESPONSE LATENCY FOR THE LETTER RECALL TASK AS A FUNCTION OF CONDITION	60
FIGURE 15.	EXPERIMENT 2: MEAN NUMBER OF TRAINING TRIALS NECESSARY AS A FUNCTION OF BRIDGE TASK STATION	72
FIGURE 16.	EXPERIMENT 2: PERCENT OF TOTAL ERRORS MADE AS A FUNCTION OF THE STEP IN THE TACTICAL TASK PROCEDURE	74
FIGURE 17.	EXPERIMENT 2: PERCENT OF TOTAL ERRORS MADE AS A FUNCTION OF THE STEP IN THE TRANSPORTER TASK PROCEDURE	75
FIGURE 18a.	TACTICAL TASK SCREEN AND STEP SUMMARY FOR THE POSTCOMPLETION VERSION	76
FIGURE 18b.	TRANSPORTER TASK SCREEN AND STEP SUMMARY	76
FIGURE 19.	EXPERIMENT 2: FREQUENCY OF OVERALL ERROR ACROSS THE FOUR TEST PERIODS AS A FUNCTION OF CONDITION . . .	79
FIGURE 20.	EXPERIMENT 2: MEAN TASK COMPLETION TIME ACROSS ALL BRIDGE STATION TASKS AS A FUNCTION OF CONDITION	80
FIGURE 21.	EXPERIMENT 2: MEAN TASK COMPLETION TIME FOR THE TACTICAL TASK AS A FUNCTION OF CONDITION	81
FIGURE 22.	EXPERIMENT 2: FREQUENCY OF ERROR AT THE “TURN OFF TRACKING” STEP IN THE TACTICAL TASK	82
FIGURE 23.	EXPERIMENT 2: THE PROPORTION OF POSTCOMPLETION ERROR AT THE “TURN OFF TRACKING” STEP TO TOTAL ERROR IN THE TACTICAL TASK	83
FIGURE 24.	EXPERIMENT 2: MEAN STEP COMPLETION TIME AS A FUNCTION OF CONDITION AT THE “TURN OFF TRACKING” STEP IN THE TACTICAL TASK	87

FIGURE 25.	EXPERIMENT 2: COMPARISON OF THE FREQUENCY OF ERROR AT "POWER CONNECTED 1" IN THE TACTICAL TASK AND "SCANNER ON" IN THE TRANSPORTER TASK	90
FIGURE 26.	EXPERIMENT 2: COMPARISON OF THE PROPORTION OF ERROR AT THE "POWER CONNECTED 1" STEP IN THE TACTICAL TASK AND THE "SCANNER ON" STEP IN THE TRANSPORTER TASK	91
FIGURE 27.	EXPERIMENT 2: MEAN STEP COMPLETION TIME FOR THE "POWER CONNECTED 1" STEP IN THE TACTICAL TASK AND THE "SCANNER ON" STEP IN THE TRANSPORTER TASK	92
FIGURE 28.	EXPERIMENT 2: COMPARISON OF THE FREQUENCY OF ERROR AT "POWER CONNECTED 2" IN THE TACTICAL TASK AND "SCANNER OFF" IN THE TRANSPORTER TASK	94
FIGURE 29.	EXPERIMENT 2: COMPARISON OF THE PROPORTION OF ERROR AT "POWER CONNECTED 2" IN THE TACTICAL TASK AND "SCANNER OFF" IN THE TRANSPORTER TASK	95
FIGURE 30.	EXPERIMENT 2: MEAN STEP COMPLETION TIME FOR THE "POWER CONNECTED 2" STEP IN THE TACTICAL TASK AND "SCANNER OFF" IN THE TRANSPORTER TASK	96
FIGURE 31.	EXPERIMENT 2: MEAN PERCENT OF TRIALS CORRECT ON THE LETTER RECALL TASK	97

INTRODUCTION AND OVERVIEW

An elderly man arrived at the emergency department of a large urban hospital after falling and hitting his head during the night. An examination revealed that the man had lost a large amount of blood from a cut on his head, his blood pressure was fluctuating, and he was in a great deal of pain from his fall. The doctor decided to admit the man to the hospital for observation and wrote up a set of admitting orders. As she finished the orders, a nurse called the doctor away to attend a patient who arrived in cardiac arrest. Finished with the orders, the doctor left to attend to the new patient. An hour later, a nurse checked on the elderly man and noticed that the doctor failed to sign the admitting orders. The nurse spent the better part of another hour tracking down the doctor to sign the orders. During this time, the elderly man suffered acute distress from the pain caused by his injuries; but he could not be administered a prescribed pain reliever until his doctor was found and signed the orders (D. Snow, personal communication, January 24, 2001) .

In a different hospital, a nurse prepared a new IV bag of heparin, a commonly used anticoagulant, for a patient. The nurse turned the IV pump to “hold,” and removed the old IV bag and tubing. The nurse spiked and hung the new IV bag. As a patient in the next bed began complaining of stomach upset, the nurse changed the pump settings to a new infusion rate and turned the pump back on. Concerned about the patient in the next bed, the nurse left the first patient. Later, it was discovered that the nurse failed to break a seal in the IV bag that allows the drug to mix with a dextrose diluent that provides the final admixture for administration. The patient developed impaired circulation and had to

have his leg amputated (D. Snow, personal communication, January 24, 2001; ISMP, 1999).

What happened in these events? Did the doctor simply forget to sign the admitting orders? Did the nurse simply forget to break the seal in the IV bag? Or is something much more complex occurring than random acts of forgetting?

Errors are not made in a vacuum. Despite conventional wisdom that errors are the result of negligence, laziness, and the like, humans are remarkably reliable as a “stand-alone” system. Problems tend to arise, however, when humans become part of a system, particularly when they interact with technology (Van Cott, 1994). Errors can be triggered by technology, its environment, and the conditions, conventions, and procedures for the use of technology and the way these factors interact and challenge the nature of human abilities and limitations. Without a proper evaluation of human abilities and limitations, people can be “set up” to make errors by the failures of designers, managers, and planners to take into account what is known about human behavior and how to apply that knowledge to error-reducing designs.

Errors are generally not catastrophic in their consequences; in terms of everyday life, they are often just merely annoying. However, when errors are introduced into complex systems such as aviation or healthcare, errors at the “wrong time and place” can have catastrophic outcomes. The events outlined above demonstrate just how potentially serious the commission of errors in these complex environments can be. The discipline of human factors provides a possible approach for addressing the problem of errors in complex environments.

Human factors, at its best, employs a systems approach that allows us to address the circumstances that may influence error commission in complex systems. The systems approach takes an entity under consideration and defines it as a system with interrelated subsystems. This method allows the issue under consideration to be analyzed with respect to the contributions of various system components to a particular problem. Specifically, a systems approach allows us to address the manner in which a person's performance might be improved within the existing constraints of the system. It is important to note that this does not mean changing the person to meet the needs of the system, rather, it is the other way around; the system is changed to address and meet the constraints of the people in the system. The human factors discipline utilizes several methods to analyze the various interactions of humans and the tools they use within a system, employing knowledge about the cognitive and physical abilities of people within the system to ensure system safety. This knowledge includes information about cognitive activities such as perception, language, learning, memory, concept formation, problem solving and thinking (Bogner, 1994).

Human Error

What are Errors?

Several relatively recent disastrous and near-disastrous large-scale accidents have illustrated that the adverse effects of human error in many of the technologies upon which we rely may well have far-reaching consequences. The scope of accident consequence can range from one person, as in the death of *Boston Globe* reporter Betsy Lehman from a four-fold overdose of chemotherapy medications, to the thousands who suffered

immediate or long-term effects of radiation exposure at Chernobyl. Error has become a topic of serious inquiry in response to the potential for devastating accidents as a result of human error. The adoption of error as a topic of serious inquiry signals a shift from a “conventional wisdom” perspective to one that takes a systems view of error. Thus, as a phenomenon in its own right, a definition of human error has been established, where “something has been done which was: *not intended by the actor; not desired by a set of rules or an external observer; or that led the task or system outside its acceptable limits*” (Senders & Moray, 1991, p. 25).

Two Perspectives on Error

The “Conventional Wisdom” perspective of human error holds that human beings are intrinsically unreliable, and that when “something” goes wrong, “someone” must have made an error. In other words, the person(s) judged closest to the occurrence of the event must have done “something they shouldn’t have” that resulted in the unintended or “bad” outcome. These errors are assumed to be the result of inattention, laziness, carelessness, and negligence on the part of the human(s) deemed “responsible” (Van Cott, 1994).

The newly emerging systems perspective on error, however, proposes that error is not the result of an intrinsic unreliability, but rather, springs instead from the same psychological processes as successful performance (Reason, 1990; Baars, 1992). According to this view, error and correct performance are essentially diametrical opposites. Reason (1990) likens this natural opposition to a “cognitive balance sheet,” where correct performance and systematic errors are two sides of the same balance sheet. Each cognitive ability that can be entered as an asset contributing to correct performance

is matched by a corresponding deficit toward human error. One ability that seems particularly linked to the occurrence of systematic error is a rapid retrieval system capable of locating relevant items within a virtually unlimited knowledge base. This same ability, however, allows for interpretations of the past and predictions of the future to be shaped too much by the perceived regularities of the past.

Indeed, Baars (1992) argues that many spontaneous action slips (i.e., slips of the tongue) appear to be “habit intrusions.” The more habitual and automatic are the components of an action, the more likely they are to take the place of less habitual parts of an intended act. Reason’s (1984) law of error summarizes this phenomenon:

Whenever our thoughts, words, or deeds depart from their planned course, they will do so in the direction of producing something that is more familiar, more expected and more in keeping with our existing knowledge structures and immediate surroundings, than that which was actually intended. (p.184)

Despite the possible existence of a theorized cognitive “balance sheet” and the deficits it implies, errors are relatively rare events when compared with correct and successful performance. Only a few behavioral mechanisms appear to be responsible for most errors, and although the forms an error can take are limited, it may appear in several different contexts. For example, there appear to be comparable forms of error found in speech, perception, decision-making, problem solving, and action (Van Cott, 1994).

Another hallmark of the systems perspective on human error is the classification of error. Errors have traditionally been broadly classified into two general types: mistakes and slips/lapses. Mistakes are a type of error where the intended actions are not the

correct actions. Senders and Moray (1991) define a mistake as “an incorrect intention, an incorrect choice of criterion, or an incorrect value judgement” (p. 27). Mistakes essentially represent planning failures. They are deficiencies or failures in the judgmental and/or inferential processes involved in the selection of an objective or in the specification of the means to achieve it. Thus, while the developed plan itself is correctly executed, the plan is not the correct plan for the particular situation.

Slips and lapses, on the other hand, generally encompass actions that do not go as intended; a mismatch exists between the reportable intention and the covert performance of that intention (Baars, 1992). Slips and lapses are somewhat different from each other. Slips are actions “not in accord with the actor’s intention, the result of a good plan but a poor execution” (Senders & Moray, 1991, p. 27). One of the defining characteristics of a slip appears to be that it is potentially observable, and is generally revealed as an externalized action not-as-planned (i.e., a slip of the tongue). Lapses, however, seem to encompass a more covert error form than slips and largely involve failures of memory that do not necessarily manifest themselves in actual behavior. A lapse may only be apparent to the person experiencing it, and is akin to having something “slip your mind” (Reason, 1990).

Systems Theories of Error

Given a systems perspective as the foundation of a discussion on error, there are two predominant theoretical treatments on the nature of human error. The first is Rasmussen’s (1987) “Skill-Rule-Knowledge” classification of error, and the second is Reason’s (1987, 1990) “Generic Error-Modelling System” (GEMS). Because GEMS is

founded on Rasmussen's skill-rule-knowledge framework, they will be discussed together.

Rasmussen (1987) adopts a human information-processing point of view for discussing error and argues for a set of cognitive control mechanisms to explain the occurrence of human error. According to this argument, there are three levels of cognitive control, and the cognitive mechanism responsible for error depends on the actor's knowledge about the environment and interpretation of available information in a given situation. These levels of cognitive control are identified and discussed within the skill-rule-knowledge framework. Each of these components (skills, rules, and knowledge) represents a different level of cognitive control within the environment, as well as varying degrees of familiarity with the task and its environment. Knowledge-based behavior represents the least degree of cognitive control and task familiarity, while skill-based behavior represents the highest. As each level of behavior builds on the preceding level, the framework will be discussed from the lowest level (knowledge) to the highest (skill).

At the lowest level of cognitive control, knowledge-based behavior represents behavior that demonstrates a very low level of familiarity with the task at hand. Knowledge-based behavior is predicated on the actor being unfamiliar with a given situation, with no available skills or rules for control of and response to the situation. In this instance, performance becomes specifically goal-oriented and controlled: "the goal is explicitly formulated. . . Then a useful plan is developed - by selection, such that different plans are considered and their effect tested against the goal, physically by trial and error,

or conceptually by means of understanding the functional properties of the environment and prediction of the effects of the plan considered” (Rasmussen, 1987, p. 55).

Reason (1987) claims that failures at the knowledge-based level derive from two basic sources: “bounded rationality” and inaccurate mental models of the problem space. Of these, the latter is relatively self-explanatory. Reason describes bounded rationality as:

. . . a beam of light (the working database of the attentional resource) being directed on to a large screen (the problem space). The difficulties are that the illuminated portion of the screen is very small compared to its total area, that the information potentially available on the screen is inadequately and inefficiently sampled by the tracking of the light beam, and that, in any case, the beam is continually changing its direction in a manner that is only partially under the control of its operator. The beam is repeatedly drawn to certain parts of the screen, while other parts remain in darkness. (p. 77)

Thus, errors arising from bounded rationality may come from attending to the wrong features, giving weight to facts that come to mind while ignoring those not immediately present, and either superficially examining the issues or lingering too much over the details.

In rule-based behavior, the actor is familiar with the situation, but not at a level of performance found in skill-based behavior. In this case, the actor relies upon rules to respond to the situation at hand. The rules utilized may have resulted from a process of “survival of the fittest,” where previous successful experience with particular rules may

make them more likely to be selected than other rules. These rules essentially become heuristics, called upon as needed by the situation at hand. Thus, rule-based behavior represents a “composition of a sequence of subroutines in a familiar work situation . . . controlled by a *stored rule* or procedure which may have been derived empirically during previous occasions, communicated from other persons’ know-how as an instruction or cook book recipe, or it may be prepared on occasion by conscious problem-solving and planning” (Rasmussen, 1987, p. 54).

Mechanisms that shape errors at the rule-based level generally bias a problem solver to decide that a particular rule is relevant to the situation, when in fact, it is not. Mind-set is one such mechanism at the rule-based level; mind-set represents a mechanization of thinking where one rule is applied consistently to all problems (e.g., when all you have is a hammer, everything looks like a nail). Because this rule has been successfully applied in the past, the actor applies it again, despite the fact that it may not be warranted by the situation. Availability is another mechanism that affects rule-based problem solving and behavior. In this case, those heuristics that come to mind first are generally preferred for utilization, even if they are not always applicable or relevant. Availability and mind-set are somewhat related, as the rules that come to mind first will most likely be those that have successfully worked in past situations (Reason, 1987, 1990).

Finally, skill-based behavior is relatively automatic and is defined as “sensori-motor performance during acts or activities which, following a statement of intention take place without conscious control as smooth, automated and highly integrated patterns of

behavior” (Rasmussen, 1987, p. 53). Thus, skill-based behavior demonstrates relative mastery of the task being performed and implies that the actor needs to perform little or no monitoring of the task once it has been intended and initiated.

Reason (1990) claims that slips and lapses at the skill-based level are shaped by several psychological and situational mechanisms. Four primary factors are identified: the recency and frequency of use, environmental control signals, shared schema properties, and concurrent plans. According to the recency and frequency of use factor, the “more recently and frequently a particular routine is set in motion and achieves its desired outcome, the greater its likelihood of recurring unintended as a slip of habit” (Reason, 1987, p. 72). In other cases, a familiar environment will elicit certain action routines that do not correspond with the current intention. With shared schema properties, a given action schema will increase the activation of other schema possessing shared or similar features. Due to inattention (i.e., absentmindedness) or other factors, one of the shared schema may take over the action routine, resulting in an unintended outcome. Finally, concurrent plans may result in slips and lapses at the skill-based level of behavior. The appearance of concurrent plans “can take the form of *blends* in which two active plans become intermingled in the same action sequence. . . Alternatively they can involve *reversals* in which the right actions are applied to the wrong objects” (Reason, 1987, p. 74).

Training on a particular task is theorized to follow the skill-rule-knowledge framework. Initially, task performance represents knowledge-based behavior, with little control over and familiarity with the task. With practice, however, the actor’s control

over task performance moves from the knowledge-based level up through rule-based, to the relatively automated performance at the skill-based level. This progression through the performance framework brings to bear different cognitive control mechanisms at each level, as task performance improves. Accordingly, the various error forms that can manifest during task performance change as the cognitive control mechanisms change through the progression.

While Rasmussen's skill-rule-knowledge framework and Reason's GEMS theory of human error provide a scientific framework for discussing error, they lack the specificity necessary to predict the occurrence of systematic error. This is not a trivial concern. In some complex mechanical systems, a practice known as reliability-centered maintenance may be followed. According to reliability-centered maintenance, certain metrics such as failure rates and time to failure are calculated for various system components. Because of the relative importance of certain components to the system, "perfectly good" components are removed and replaced before they are expected to fail because the system cannot afford to have them fail (D. Serig, personal communication, January 26, 2001). The same consideration, however, is seldom applied to the human "components" of the system. Thus, while "to err is human," there seems to exist an unrealistic expectation that humans are the singular, one-hundred-percent reliable component of any system; we do not expect human error, nor do we tend to make allowances for its occurrence within complex systems.

What Rasmussen (1987) and Reason (1990) do provide is a confirmation that people will indeed make errors, and that various types of errors will be made dependent

on the actor's particular level of skill and familiarity with a task. What their theories cannot do, however, is say precisely that "due to the nature of the task at hand, people will be more likely to commit errors at x point in the process, rather than at y or z points." The true value of a theory of human error is to provide the ability to predict the likelihood of error at various points in a process, just as various metrics are used to predict failure rates and mean time to failure for mechanical components of a system.

Postcompletion Error

Recognizing a lack of specificity in existing theories on human error, Byrne and Bovair (1997) turned to a computational theory for a specific type of error: postcompletion error. Postcompletion errors are a type of procedural error that occurs in the execution of a routine procedure. According to the current thinking in error theory outlined above, most procedural errors are described in terms of rule- and knowledge-based errors, where either an incorrect rule is selected for executing the procedure, or gaps in knowledge regarding the correct procedure lead to a flawed execution (Rasmussen, 1987; Reason, 1987, 1990). Postcompletion errors, however, represent a special case of procedural error, where the actor possesses the correct knowledge necessary to execute a task. Not only does the actor possess this knowledge regarding task execution, but the task is one that is frequently and routinely performed correctly. Thus, postcompletion errors seem to comprise a distinct class of common errors where the actor possesses the knowledge required to perform the task correctly, yet manages to occasionally fail in correctly executing the task.

Specifically, Byrne and Bovair (1997) define postcompletion errors as errors in which the task structure demands “that some action . . . is required *after* the main goal of the task . . . has been satisfied or completed” (p. 32). Thus, in the events described at the beginning of this paper, failing to sign the admitting orders was an action required after the main goal of making the orders was satisfied, and breaking the seal in the IV bag was an action that was required after the main goal of preparing the IV bag had been completed. More commonplace examples of postcompletion error include forgetting to reset the photocopier to default settings after a custom job, and sending an e-mail message without an indicated file attachment.

General Theory of Postcompletion Error

Byrne and Bovair (1997) hypothesized that postcompletion errors arise from goal forgetting that is caused by an excessive working memory load, despite the presence of the correct procedural knowledge necessary to perform the task at hand. Goal forgetting simply means that the actor forgot to perform some portion of the task in question. Quite often, the actor typically knows the correct procedure for the task in question. Thus, the actor does not forget *what to do*, he simply forgets to do *it*. Had the actor forgotten the correct procedure (what to do), postcompletion error would be linked to long-term memory. However, since the actor simply forgets to do some part of the procedure, working memory is a much more likely candidate for influencing the occurrence of postcompletion error.

Byrne and Bovair (1997) claim that “postcompletion omissions can be explained in a relatively straightforward way as goal loss from working memory” (p. 38).

According to this claim, errors are described in terms of a goal hierarchy. Thus, a goal resident in working memory supplies activation to its subgoals. These subgoals are also resident in working memory at the same time as the primary goal. These subgoals receive activation only so long as the parent goal is active and resident in working memory.

When a parent goal is satisfied, it is eliminated from working memory. Any subgoal of that parent goal that remains in working memory loses activation as the satisfied parent goal is eliminated from working memory. For example, in the IV bag example, the main goal is to “start IV.” There are several subgoals associated with this main goal, including “remove old IV and tubing,” “spike and hang new bag,” “set infusion rate,” and “break seal between drug and diluent.” Thus, in this example, when the “set infusion rate” subgoal is met, the main goal of “start IV” is essentially met. Any subgoals that were not already satisfied, for example, the “break seal between drug and diluent” subgoal, lose activation once the satisfied parent goal of “start IV” is eliminated from working memory. Under conditions of high working memory load, the loss of activation from this subgoal (“break seal between drug and diluent”) increases the likelihood that this subgoal will not be satisfied.

In cases where the working memory load is high, the loss of activation from the parent goal may lead to the loss of any as-yet unsatisfied subgoals of the parent goal. If the parent goal falls below threshold in working memory too soon, unsatisfied subgoals will not receive enough activation to reach threshold. If the parent goal remains active for some time, however, the associated subgoals will safely reach threshold and be executed. The effects of a higher working memory load are assumed to be associated with a faster

decay or displacement of information from working memory. Thus, a higher working memory load should be associated with the commission of more postcompletion errors (Byrne & Bovair, 1997).

Computational Model of Postcompletion Error

This general theory of postcompletion error was developed as a computational model in the Collaborative Activation-based Production System (CAPS). The value of developing a computational model of postcompletion error lies in its ability to provide greater specificity of the phenomenon than verbal theories (i.e., Rasmussen's and Reason's theories). Furthermore, a computational model facilitates the examination of interactions between mechanisms, and provides a test of the theory's internal coherence. The improvement of a computational theory over a verbal model is the ability to provide a specificity of prediction. As noted earlier, it is difficult to quantify the occurrence of errors given Rasmussen's and Reason's treatment of error. A computational model, however, can facilitate quantitative prediction, due to the necessary precision of the theory's mechanisms and parameters (Byrne, 1998).

Byrne and Bovair (1997) created a unique experimental task that was designed specifically to be instantiated as a CAPS model. CAPS is described as a production system that has two memory systems: "a long-term memory and a working memory. Working memory contains elements such as goals and propositions. . . . each element in working memory has some continuous activation value associated with it" (Byrne & Bovair, 1997, p. 39). The model's working memory capacity is represented as a ceiling

of the total amount of element activation that can be present in the system at any given time.

One particular step in the created task was designed to be a postcompletion step. Both a control version and a postcompletion version of the task were modeled. Running the two models resulted in different outcomes, as predicted by the general postcompletion theory. In cases where the goal activation ceiling in the model was high (equivalent to high working memory capacity/low memory load), both versions of the model executed the task without error. When the activation ceiling was lowered (equivalent to low working memory capacity/high memory load), a critical postcompletion step was omitted in the postcompletion version of the model. Thus, with a high activation ceiling, the error never occurred, and with a low activation ceiling, the error always occurred in the postcompletion version of the task. This supports the claim of a relation between working memory capacity and postcompletion error (Byrne & Bovair, 1997).

A series of empirical laboratory experiments were designed to determine whether such postcompletion errors could be generated experimentally and if the experimental results would support the findings from the CAPS model. Participants performed the same task as modeled by CAPS. The experimental results supported those produced by the CAPS model. When working memory was not taxed, postcompletion errors were extremely rare. However, when working memory demands were increased, either through task complexity or an external load, the postcompletion error became much more common (Byrne & Bovair, 1997). Thus, the computational model of postcompletion error was able to predict the relative occurrence of a particular error at a particular place

within a procedure. Not only did it predict the systematic occurrence of an error, it tied error commission to a very specific phenomenon, namely, working memory load and capacity. Given the power of knowledge that such computational models provide, it should be possible to predict the circumstances under which certain types of errors could occur and affect system operations.

Organizational Responses to Error

In the course of everyday life, when we catch ourselves committing an error (such as a slip of the tongue), we are quite likely to be “tough” on ourselves, berating the perceived inadequacies that led to making the error. Organizations, when faced with the errors of its employees, tend to engage in the same type of behavior that we, as individuals, do. Reason (1994) calls this response the “blame trap.” According to this phenomenon, those fallible individuals that were in direct contact with the vulnerable parts of the system that failed are blamed for that failure. This response is described as one that is natural, universal, emotionally satisfying, and legally convenient. The pressure to place blame, much of which comes from entities external to the organization, often leads to ineffective countermeasures, such as disciplinary action, demands to be “more careful,” re-instruction, and writing new procedures to proscribe those actions implicated in the system failure.

Several well-known psychological factors are assumed to influence the need to assign blame in the face of the catastrophic consequences of an error. These factors include the fundamental attribution error, the illusion of free will, and the similarity bias (Reason, 1994). According to the fundamental attribution error, we tend to attribute

another individual's behavior to the dispositional factors of that individual, while the individual in question tends to attribute his own behavior to various situational factors. The illusion of free will follows a line of thought parallel to that of the fundamental attribution error: when we find that someone has committed an error, we tend to believe that the individual deliberately and knowingly chose an error-prone course of action. Finally, the similarity bias assumes a symmetry of magnitude between causes and consequences. Thus, in the face of horrific man-made catastrophes, it is often assumed that some equally egregious act of incompetence or irresponsibility was the primary cause.

Model of Organizational Discipline

Despite a relative lack of research in the area of organizational response to error, Arvey and Jones (1985) developed a model of organizational discipline that can be applied generally to instances of "inappropriate" organizational behavior. According to this model, the function of discipline within an organization is to operate as a mechanism for the direct control of behavior. A system of organizational discipline helps establish and maintain an "organizational boundary system," by which employees learn what behaviors will and will not result in aversive consequences.

Organizational discipline is defined as the "presentation of an aversive event or the removal of a positive event following a response which decreases the probability of that response" (Kazdin, 1975; as cited in Arvey & Jones, 1985, p. 367). Discipline can occur under at least two sets of circumstances. Under the first set of circumstances, a primary aversive event is paired directly with a behavior, and may involve such aversive

events as termination or demotion. In the second set of circumstances, a conditioned, or secondary, aversive event may become aversive through repeated pairings with a primary aversive event. This secondary punisher usually operates by warning of impending aversive consequences if a particular response or pattern of behavior persists. Many of the disciplinary events in organizations fall under this second type; for example, reprimands and verbal warnings are often issued before resorting to such primary aversive events as termination or demotion.

Arvey and Jones (1985) propose that there is a general pattern of events that takes place during a disciplinary episode. First, a specific inappropriate action or pattern of behavior occurs and is perceived by or brought to the attention of the individual's supervisor. The supervisor must then make a decision regarding whether or not an infraction against the organization's policies has occurred, and if so, whether to take disciplinary action. The resulting action is then perceived and responded to by the employee. The following provides a more detailed account of Arvey and Jones' model of organizational discipline.

The identification of a policy violation in the organizational setting may either be direct or indirect. In a direct observation sequence, a supervisor actually witnesses the occurrence of a specific behavior. In the case of indirect observation, a "flag" or indicator of some other sort suggests that a violation has occurred and directs attention toward the probable cause of the violation. Once an event or pattern of behavior has been identified as inappropriate, a decision must be made about what, if any, action should be taken.

Each of these decisions, the decision to act, and what action to take, will be discussed separately.

In terms of deciding to take action in response to inappropriate behaviors, certain employee behaviors are more likely than others to be punished. In particular, Arvey and Jones (1985) identified several behaviors that are likely to be punished, including absenteeism, poor performance (incompetence, negligence, and poor workmanship), violation of safety rules, and unethical behavior. The decision to take disciplinary action, however, does not always follow from the direct observation of inappropriate behavior. In many cases, organizational leaders tend to make attributions regarding the perceived causes of employee behavior, and make decisions based on these attributions, rather than in response to the actual behavior. Thus, managers are more likely to fire, suspend, or demote an employee when they attribute the employee's poor performance to a lack of motivation, interest, or drive, rather than to a lack of ability or technical competence. Leaders make two primary attributions regarding employee behavior: whether the employee's behavior was due to internal or external factors, and whether the expressed behavior was under the control of the employee. To the extent that inappropriate behavior is viewed as due to internal causes (i.e., a lack of effort or motivation) and under the employee's control, organizational leaders are more likely to respond in a punitive manner. Attributions to an external cause, however, tend to prompt organizational leaders to focus on changing the situation that brought about the inappropriate behavior.

Once a supervisor has identified that a policy violation has indeed occurred, and has made an attribution of cause that is consistent with the use of punishment, there is

still a broad range of actions that could be taken. There appears to be two major classes of variables that seem to affect a supervisor's choice in the use of rewards or punishment: contextual variables and subordinate factors (Arvey and Jones, 1985). The factor of contextual variables takes into account the fact that a supervisor's behavior occurs in the broader context of the task and organizational environments in which they work. Thus, supervisors are more likely to use punishment when they have a relatively large span of control, possess limited or restricted reward power, and encounter explicit, but not excessively harsh, organizational policies that encourage punishment. Another set of contextual variables that may need to be taken into account are situational and task variables. The nature of the job or task on which an employee is working may be a factor in influencing the nature of the discipline received. For example, some jobs and tasks may be more critical or important than others because of their possible impact on the safety of others, or because they are more vital to the production of the organization's product. Policy violations by these employees may be dealt with more strictly than policy violations from employees working in less critical jobs. Subordinate characteristics may also play a role in the supervisor's selection of discipline. In particular, individual personality characteristics may influence the course of action taken. Often, individuals self-administer rewards and punishments in ways that are consistent with their own self-efficacy; this self-administration may often offset many aspects of the external reward/punishment process.

Once the decisions to act, and what action to take, have been made, several variables are salient in the actual application of discipline. These variables include the

timing, intensity, schedule of punishment, provision of a rationale, and availability of alternative responses. Disciplinary episodes where the administration of punishment occurs soon after the infraction occurs, is administered consistently among and within employees, is accompanied by a clear explanation for the discipline, and applies a penalty that is not unduly harsh, are perceived as being much more effective than episodes where punishment is administered haphazardly and with little explanation. Another factor in the administration of discipline is whether the punishment is applied through a relatively formal system specified by the organization, or whether it is applied through informal mechanisms. Formal methods of punishment include sanctions such as days off without pay, and written reports/reprimands, while informal methods include yelling at employees, ridicule, and assigning onerous tasks.

A variety of outcomes might be influenced by the administration of discipline. The outcome of immediate concern focuses on the “target” behavior. Arvey and Jones (1985) report the results of several studies investigating the use of punitive systems to control specific behaviors. In general, these systems tended to have a positive effect on the target behavior when disciplinary tactics were used in combination with a reward system of some type. For example, in a study of absenteeism, Kemper and Hall (1977; as cited in Arvey and Jones, 1985) found that a mixed-consequence system that rewarded employees for good attendance and progressively disciplined employees for excessive absenteeism substantially reduced absenteeism at two organizations.

Organizational Discipline in Healthcare

While Arvey and Jones (1985) provide a theoretical model for the administration of punishment within organizations, what evidence exists that reveals actual organizational responses to error? In one particular organizational environment, namely healthcare, there is some evidence that punitive disciplinary measures are often taken in response to the commission of error. According to the mores of the culture of medicine as proposed by Leape (1994), physicians are socialized in medical school and residency to strive for error-free practice. A strong emphasis is placed on perfection, both in the diagnosis and treatment of patients. The message behind this emphasis is simple: mistakes are unacceptable in everyday hospital practice. Physicians are expected to function without error, and this expectation is translated into a need to be infallible. Thus, physicians come to view error as a failure of character, where someone was not “careful enough,” or did not “try hard enough.” These beliefs further lead to a common reaction by physicians to error - there can be no error without negligence.

Leape (1994) calls the medical approach to error prevention a perfectibility model: “if physicians and nurses could be properly trained and motivated, then they would make no mistakes” (p. 1853). The methods used to reach this goal of perfection are training and punishment. Training is directed at teaching people to do the right thing; in nursing, this means an emphasis on the rigid adherence to protocols, while in medicine, there is an emphasis less on rules and more on knowledge. Punishment generally occurs through social censure or peer disapproval. The professional cultures of nursing and medicine tend to use blame to motivate “proper” performance. When errors do occur,

they are viewed as someone's fault, caused by a lack of sufficient attention or lack of caring enough to make sure they are correct. The medical approach to error prevention is largely reactive. In healthcare, errors are usually only discovered when there is an incident - the untoward effect or injury to a patient. Corrective measures are then directed toward preventing the recurrence of a similar error, often by attempting to prevent the *individual in question* from making a repeat error. The underlying causes of an error are seldom explored.

Supporting Leape's (1994) assessment of error in medicine are several studies and anecdotes from healthcare regarding the administration of discipline in response to error. In a study of radiopharmaceutical misadministrations, the U. S. Nuclear Regulatory Commission (NRC) compiled a database of misadministration reports from its licensees. A radiopharmaceutical misadministration is a medical error that involves the use of a material that contains a small amount of radioactivity for diagnostic or therapeutic medical purposes. The NRC database contained 902 reports, collected over the years 1989 and 1990, and included information regarding organizational responses to the reported incidents (Serig, 1994). Table 1 summarizes the corrective actions that licensees indicated that they had taken (D. Serig, personal communication, November 21, 2000). Licensees could take a number of actions, therefore, the total of corrective actions in Table 1 exceeds the number of reported events.

Table 1

Corrective actions taken by NRC licensees in response to radiopharmaceutical misadministrations

Corrective Action	Total
Implement new procedure	245
Re-instruct personnel	510
Reprimand personnel	146
Improve supervision of personnel	62
No action	36
Other	176

A study of emergency medicine residents revealed some of the reasoning behind why disciplinary actions are pursued in response to the occurrence of error. A survey distributed to emergency medicine residency directors revealed that they reported the observation of errors to the emergency medicine residents for many reasons, including a perceived need to help the residents improve performance, change behavior, and enhance perceptions of personal responsibility (Hobgood, Ma, & Swart, 2000).

Another study investigated organizational factors that may account for variations in drug error rates across hospital units (Edmondson, 1996). It was found that high reported error rates were strongly associated with high scores on nurse manager “direction setting,” coaching, perceived unit performance outcomes, and the quality of unit relationships. It was hypothesized that in certain units, leaders may have established a climate of openness that facilitated the discussion of error, thereby having an effect on detected error rates. From comments made by the nurses and nurse managers on the investigated units, various organizational patterns of administering discipline could be

discerned. For example, in a unit identified as “Memorial 3,” the nurse manager indicated that the unit prided itself on being clean, neat, and professional-appearing. Nurses in this unit reported that making a mistake meant that they would get in trouble, and that the nurse manager often made the nurses feel degraded and persecuted when they were found to have made an error. An apparent fear of the responses to error resulted in a low overall rate of reported drug errors. The climate in another unit, “Memorial 1,” however, was drastically different. In this unit, the nurse manager indicated that she expected a certain amount of error to occur and that a non-punitive environment was essential to dealing with errors productively. In many cases, she admitted that nurses were harder on themselves than she would have been when an error was committed. Nurses indicated that the unit did indeed support a non-punitive environment, where there was no punishment for the commission of errors, and if there was a problem, the nurse manager stood up for her employees. In contrast to the Memorial 3 unit, the overall reported drug error rate in the Memorial 1 unit was relatively high. The corresponding outcome of this high reported error rate, however, was that the underlying causes of the errors were discussed and steps taken to address these underlying causes.

In conjunction with the results reported by these studies regarding organizational responses to error, there is also anecdotal evidence to support the patterns of punishment in healthcare through organizational discipline. A recent television series entitled *Why Doctors Make Mistakes* (Day, 2000a) claimed that “Deep within the culture of medicine, error and incompetence are seen as inseparable. Get rid of the failing doctor or nurse -

get rid of the problem.” This statement is a direct affirmation of the assessment made by Leape (1994) of the culture of medicine and its response to error.

Two anecdotes from the *Why Doctors Make Mistakes* series provide examples of both the informal and formal administration of discipline in response to error in healthcare. In the first case, Dr. Paul Barach, an anesthetist at Massachusetts General Hospital in Boston, related an event that occurred while he was a junior doctor. Dr. Barach was asked to perform a procedure that he had never before performed. A senior colleague was supervising the procedure, but was called away on an emergency, leaving Dr. Barach to complete the procedure, which involved inserting a needle just under the clavicle to find a vein that was approximately one inch under the skin. During the procedure, the patient began to have trouble breathing - Dr. Barach had inadvertently pierced the patient’s lung. According to Dr. Barach, “the senior resident came by and he started yelling at me, why I did it . . .” (Day, 2000b). This example illustrates the administration of an informal punishment tactic (yelling) as described by Arvey and Jones (1985).

A more formal disciplinary action was carried out in another case. Michele Johnson, a nurse, was asked to resign and had her nursing license suspended after she administered the incorrect medication to a patient in acute congestive heart failure. The patient died as a result of this misadministration (Day, 2000a). In this case, Ms. Johnson incorrectly picked up a vial of potassium chloride, instead of the intended vial of lasix. These two medications were stored next to each other on the ward, and are often administered together. Potassium chloride is usually administered over a long period of

time in dilute form. When the “strong” potassium chloride that Ms. Johnson retrieved was injected into the patient, the patient suffered a massive cardiac arrest.

The Underlying Assumption of Organizational Responses to Error

The underlying theme in the previous discussion is that the administration of discipline is predicated on the assumption that the inappropriate behavior that needs to be corrected is under the direct control of the individual. This assumption is essentially another rendition of the “conventional wisdom” perspective on error (humans are lazy, inattentive, careless and negligent) and the “blame trap,” where it is assumed that the person closest to the accident deliberately and knowingly carried out an error-prone action. In many cases, the behaviors that organizations tend to discipline on a regular basis (i.e., absenteeism, fraud, theft, egregious negligence, and unethical behavior) do tend to fall under the direct control of the individual in question; some decision was made on the part of the employee to behave in the inappropriate manner in question. Because these are often the types of issues that organizations deal with on a regular basis, organizations appear to have developed a general model of discipline that is based on the idea that all employee behavior is under the direct control of the individual. It is therefore logical, from the perspective of the organization, to assume that errors are also the result of a direct, conscious process on the part of the employee. According to the model developed by Arvey and Jones (1985), such an attribution of error to internal causes results in generally more punitive actions toward the employee, in order to “motivate” the employee to change his behavior to “proper” behavior patterns. The pervasiveness of this belief and approach to disciplining error is well illustrated in the *Blondie* newspaper

comic strip shown in Figure 1. Leape (1994) so much as admits that the culture of medicine views the commission of error as an internal flaw and that it essentially tries to “blame and train error out” of physicians and nurses.



Figure 1. Evidence in the popular press for the belief that the occurrence of error is under the control of the individual and that punitive methods “motivate” employees to error-free behavior (Young & Lebrun, 2001).

The assumptions of the “conventional wisdom” model of error and the “blame trap,” however, are seldom true. Byrne and Bovair (1997) found that postcompletion errors had a primary causal agent in working memory. The capacity of working memory is almost certainly *not* under an individual’s direct control, and therefore, the individual cannot control the likelihood of error under certain conditions such as high working memory load. Furthermore, individuals certainly do not want to commit errors; there is no “sense of decision” to commit an error that is associated with other inappropriate behaviors such as absenteeism and unethical behavior. The model of a non-punitive environment in the Memorial 1 nursing unit illustrates the difference that a shift in attribution can make on the occurrence of error (Edmondson, 1996). In this unit, error was recognized as a normal occurrence that was not the “fault” of the responsible nurse.

As a result, the underlying causes of errors were identified and addressed, leading to a cohesive, motivated nursing unit. In addition, this unit illustrated a point made in the model of organizational discipline developed by Arvey and Jones (1985). Arvey and Jones point out that the personality characteristics of employees may influence the selection of a disciplinary action. Because employees often self-administer punishment consistent with their own sense of self-efficacy, external punishments may not have the intended effect. The nurse manager on Memorial 1 recognized this, when she mentioned that the nurses were harder on themselves than she would have been. Thus, she essentially allowed this self-administration to stand as all the punishment a nurse received, while the unit worked together to discover the underlying cause of an error.

If one is to accept the argument that the commission of errors is not under the direct control of the individual, then one must adopt a systems perspective of error. According to this perspective, error occurs as a consequence of human interactions with complex systems, where the demands of the system challenge human abilities and limitations. The response to error is generally to redesign the system to be consistent with and accommodate the abilities and limitations of the people interacting with the system. This assessment is consistent with another point made by Arvey and Jones (1985) in their model of organizational discipline: when the cause of an error is attributed to external causes, organizational leaders are more likely to investigate and change the environment that prompted the error. One of the anecdotes described earlier relates to this concept. In the case of Ms. Johnson, recall that she inadvertently retrieved a vial of potassium chloride, which was stored next to the vials of lasix, which she had intended to retrieve.

Recognizing that the potential for error exists when potassium chloride and lasix are stored together, another unit in the same hospital as Ms. Johnson worked stored the two drugs on physically separate sides of the storage room. Thus, while the “conventional wisdom” approach attributes error to internal qualities of the individual, a systems approach argues that the cause of an error almost always lies in a combination of external causes and their effect on internal cognitive processes.

Thus, it is suggested that under the current perspective of “conventional wisdom,” organizations tend to develop an incorrect attribution of the cause of error, and therefore select inappropriate corrective actions to address the problem. Because organizations assume that error is under the control of the individual, methods such as reprimands and re-instruction are often administered to “motivate” individuals to make different choices regarding their behavior. Essentially, an organizational approach is applied to what amounts to a cognitive problem. This paper, however, argues for a different approach to error - a systems approach that attributes the occurrence of error to a combination of external factors and the effect of these external factors on internal cognitive processes.

EXPERIMENTS

Two experiments were conducted to assess the influence of several organizational responses to task performance on the commission of a very particular type of error - postcompletion error. The first experiment served three primary purposes: to replicate the findings of Byrne and Bovair (1997) regarding the occurrence of postcompletion error, to assess the effects of learning on task performance over time, and to establish the protocols and methodology for the second experiment. As mentioned previously, Byrne and Bovair found that an increased load on working memory resulted in an increase of the commission of postcompletion errors. The second experiment was directed at trying to manipulate various organizational responses to task performance and their effect on the commission of postcompletion error. In this second experiment, so-called “typical” organizational responses to task performance were used to manipulate an individual’s motivation for task performance and the occurrence of subsequent errors. These methods included issuing reprimands, re-instructing individuals on “correct” behavior, and praising individuals with good performance. Methods such as issuing reprimands to, or re-training, individuals who make errors essentially amount to exhortations to “pay attention, and don’t do it again.” Praising individuals for good performance has the similar effect of merely “thanking” individuals for paying attention. As these methods are primarily motivational in nature, they assume that the an individual has made a very deliberate choice about whether they choose to commit errors. However, as discussed earlier, these responses should have relatively little effect on the commission of error that has at its root a cognitive limitation (i.e., working memory capacity). Redesigning an

error-prone procedure from a systems perspective that addresses the underlying challenge to cognitive limitations, however, should result in a subsequent decrease in the occurrence of systematic error.

Following the precedent established by Byrne and Bovair (1997), novel tasks were created specifically for the experiments to assess the occurrence of postcompletion errors. As in the original Byrne and Bovair experiment, these tasks were set in the science-fictional “Star Trek” universe. Participants engaged in four tasks that were introduced as part of qualification for “Bridge Officer’s Command School.” It was hoped that this fictional setting would be both engaging and entertaining to the participants. Refer to Appendix A for the “Operations Officer Qualifying Exams: Cadet Manual” that established the experimental setting for the participants.

Participants engaged in tasks at four different bridge “stations:” tactical, transporter, conn, and operations. Qualifications took place over a four week period, with a one week period between sessions. Participants trained on the tasks during the first session, and were tested on their performance during the remaining three sessions. This differs from the original Byrne and Bovair experiment, where participants only participated for two weeks, receiving training during the first week, and being tested on their performance during the second week. The extended testing period in the current set of experiments was intended to facilitate assessment of the potential for change in error occurrence over time, as well as to facilitate the execution of the experimental manipulation of response to task performance.

As mentioned above, tasks at four bridge stations were used to assess participant performance: tactical, transporter, conn, and operations. The tactical and transporter tasks were taken from the Byrne and Bovair (1997) experiment. In that experiment, tactical was referred to as the “phaser” task. In the Byrne and Bovair experiment, there were two versions of the phaser and transporter tasks - a control version and a postcompletion version. Participants received a task mix of the control version on one task and the postcompletion version of the remaining task. In the current research, however, only the tactical (“phaser”) task had both the control and postcompletion versions.

The tactical task was a relatively complex one and required participants to execute a procedure that involved several subgoals and a tracking task. The difference between the two versions of the tactical task occurred at the end of the procedure, where the main goal (destroying a target) was satisfied in the postcompletion version before the entire sequence of the task procedure was complete. The specific point of differentiation occurred at the “Turn Off Tracking” step in the tactical procedure - the next to last step in the sequence. In the postcompletion version, participants were informed that they had destroyed the target (thereby satisfying the main goal of the task) before they turned off the target tracking. In the control version of the task, however, participants did not learn the outcome of their actions (whether or not the target was destroyed and the main goal therefore met) until they had executed the “Turn Off Tracking” step. This difference is illustrated in Figures 2a, 2b and 3. Figures 2a and 2b are diagrammatic representation of the task structure for both versions of the phaser task (control and postcompletion,

respectively). Figure 3 contains the tactical task screen display and a summary of the procedure for both the control and postcompletion versions of the tactical task.

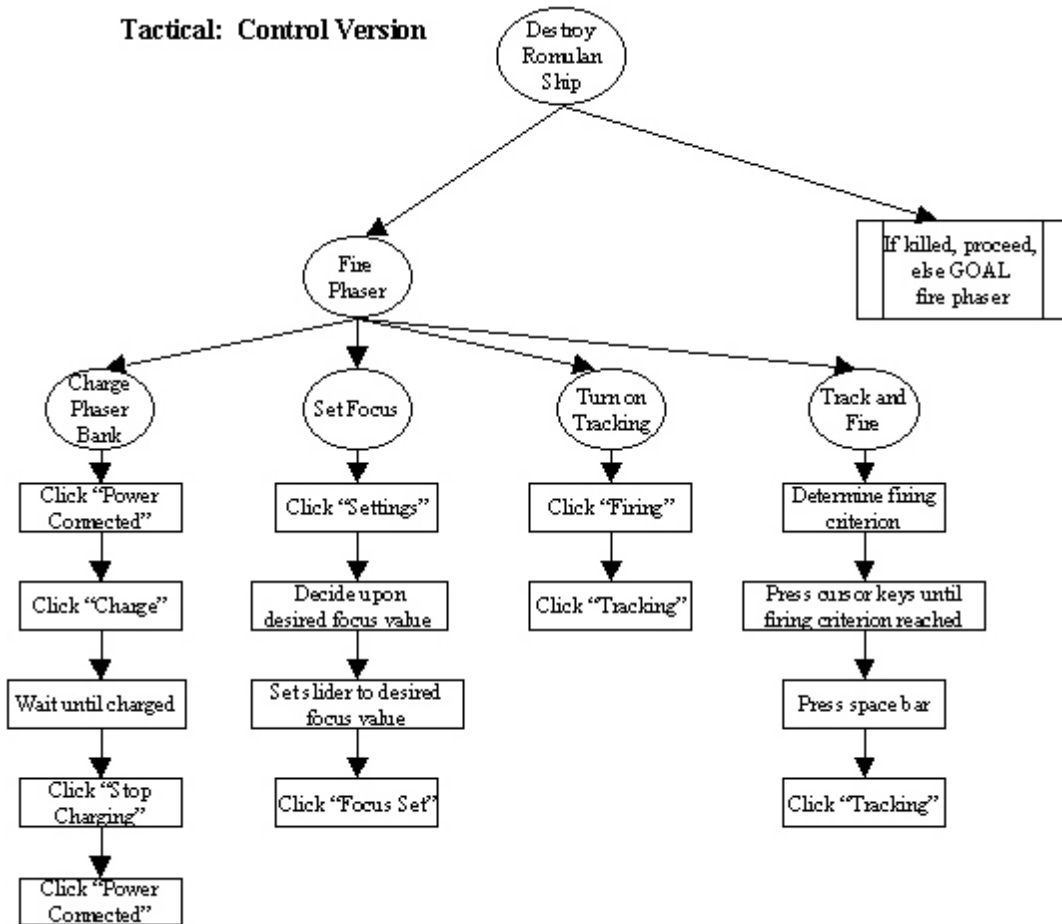


Figure 2a. Tactical task structure - control version.

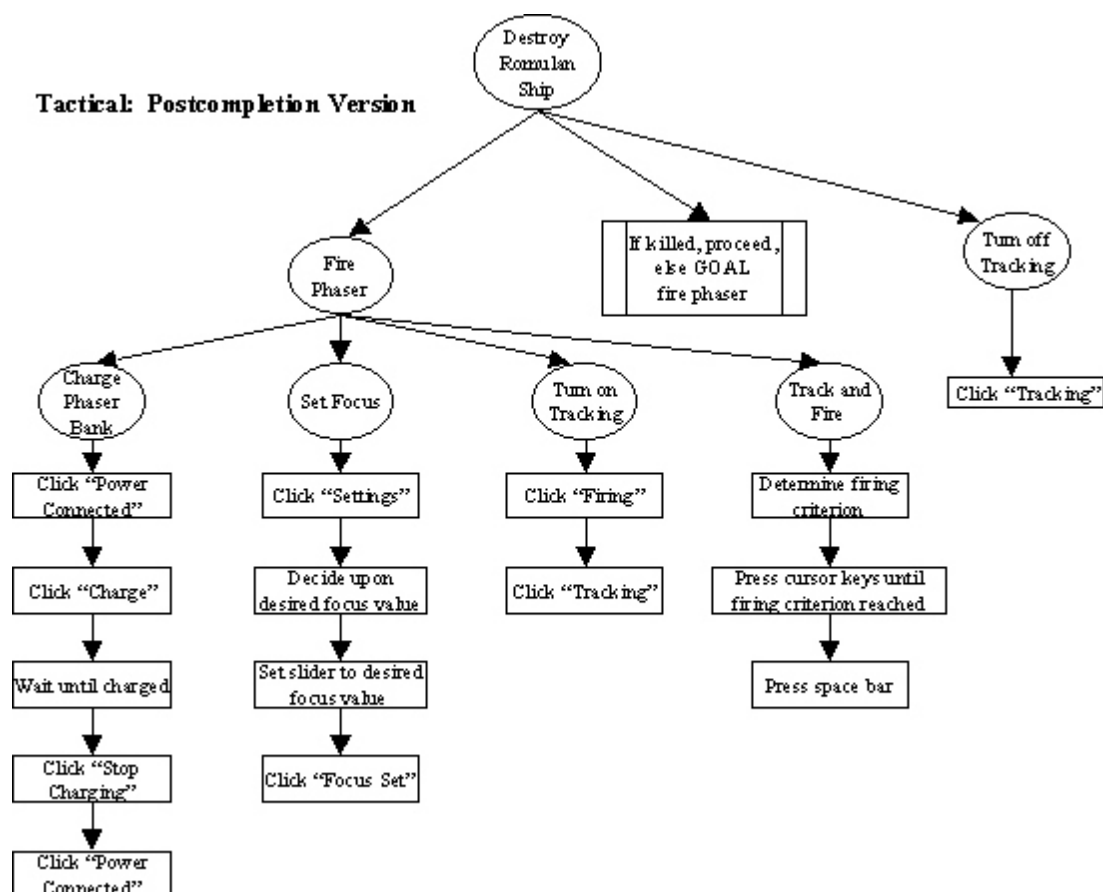
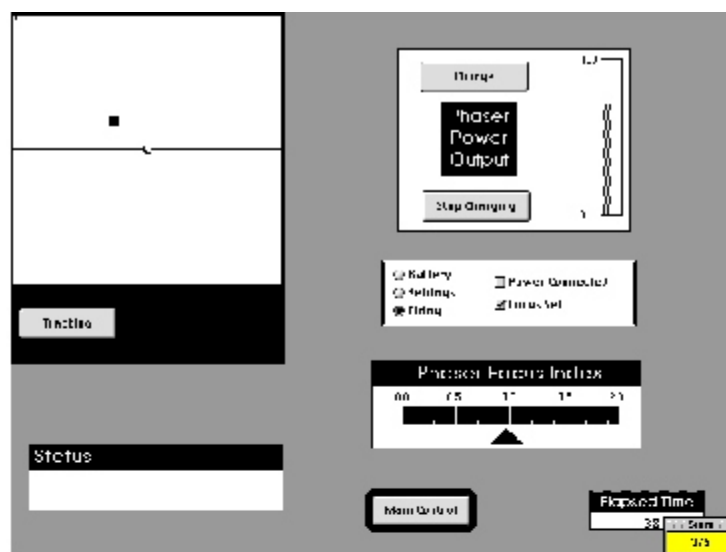


Figure 2b. Tactical task structure - postcompletion version.



Summary of Phaser Bank Operating Procedures
(Postcompletion Version):

Step 1. Charge the phaser:

- Click "Power Connected"
- Click "Charge"
- Wait until phaser charges the appropriate amount
- Click "Stop Charging"
- Click "Power Connected"

Step 2. Set phaser beam focus:

- Click "Settings"
- Adjust location of slider to desired focus
- Click "Focus Set"

Step 3: Track the target:

- Click "Firing"
- Click "Tracking"
- Use arrow keys to adjust location of the target indicator

Step 4: Fire the phaser:

- Press the space bar
- **Determine if the target has been destroyed**
- **If so, click "Tracking"**
- If not, return to Step 1

Then return to Main Control

Summary of Phaser Bank Operating Procedures
(Control Version):

Step 1. Charge the phaser:

- Click "Power Connected"
- Click "Charge"
- Wait until phaser charges the appropriate amount
- Click "Stop Charging"
- Click "Power Connected"

Step 2. Set phaser beam focus:

- Click "Settings"
- Adjust location of slider to desired focus
- Click "Focus Set"

Step 3: Track the target:

- Click "Firing"
- Click "Tracking"
- Use arrow keys to adjust location of the target indicator

Step 4: Fire the phaser:

- Press the space bar
- **Click "Tracking"**
- **Determine if the target has been destroyed**
- If not, return to Step 1

Then return to Main Control

Figure 3. Tactical task screen and step summary for both the control and postcompletion versions.

The transporter task only appeared in a “control” version (no postcompletion step) in the current research. The task structure for this task was identical to the control version of the tactical task, and is shown in Figure 4. Figure 5 contains the transporter task screen display and a summary of the transporter task procedure.

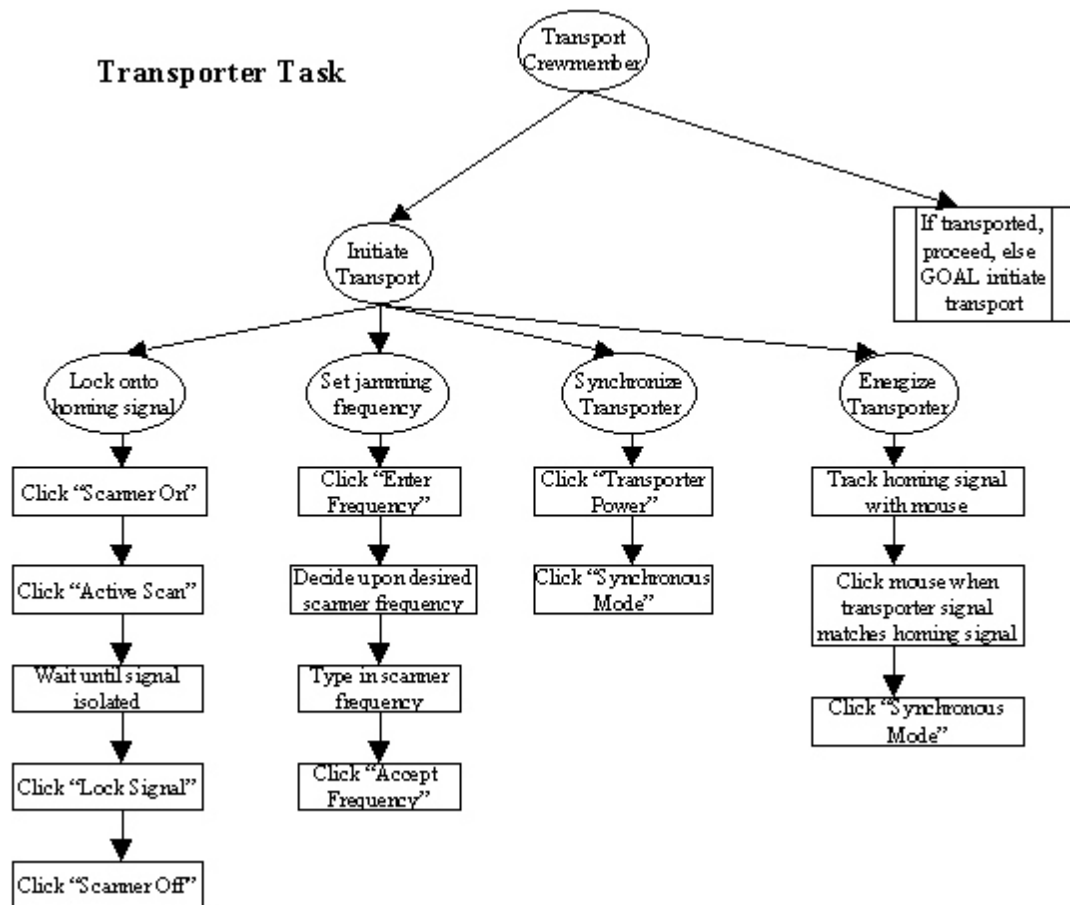
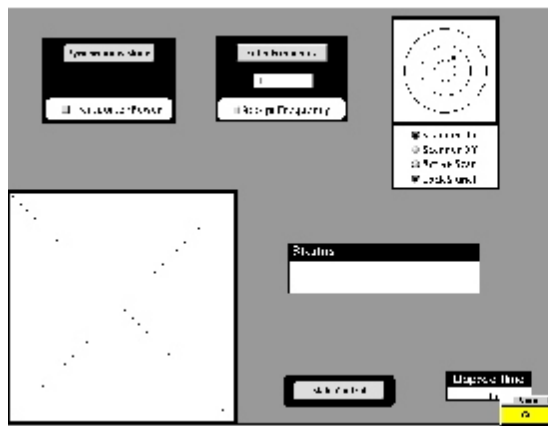


Figure 4. Transporter task structure.



Summary of Manual Transporter System Operating Procedures:

Step 1: Lock on to the homing signal:

- Click "Scanner On"
- Click "Active Scan"
- Wait until: carrier homes in on valid signal
- Click "Lock Signal"
- Click "Scanner Off"

Step 2: Setting the jamming frequency:

- Click "Enter Frequency"
- Type in the desired carrier frequency
- Click "Accept Frequency"

Step 3: Synchronise the transporter and homing signal:

- Click "Transporter Power"
- Click "Synchronise Mode"
- Use the mouse to track the homing signal

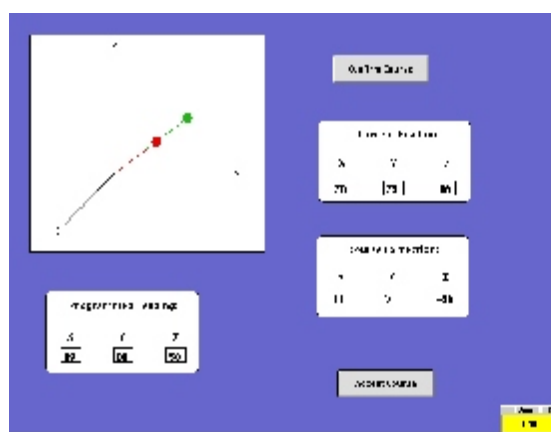
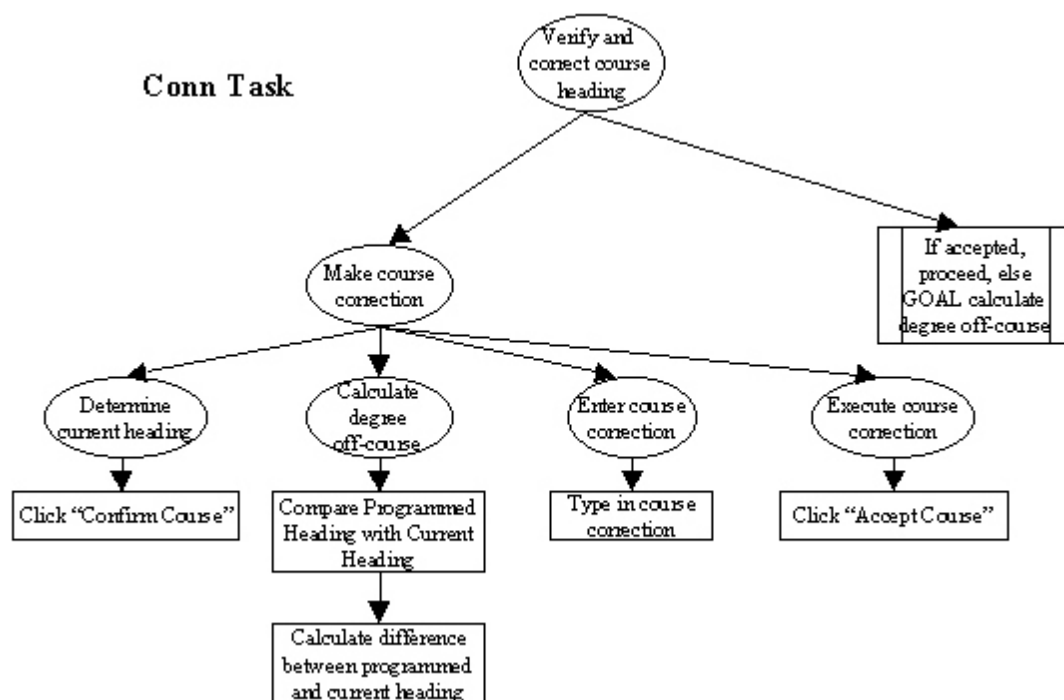
Step 4: Energising the transporter:

- Click the mouse button
- Click "Synchronise Mode"
- Determine if the beam has been successful
- If not, return to Step 1

Then return to Main Control

Figure 5. Transporter task screen and step summary.

The remaining tasks, conn and operations, were introduced primarily as distractors so that participants would be less likely to monitor their performance on the tasks of interest (tactical and transporter). These tasks were structurally much simpler than the tactical and transporter tasks. Figure 6 shows the diagrammatic representation of the conn task, as well as the task display screen and procedure summary. Figure 7 shows the diagrammatic representation of the operations task, as well as the task display screen and procedure summary for this task.



Summary of Navigation Console Operating Procedures:

Step 1. Determine Ship's Course Heading:

- Click "Confirm Course"

Step 2. Compute Course Difference from Programmed Heading:

- Compare course heading in "Programmed Heading" with course heading in "Current Heading"
- Calculate the difference between Programmed Heading and Current Heading by subtracting Current Heading from Programmed Heading ($C\text{ correction} = \text{Programmed} - \text{Current}$)
- Enter course correction in "Course Correction" text boxes

Step 3: Enter course correction into navigation system:

- Click on "Accept Course"

Figure 6. Conn task structure and conn task screen with step summary.

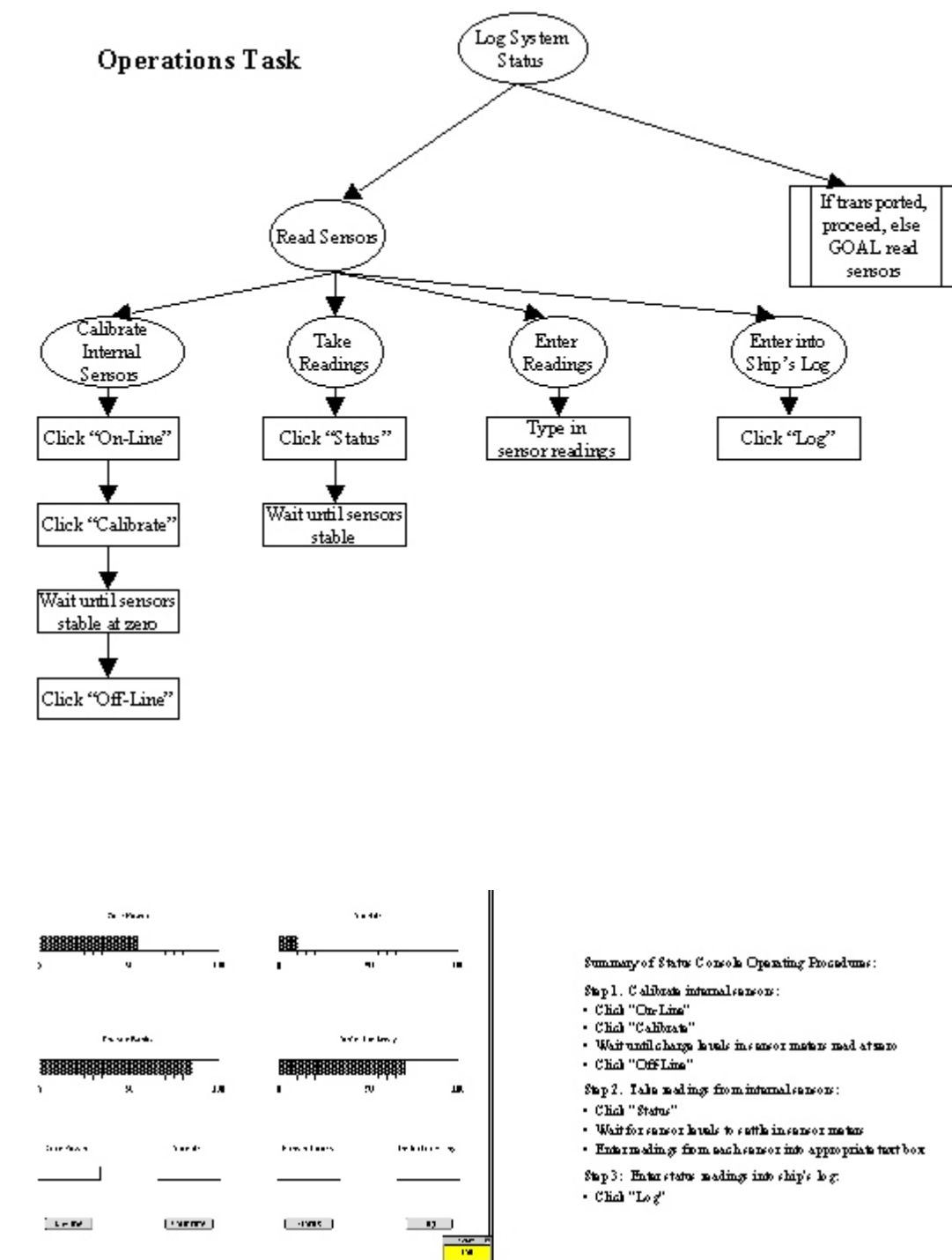


Figure 7. Operations task structure and operations task screen with step summary.

Experiment 1

As stated previously, this experiment served three primary purposes: to replicate the findings of the Byrne and Bovair (1997) experiment regarding the occurrence of postcompletion error, to assess the effects of learning on task performance over time, and to establish protocols and methodology for the second experiment.

Method

Participants

Twenty-two undergraduates from Rice University participated for course credit in a psychology course. There were 12 males and 10 females, with a mean age of 19.8 years, s.d. of 1.2 years.

Materials

The materials for the experiment consisted of a paper instruction manual for each of the tasks, Apple iMac computers running an application written in Macintosh Common Lisp Version 4.3 and another application written in PsyScope, and lightweight stereo headphones. Refer to Appendices B - F for copies of the participant instruction manuals. Appendix B is the Tactical (Postcompletion) manual, Appendix C is the Tactical (Control) manual, Appendix D is the Transporter manual, Appendix E is the Conn manual, and Appendix F is the Operations manual.

Design

This experiment was a two-factor mixed within- and between-subjects design that consisted of the following two factors: task assignment and test period. The first factor, task assignment, had two conditions: tactical postcompletion and tactical control.

Participants were randomly assigned to one of these two conditions. The second factor, test period, was a within-subjects factor representing the days of testing. This measure facilitated the assessment of the learning effects on the bridge station tasks over time. Working memory capacity was assessed as a between-subjects independent continuous individual variable. This measure was used to correlate with other “in-task” measures such as the frequency of errors during task performance, frequency of errors during a letter recall task, and the reaction time for task completion.

The Bridge Officer Qualification software application recorded information on several independent variables. Three different sets of dependent measures were collected: trial summaries, trial protocols, and letter recall summaries. The trial summaries recorded the following information for each experimental trial: the type of bridge station task (i.e., tactical, transporter, conn, or operations), the overall time in seconds for a participant to complete the trial, and the total number of incorrect actions executed during the trial. The trial protocol summaries recorded information on task execution in much more depth. For each trial, the correct, or expected, action was identified, along with the corresponding action (correct or incorrect) executed by the participant at that step in the procedure, and the time elapsed (in milliseconds) since the execution of the last action, accurate to approximately ± 17 milliseconds. The letter recall summaries recorded information on a secondary letter recall task, including the letter string that had been presented, the participant’s response, and the latency (recorded in milliseconds) from the appearance of a response prompt to the when the participant logged a response.

These raw measures were used to determine the primary dependent measure: number of errors. This general “number” of errors dependent measure was broken down into two specific types of error: the number of overall errors during task performance, and the number of postcompletion errors that occurred during task performance on the tactical task. The overall error count for a trial was calculated as the *number of steps* in the trial that had incorrect actions, rather than as the total number of incorrect actions. For example, a participant may have made an error at one particular step in a task; at this one particular step, it was possible for the participant to execute several incorrect actions as he attempted to find the correct response. Even though the participant committed several incorrect actions, they all occurred at the same step, thus, error was counted as the number of steps where an error occurred. Postcompletion errors were assessed directly at the step in question by counting only those occurrences of error where the action executed by the participant was the postcompletion action. For example, in the tactical task, the “Turn Off Tracking” step was the postcompletion step, and the postcompletion action was to incorrectly execute the “Return to Main Control” step instead of “Turn Off Tracking.” Only errors where “Return to Main Control” was the first action executed at the “Turn Off Tracking” step were counted as postcompletion errors.

Completion time was also another dependent measure of interest. This was assessed in two different ways: the time it took a participant overall to complete a trial, and the time it took for a participant to execute a specific step in the task procedure. Finally, the percent of trials correct was used to assess performance on the secondary letter recall task.

Procedure

The experiment was conducted in four sessions, each spaced one week apart. The first session served as the training session for the experimental tasks. In the training procedure, the order of training on the bridge station tasks was randomized for each participant. According to the training protocol, participants first read the instruction manual for a task as identified by the computer. After reading the manual, participants completed one trial of the designated task, during which they were allowed to use the manual. After successfully completing this first trial (the main goal was met and the procedure was executed error-free), participants were instructed to return the manual to the experimenter. This was done to encourage familiarization and memorization of the task procedures on subsequent training trials. Training continued with the designated task until three trials were completed without error. On trials where an error was committed, the trial was terminated, and the participant was informed of the correct action that should have been performed. Once the three error-free trials were logged, participants proceeded to train on the next designated task. The training procedure for all tasks was identical. Participants were allowed to leave the testing center when they completed training for all four experimental tasks. In further discussion, this first session will be referred to as “Training.”

The second, third, and fourth sessions consisted of test trials for the four tasks. These sessions will be referred to as Days 1, 2, and 3 of testing in further discussion. For each test session, participants completed ten trials of each bridge station task (for a total of 40 trials), presented in random order. During these sessions, the experiment program

emitted a warning signal (beep) when an error was made. This allowed participants to know when they had committed an error in the task procedure. The trial, however, continued until the task goal was met. This is consistent with “real-world” tasks; we often receive feedback about our performance, and must continue until we successfully complete a task.

Participants were encouraged to work both accurately and quickly. To this end, the display screen for each task displayed the amount of time in seconds that had passed since the beginning of the trial. Upon successful completion of a trial, participants were again told the time it took to complete the task and the number of errors they had committed.

A letter recall task was introduced on Day 1 of testing. This task was performed concurrently with the test trials on the bridge station tasks and was intended to increase the load on working memory during task performance. The letter recall task took the form of randomly ordered letters spoken at the rate of one letter every four seconds. A tone was presented randomly at intervals ranging from 12 to 45 seconds. Upon hearing the tone, participants recalled the last three letters they heard, in the order they heard them, and typed them into a text box that appeared on the computer screen. All participants wore headphones to avoid “cross-contamination” from other participants on this task.

At the end of testing on Day 3, participants completed a working memory capacity assessment. A reverse digit-span task was administered via a computer program written in PsyScope to measure working memory capacity. Participants performed recall on digit

strings ranging in length from three to ten digits. The strings were presented on the computer screen one digit at a time, with an interval of 500 milliseconds between digits. Participants were prompted to recall and type in the digit string (in reverse order - backwards from the presentation order) by a question mark prompt that appeared on the screen after the last digit in a string. A participant's working memory capacity was interpolated in the interval occurring between the last string length with greater than 50 percent accurate recall and the first string length with less than 50 percent accurate recall.

Results

Training

Analyses were conducted on both the data from the training session as well as the three testing sessions. Data from Training indicate that task complexity influenced task learning: the tactical task required the most training trials, followed by the transporter task, operations task, and conn task. Because each training trial ended with either successful completion or an error, and participants were required to complete three successful trials for each task, the number of errors committed during training corresponded directly to the number of trials required for training (e.g., Number of errors = Number of trials - 3). Thus, the tactical task also reflected the highest number of errors (as each trial ended either with an error or correct completion). Figure 8 depicts the number of trials necessary for training as a function of task station.

As the tactical task was the primary task of interest for task performance (for comparison of the Control and Postcompletion conditions), the mean number of training trials on the tactical task was calculated by condition (control or postcompletion). On the

tactical task, the control group had a mean of 6.54 training trials, s.d. = 1.81, while the postcompletion group had a mean of 10.27 training trials, s.d. = 5.36. A t-test revealed that there was evidence for a reliable difference between the number of training trials required in the Control and Postcompletion conditions: $t(12.24) = -2.18$, $p = 0.049$. The heterogeneity of variance between the conditions (s.d. = 1.81 for Control and s.d. = 5.36 for Postcompletion) was due mainly to individual differences in the participants, rather than a definitive difference between the conditions.

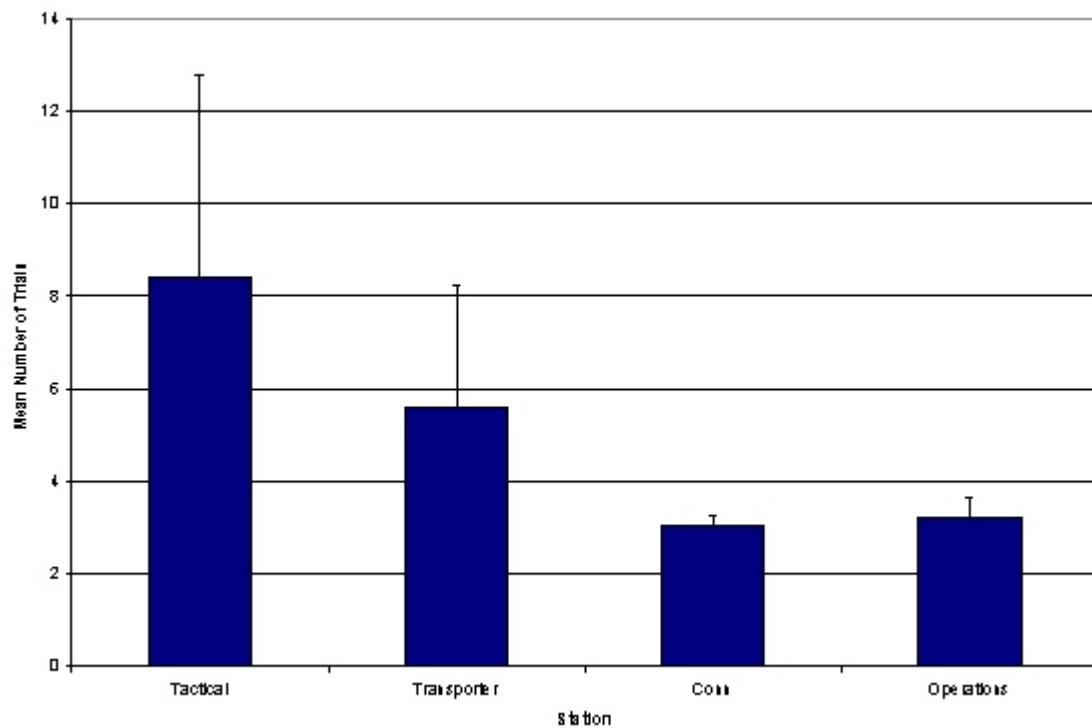


Figure 8. Experiment 1: Mean number of training trials necessary for training as a function of bridge task station.

Testing

The primary phenomenon of interest in the current research was the occurrence of error, in particular, the commission of error at the “Turn Off Tracking” step (the postcompletion step) in the tactical task. When looking at the occurrence of error, there are two different measures that can be used. The first measure looks at the occurrence of error at a particular step in relation to the number of opportunities for that error to occur. This measure of error is referred to as the frequency of error, and is calculated as follows: Error Frequency = Number of Errors at Step X_i / Number of Opportunities for Error at Step X_i . The second measure of error is more specific in scope, as it looks at the occurrence of error at a particular step in relation to the total number of errors that occurred for a task. This measure of error is referred to as the proportion of error occurrence and assesses whether error at a particular step occurs systematically within the task structure. Proportion of error is calculated as follows: Error Proportion = Number of Errors at Step X_i / Total Number of Errors in Task X. Both the frequency and proportion of error measures were calculated as percentage values.

One of the primary goals of the current research was to assess whether the postcompletion error (“Turn Off Tracking” in the tactical task) occurred systematically during task performance, or if error commission reflected a stochastic process. A stochastic error hypothesis would predict that, in the tactical task, any particular step would be responsible for 1/12 (12 steps in the procedure) or 8.3 % of the errors made. Across the three days of testing, the “Turn Off Tracking” error accounted for approximately 14 % of all errors made by participants on the tactical task, $p < 0.001$ (34

out of 246 errors). Analyzed by test day, postcompletion errors accounted for 16.52 % of all tactical errors on Day 1, $p < 0.001$; 16.12 % of all tactical errors on Day 2, $p = 0.015$; and only 6.25 % of tactical errors on Day 3, $p = 0.17$. Thus, the error occurred much more reliably across the three test days, and specifically on Days 1 and 2, than can be accounted for by a stochastic error hypothesis.

Because the participants were tested across three days, another question of interest relates to the change in occurrence of error at the postcompletion step (“Turn Off Tracking”) in the tactical task. The occurrence of error at this step was assessed with both measures of error occurrence introduced above: frequency of error and proportion of error. Figure 9 shows the frequency of error over the test period as a function of condition, and Figure 10 shows the proportion of error at the postcompletion step to total error as a function of condition.

To assess whether a reliable difference existed between the conditions for occurrence of error at the postcompletion step, both the frequency of error and the proportion of error were submitted to a repeated-measures mixed within- and between-subjects Analysis of Variance (ANOVA). The within-subjects variable of interest was test day, and the between-subjects variable was condition (Control version of tactical task or Postcompletion version of tactical task).

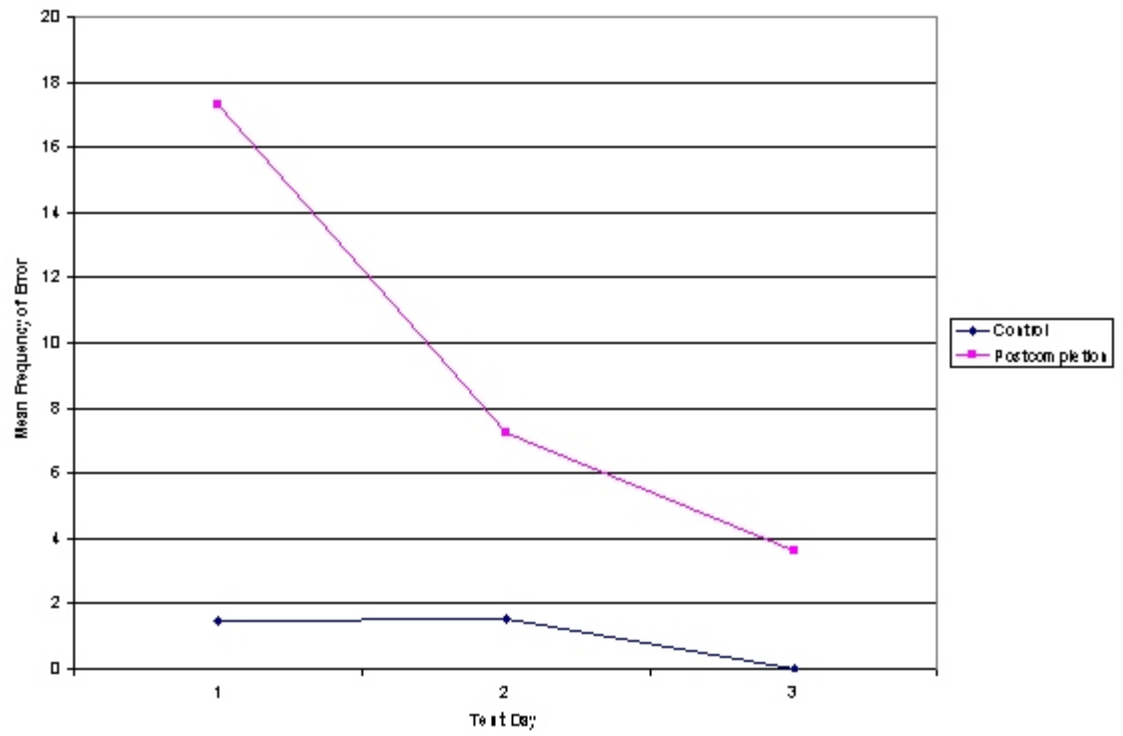


Figure 9. Experiment 1: Frequency of error at the postcompletion step (“Turn Off Tracking”) in the tactical task as a function of experimental condition.

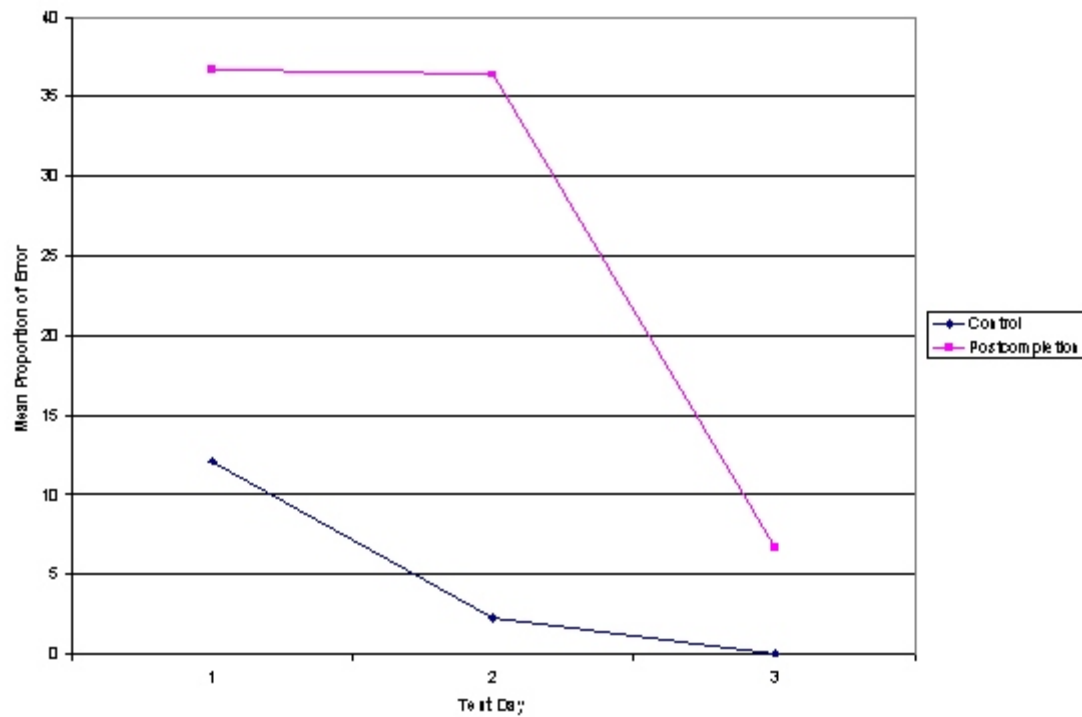


Figure 10. Experiment 1: Proportion of error at the postcompletion step (“Turn Off Tracking”) to total error in the tactical task as a function of condition.

For the frequency of error at the “Turn Off Tracking” step, there was a reliable effect of test day, $F(1.48, 40) = 5.12$, $p = 0.02$, as well as a reliable linear trend, $F(1, 20) = 7.06$, $p = 0.015$. Tests of sphericity for error occurrence as a function of test day resulted in a Huynh-Feldt epsilon of 0.74, indicating a mild violation of sphericity. The corrected Huynh-Feldt values were used for the within-subjects analysis of the main effect of test day on the frequency of error occurrence, as reported above. A reliable between-subjects effect of condition was also found, $F(1, 20) = 14.81$, $p = 0.001$. For the proportion of postcompletion errors per day to the total number of errors for the tactical task, there was also a reliable within-subjects main effect of test day on error occurrence,

$F(2, 40) = 5.991, p = .005$, as well as a reliable linear trend, $F(1, 20) = 13.62, p = 0.001$.

A reliable between-subjects main effect of condition on error proportion was also found, $F(1, 20) = 8.58, p = 0.008$.

As can be seen in Figures 9 and 10, and as supported by the reliable linear trends for both the frequency and proportion of error occurrence, the data exhibit a strong linear trend, as error occurrence decreased over time. These strong linear trends seem to indicate that learning occurred over time, as participants became more practiced with the tasks. Despite a learning trend over the three days, where participants apparently learned to avoid the postcompletion error at the “Turn Off Tracking” step, 27 % of the participants still committed the error on Day 3 of testing. This speaks to the relative robustness of the postcompletion error, as the participants were not successful at eliminating it, even after three days of testing.

The occurrence of error during task performance (as a measure of accuracy on task performance) is one variable of interest during task performance on “timed” tasks; the other variable is the time it takes to complete the task in question. These two variables form the basis of a trade-off known as the speed-accuracy trade-off. During speeded task performance (as in the bridge station tasks, where participants were trying to improve their task completion time from trial to trial), people tend to make errors as they try to respond more quickly to the task (Wickens, 1992). Thus, a question of interest is whether, when participants improved their task performance in terms of error occurrence, did they move to a different location on this speed-accuracy trade-off? In other words, did their improved performance come at the cost of a slower completion time? This

could manifest itself in at least two different ways. First, there could be an overall difference between the conditions for mean task completion time on the tactical task. Figure 11 shows the mean task completion time for the tactical task as a function of condition. While the Control group appeared to have a much slower task completion time, particularly on Day 1, this was the result of an outlier, which had no reliable effect on analyses of differences between the conditions. The second way a move on the speed-accuracy trade-off could manifest itself was at the specific step in question. Figure 12 shows the mean step completion time at the “Turn Off Tracking” step in the tactical task as a function of condition, where in this case, it appeared that the step completion time for the Control condition was much faster than the step completion time for the Postcompletion condition.

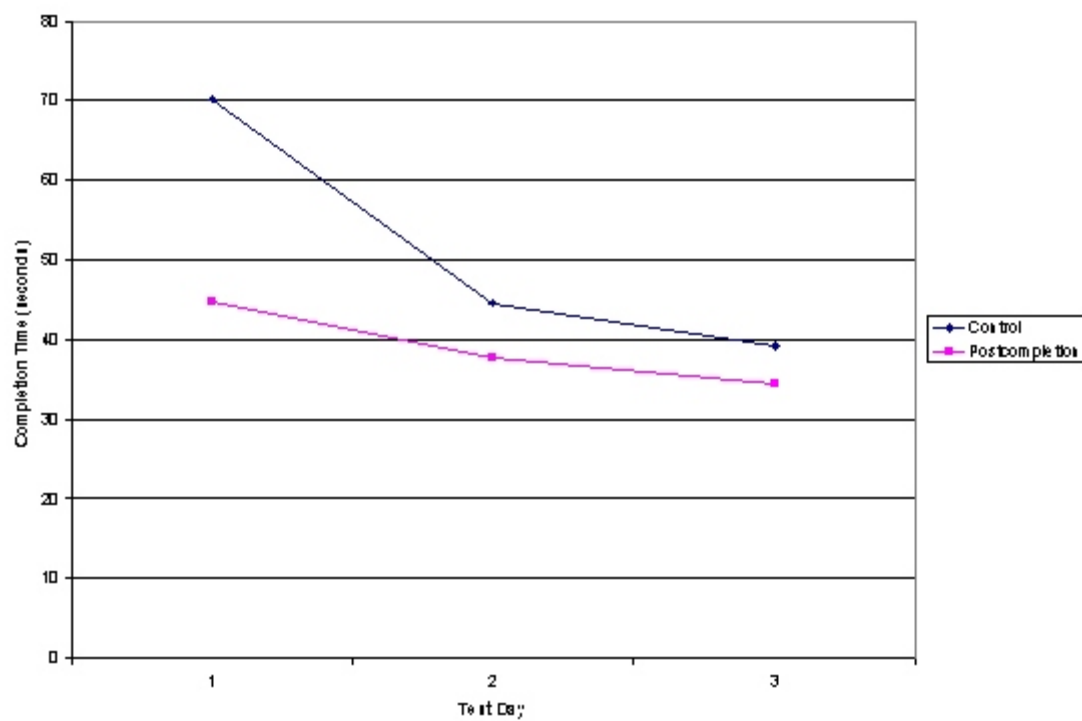


Figure 11. Experiment 1: Mean task completion time on the tactical task as a function of condition.

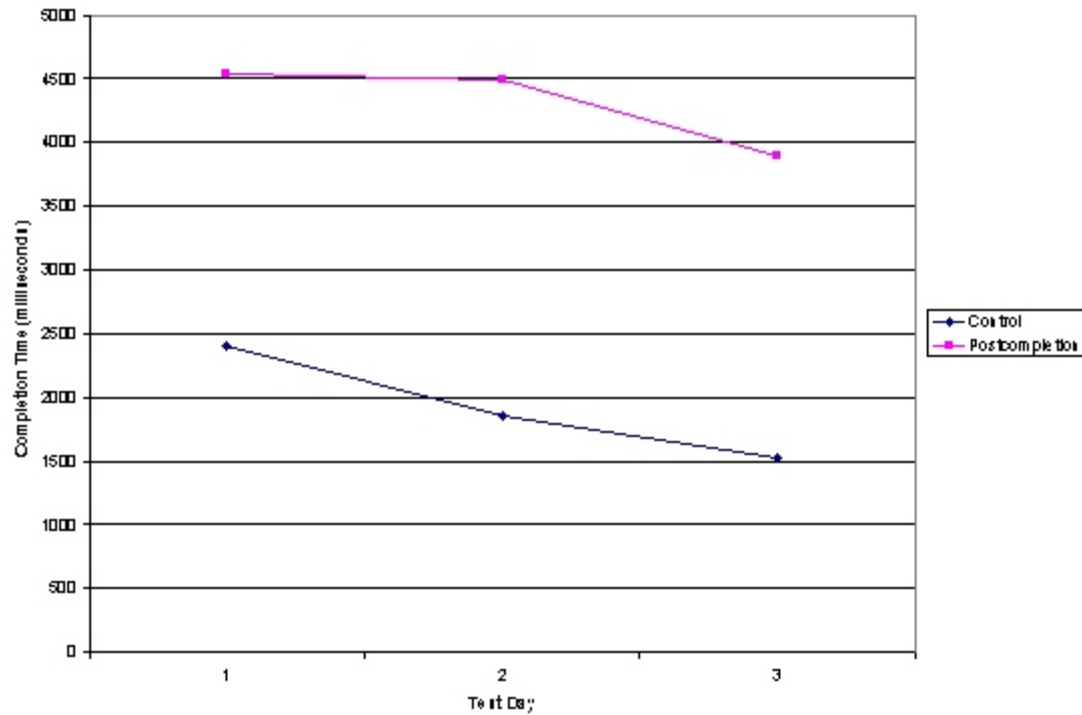


Figure 12. Experiment 1: Mean step completion time at the “Turn Off Tracking” step in the tactical task as a function of condition.

To assess whether there was an effect on overall task completion time, repeated-measures ANOVA was used. For the overall mean task completion time on the tactical task, there was a reliable main effect of day, $F(2, 40) = 4.92$, $p = 0.01$, as well as a reliable linear trend, $F(1, 20) = 5.42$, $p = 0.03$. However, a reliable effect of condition was not found, $F(1, 20) = 2.58$, $p = 0.12$, indicating that there was not a reliable difference between the Control and Postcompletion conditions for the overall mean task completion time for the tactical task.

When step completion time data were analyzed for the “Turn Off Tracking” step, a reliable main effect was found for the effect of test day, $F(2, 40) = 6.83$, $p = 0.003$, as

well as a reliable linear trend, $F(1, 20) = 11.64, p = 0.003$. This is hardly surprising, when evaluated in conjunction with the evidence for the strong linear decrease in error occurrence over the test period. Across the three test days, not only were the participants improving their performance in terms of error, but the practice effects also reduced the amount of time that it took participants to complete the task and its associated steps. In addition to these effects, a reliable main effect of condition was found for differences in step completion time, $F(1, 20) = 74.25, p < 0.001$. Thus, as shown in Figure 12, the participants in the Control condition were taking reliably less time to execute the “Turn Off Tracking” step than the participants in the Postcompletion condition. This is particularly interesting when compared to the fact that there was not a reliable difference between the conditions for the overall task completion time for the tactical task.

This finding, however, makes sense when taking into account the difference between the two versions of the tactical task. In this task, it was possible to fire at the target and miss, requiring participants to re-initiate the procedure. In the control version of the task, after the participant fired at the target, he *must* click “Tracking” to learn whether or not he had destroyed the target. Thus, for the Control condition, executing the “Turn Off Tracking” step became a forcing function that linked clicking the “Tracking” button with learning the outcome of the firing action. Essentially, these participants did not know anything about the state of the system until they executed the “Turn Off Tracking” step. On the other hand, participants in the Postcompletion condition did not have to do anything “special” to learn the state of the system after they fired at the target. As soon as they fired, these participants were informed of the outcome. Participants only

needed to execute the “Turn Off Tracking” step when the target had been destroyed; they therefore had only one opportunity per trial to make this action. Because this step was not linked to information about task performance or the state of the system, it may have taken participants longer to recall that “Turn Off Tracking” was the appropriate action once the target has been destroyed.

The concurrent letter recall task also offered insight into the differences between the conditions for error rates. The letter recall task was introduced to induce a load on working memory, thereby making it more likely that errors would occur. The learning and subsequent decrease in error rates exhibited by the participants on the bridge station tasks, reflecting a relatively skilled level of performance, indicated that perhaps the letter recall task may not have been “enough” to serve as a high load on working memory at this skilled level. If this were the case, it seemed likely that performance of the letter recall task should improve over time. Figure 13 shows however, that accurate performance on the recall task actually decreased over the three test days. Repeated-measures ANOVA revealed that this decrease in the percent of correct letter recall trials over time was reliable, $F(2, 40) = 5.51, p = 0.008$. A reliable difference, however, was not found between conditions for performance on the letter recall task, $F(1, 20) = 0.298, p = 0.6$.

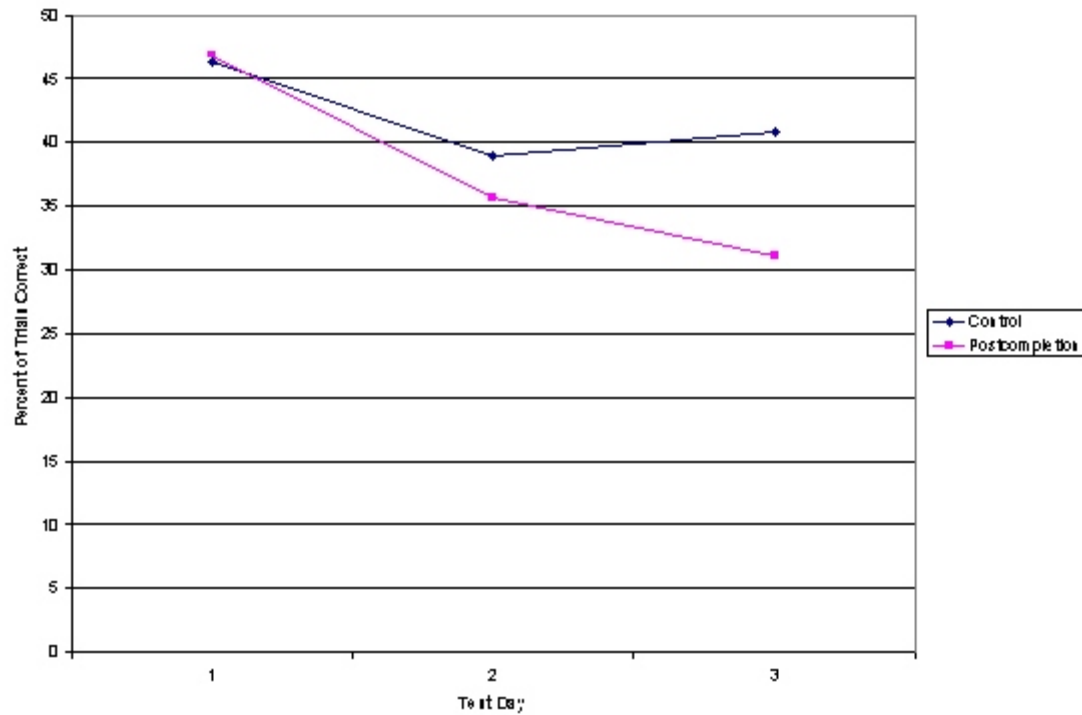


Figure 13. Experiment 1: Mean percent of trials correct on the letter recall task as a function of condition.

From the speed-accuracy trade-off perspective, these results would make sense if the participants also decreased their response time on this task over the three test days; a move along the speed-accuracy curve to a faster response time might result in a corresponding decrease in accuracy. Figure 14 shows that response time for the letter recall task did indeed decrease over the three test days. Repeated-measures ANOVA revealed that this decrease in response time for the letter recall task was reliable, $F(1.476, 40) = 11.48, p < 0.001$. As with the percent of correct trials, there was not a reliable difference between the conditions on response time, $F(1, 20) = 0.983, p = 0.33$.

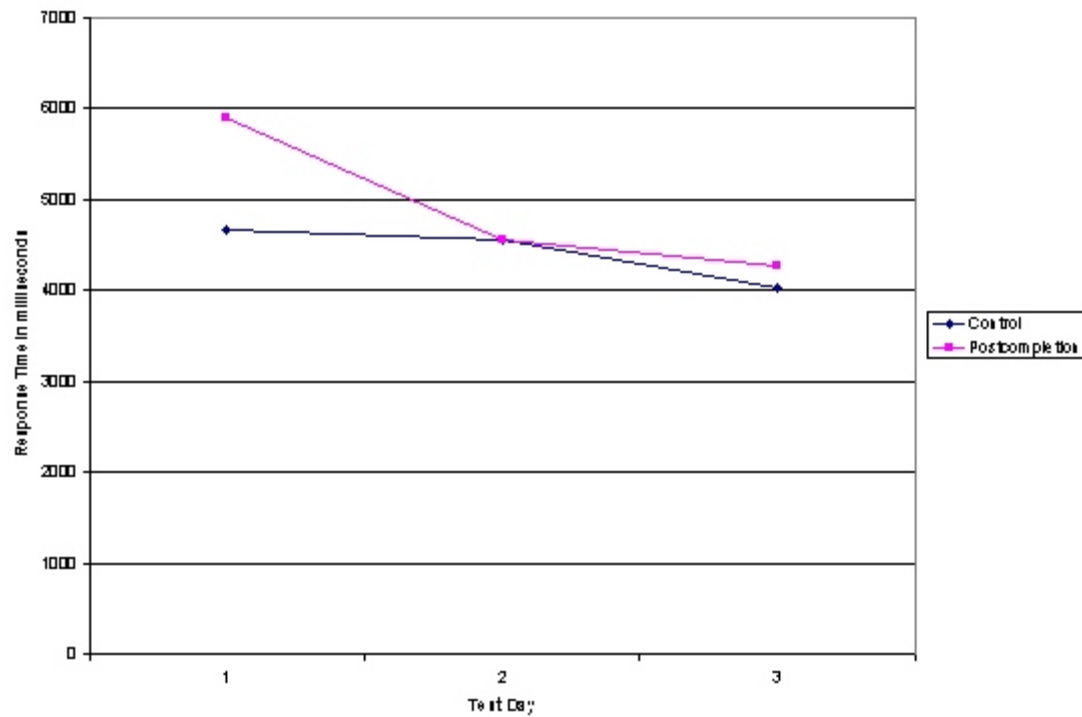


Figure 14. Experiment 1: Mean response latency for the letter recall task as a function of condition.

Another explanation presents itself when considered in conjunction with the data for error occurrence and step completion time at the “Turn Off Tracking” step in the tactical task. As noted earlier, error occurrence at the “Turn Off Tracking” step decreased over the test period, but at the cost of a longer step completion time for the participants in the Postcompletion version. It could be argued that another cost of the decreased error occurrence for the “Turn Off Tracking” step was the apparent decrease in the percent of correct letter recall trials over the test period for the Postcompletion participants; as attention was devoted to the primary task of accurately completing the bridge station tasks, performance on the secondary letter recall task was affected.

Finally, several correlations were computed to assess the relations between working memory capacity and task performance. Working memory capacity, as measured by the reverse digit-span task, was not found to correlate highly with any of the task performance measures, including the occurrence of error, $r = -0.12$, $p = 0.6$; the occurrence of the postcompletion error at the “Turn Off Tracking” step, $r = 0.04$, $p = 0.8$; the mean task completion time, $r = -0.4$, $p = .07$; the step completion time at the “Turn Off Tracking” step, $r = -0.05$, $p = 0.8$; or the percent of correct trials on the letter recall task, $r = -0.34$, $p = 0.12$. It is hypothesized that the lack of reliable comparisons is due primarily to the reduced error rates and faster completion times that seem to be linked to learning effects over time.

Discussion

This experiment successfully achieved its goal of replicating the original Byrne and Bovair (1997) experiment by generating the postcompletion error in a laboratory setting. Additionally, the results of this experiment supported the finding that the postcompletion error in the tactical task (“Turn Off Tracking”) is most likely not stochastic in nature; the postcompletion error occurred at the “Turn Off Tracking” step in the tactical task more often than a stochastic error theory would predict. Thus, this supports the idea as presented earlier that postcompletion errors are different in some fashion from other procedural errors.

Furthermore, this experiment demonstrated a reduction in the commission of the postcompletion error over time. As there were no changes to the tasks themselves during the three test sessions, such a decrease in error commission suggests that some form of

learning took place, resulting in spontaneous error reduction. Thus, it appears that participants learned on their own to recover from the postcompletion step. The robustness of the error is demonstrated in the fact, however, that the error was not eliminated completely by the last day of testing. The key finding, however, regarding the occurrence of error at the “Turn Off Tracking” step relates to the step completion time. As shown in Figure 12, the participants with the postcompletion version of the tactical task took approximately two seconds longer to execute the “Turn Off Tracking” step than participants with the control version. Thus, while it can be argued that the participants in the postcompletion version reliably decreased the occurrence of error at the postcompletion step, this level of accurate performance came at the cost of speed in step execution. Given a successful replication of the postcompletion error in a laboratory setting, the following experiment will examine the effect of various disciplinary measures on the commission of the postcompletion error over time.

Experiment 2

This experiment examined the influence of “typical” organizational responses to error on the commission of a very particular type of error - postcompletion error. The first experiment established that the postcompletion error can be produced and replicated in an experimental setting, and that through practice, it was possible to reduce, though not entirely eliminate, the postcompletion error. Given this foundation, the following experiment sought to manipulate error commission as a function of response to task performance through such organizational responses as reprimands, re-instruction, and praise, and through the systems response of implementing a new procedure. These

methods were selected for implementation as they appear to be well-represented in the discussion of organizational discipline presented earlier. According to the model developed by Arvey and Jones (1985), organizations administer discipline to control employee behavior; thus, reprimands would be issued with the intent of reducing the occurrence of a particular error, while praising individuals for particular behaviors is intended to encourage the continued occurrence of a particular behavior. Re-instructing individuals on particular procedures is another method by which organizations attempt to reduce the occurrence of certain undesirable behaviors. These methods are consistent with the description of the culture of medicine, as explained by Leape (1994), where it is believed that error-prone performance can be “trained out” of doctors and nurses, and where errors are met with punishment: these methods reflect a “conventional wisdom” approach to the occurrence of error. Another method of influencing subsequent task performance is to redesign a problematical task in such a way that the underlying causes of poor performance are addressed and rectified; this method reflects a systems approach to the occurrence of error. It is hypothesized that these “typical” organizational responses (reprimands, re-instruction, and praise), which essentially amount to exhortations to “pay attention,” will have little reliable, long-lasting effect on the commission of errors, while redesigning the procedure will reliably impact the number of postcompletion errors made during task performance.

Method

Participants

Fifty-four undergraduates from Rice University participated for course credit in a psychology course; six psychology graduate students from Rice University received \$40 for their participation, for an overall total of 60 participants. There were 41 females and 19 males, with a mean age of 19.6 years, s.d. of 1.6 years.

Materials

The materials for the experiment were the same as for the pilot experiment: a paper instruction manual for each of the experimental tasks, Apple iMac computers running an application written in Macintosh Common Lisp Version 4.3, and lightweight stereo headphones. In addition, an exit questionnaire was developed for each experimental condition; the exit questionnaire is provided in Appendix G.

Design

This experiment was a two-factor mixed within- and between-subjects design consisting of the following two factors: response to error and test period. The first factor, response to error, had six conditions: Control (the control version of the tactical task), Postcompletion (the postcompletion version of the tactical task), Reprimand, Re-instruction, Praise, and Redesign. The Reprimand, Re-instruction, Praise, and Redesign conditions used the postcompletion version of the tactical task. Participants were randomly assigned to one of these six conditions. The second factor, test period, was a within-subjects factor that accounted for the several days of testing and facilitated the assessment of error commission and learning effects over time. Working memory

capacity was a between-subjects independent continuous individual variable used to correlate with the within-subjects dependent measures such as frequency of error during task performance, frequency of error during a concurrent letter recall task, and reaction time for task completion.

The Bridge Officer Qualification software application recorded information on several independent variables. Three different sets of dependent measures were collected: trial summaries, trial protocols, and letter recall summaries. The trial summaries recorded the following information for each experimental trial: the type of bridge station task (i.e., tactical, transporter, conn, or operations), the overall time in seconds for a participant to complete the trial, and the total number of incorrect actions executed during the trial. The trial protocol summaries recorded information on task execution in much more depth. For each trial, the correct, or expected, action was identified, along with the corresponding action (correct or incorrect) executed by the participant at that step in the procedure, and the time elapsed (in milliseconds) since the execution of the last action, accurate to approximately ± 17 milliseconds. The letter recall summaries recorded information on a secondary letter recall task, including the letter string that had been presented, the participant's response, and the latency (recorded in milliseconds) from the appearance of a response prompt to the when the participant logged a response.

These raw measures were used to determine the primary dependent measure: number of errors. This general "number" of errors dependent measure was broken down into two specific types of error: the number of overall errors during task performance, and the number of postcompletion errors that occurred during task performance on the tactical

task. The overall error count for a trial was calculated as the *number of steps* in the trial that had incorrect actions, rather than as the total number of incorrect actions. For example, a participant may have made an error at one particular step in a task; at this one particular step, it was possible for the participant to execute several incorrect actions as he attempted to find the correct response. Even though the participant committed several incorrect actions, they all occurred at the same step, thus, error was counted as the number of steps where an error occurred. Postcompletion errors were assessed directly at the step in question by counting only those occurrences of error where the action executed by the participant was the postcompletion action. For example, in the tactical task, the “Turn Off Tracking” step was the postcompletion step, and the postcompletion action was to incorrectly execute the “Return to Main Control” step instead of “Turn Off Tracking.” Only errors where “Return to Main Control” was the first action executed at the “Turn Off Tracking” step were counted as postcompletion errors.

Completion time was another dependent measure of interest. This was assessed in two different ways: the time it took a participant overall to complete a trial, and the time it took for a participant to execute a specific step in the task procedure. Finally, the percent of trials correct was used to assess performance on the secondary letter recall task.

Procedure

The participants were run in four sessions, spaced one week apart. The first session served as the training session for the experimental tasks. In the training procedure, the order of training for the four bridge station tasks was randomized for each participant. According to the procedure, participants first read the instruction manual for

a bridge station task as identified by the computer. After reading the manual, they completed one trial of the designated task, during which they were allowed to use the manual. After successfully completing this first trial, participants were instructed to return the manual to the experimenter. This was done to encourage familiarization and memorization of the task procedures on subsequent training trials. Training continued with the designated task until three trials were completed without error. During these training trials, if the participant made an error while executing the procedure, the trial ended and the participant was informed of the correct action that should have been performed. Once three error-free trials were logged, participants proceeded with training on the next designated task. The training procedure for all tasks was identical. Participants were allowed to leave the testing center when they completed training for all four experimental tasks. In further discussion, this session will be referred to as “Training.”

The second, third, and fourth sessions consisted of test trials for the four bridge station tasks. Due to the introduction of the organizational response interventions during the third session, these sessions will be referred to as Days 1, 2a (before the intervention), 2b (after the intervention), and 3 of testing in further discussion. Participants completed 13 trials of each bridge station task, presented in random order, for a total of 52 trials per test day. During these sessions, the experiment program emitted a warning signal (beep) when an error was made so that participants were aware of committing the error. The trial, however, continued until the task goal was met.

Participants were encouraged to work both accurately and quickly. To this end, the display screen for each task displayed the amount of time in seconds that had elapsed since the beginning of the trial, as well as a “game score.” This scoring system was cumulative across all trials and reflected the following scoring system. For every correctly executed step in a task, the score was incremented 25 points; for every incorrectly executed step in a task, the score was decremented 50 points. Finally, for every incorrect working memory recall trial, the score was decremented 75 points; no points were accumulated for successfully completing a recall trial. Upon successful completion of a trial, participants were informed of the time it took to complete the task, the number of errors they committed, and their score.

Interventions were introduced during Day 2 of testing to determine what effect they had on subsequent performance according to the condition to which a participant belonged. Participants in the Control and Postcompletion conditions did not receive an intervention at this point; these conditions served essentially as baseline conditions. Participants in the Reprimand, Re-instruction, Praise, and Redesign conditions received an intervention after completing at least six trials of both the tactical and transporter tasks. Participants in the Reprimand condition received a report of poor performance and were advised to improve their performance, participants in the Re-instruction condition were given a report of poor performance and asked to re-read the manuals for the tactical and transporter tasks before being allowed to continue with testing, and participants in the Praise condition were given a report of good performance and encouraged to “keep up the good work.” In the Redesign condition, participants switched from the postcompletion

version of the tactical task to the non-postcompletion version. Table 2 contains the messages used as the interventions to moderate subsequent behavior.

Table 2

Experiment 2: Intervention text for the Reprimand, Re-instruction, Praise, and Redesign conditions

Intervention	Message
Reprimand	<p>!!!WARNING!!!</p> <p>Cadet, your cumulative performance on the Qualification tasks is falling below the 25th percentile of all cadets undergoing Qualification! Try concentrating harder on the tasks at hand! If your performance does not improve, you will be asked to leave the Qualification center and you will not be considered for a Command Line Commission.</p>
Re-instruction	<p>!!!WARNING!!!</p> <p>!!!WARNING!!!</p> <p>Cadet, your cumulative performance on the Qualification tasks is falling below the 25th percentile of all cadets undergoing Qualifications! To help improve your performance, inform the officer present that you must undergo re-training. If your performance does not improve after re-training, you will be asked to leave the Qualification center and you will not be considered for a Command Line Commission.</p>
Praise	<p>!!!WARNING!!!</p> <p>Congratulations Cadet! Your cumulative performance on the Qualification tasks is well above the 90th percentile of all cadets undergoing Qualifications! Continue your excellent work, and you will be sure to receive a Command Line Commission.</p>
Redesign	<p>!!!WARNING!!!</p> <p>Cadet, your cumulative performance on the Qualification tasks is falling below the 25th percentile of all cadets undergoing Qualification! In particular, you appear to have difficulty at Tactical. A new model of the MB-X15 standard phaser control bank has been introduced into newer classes of Starfleet vessels. Please inform the officer present that you have been advised to continue testing on the new model of the MB-X15 standard phaser control bank. As the new model is similar to the model on which you have been testing, you will only be given the manual for the new model to use on your next trial at Tactical. After you complete this trial, the officer present will collect your manual.</p> <p>!!!WARNING!!!</p>

Participants completed the working memory capacity assessment during the fourth session. A reverse digit-span task measured working memory capacity. Participants performed digit recall on digit strings ranging in length from three to ten digits. The

strings were presented on the computer as a program written in PsyScope. A participant's working memory span was interpolated in the interval occurring between the last string length with greater than 50 percent accurate recall and the first string length with less than 50 percent accurate recall.

In addition to the bridge station tasks, participants performed a concurrent letter recall task that was introduced on Day 1 of testing. This task was intended to increase the load on working memory during task performance. Participants were presented with auditory stimuli that took the form of randomly ordered letters spoken at the rate of one letter every three seconds. A tone was presented randomly at intervals ranging from 9 - 45 seconds. Upon hearing this tone, participants were directed to recall the last three letters they heard, in the order they heard them, and type them into a text box that appeared on the computer screen. All participants wore headphones to avoid "cross-contamination" from other participants on this task.

Results

Analyses were conducted on both the data from Training as well as Days 1, 2, and 3 of testing. There were two primary results of interest: task performance at the "Turn Off Tracking" step in the tactical task, and the potential difference in task performance after the introduction of the various interventions during Day 2 of testing. The Control and Postcompletion conditions were baseline conditions against which differences in task performance could be compared. In this experiment, an intervention was judged to have made a reliable difference in task performance if the change in task performance after an

intervention was reliably different from the corresponding baseline condition (Control or Postcompletion).

Training

As in Experiment 1, task complexity appeared to influence task learning. As shown in Figure 15, the tactical task required the most training trials, followed in descending order by the transporter, operations, and conn bridge station tasks. Each training trial ended in either successful completion or an error. As participants were required to complete three successful trials for each task, the number of errors committed during training corresponded directly to the number of training trials required (e.g., Number of Errors = Number of Training Trials - 3). Thus, the tactical task also reflected the greatest occurrence of error during training.

As a check across the conditions, Analyses of Variance (ANOVA) were conducted to determine whether there were any initial differences between the various conditions on the overall mean number of training trials required; it was expected that there should be none. As expected, there were no reliable effects, $F(5, 54) = 0.55$, $p = 0.73$. Further analysis revealed that there also were no reliable differences between the conditions at each particular bridge task station for the number of training trials required. Table 3 contains the results of these ANOVAs. These results indicate that there were no initial differences between the groups on the number of trials necessary for training.

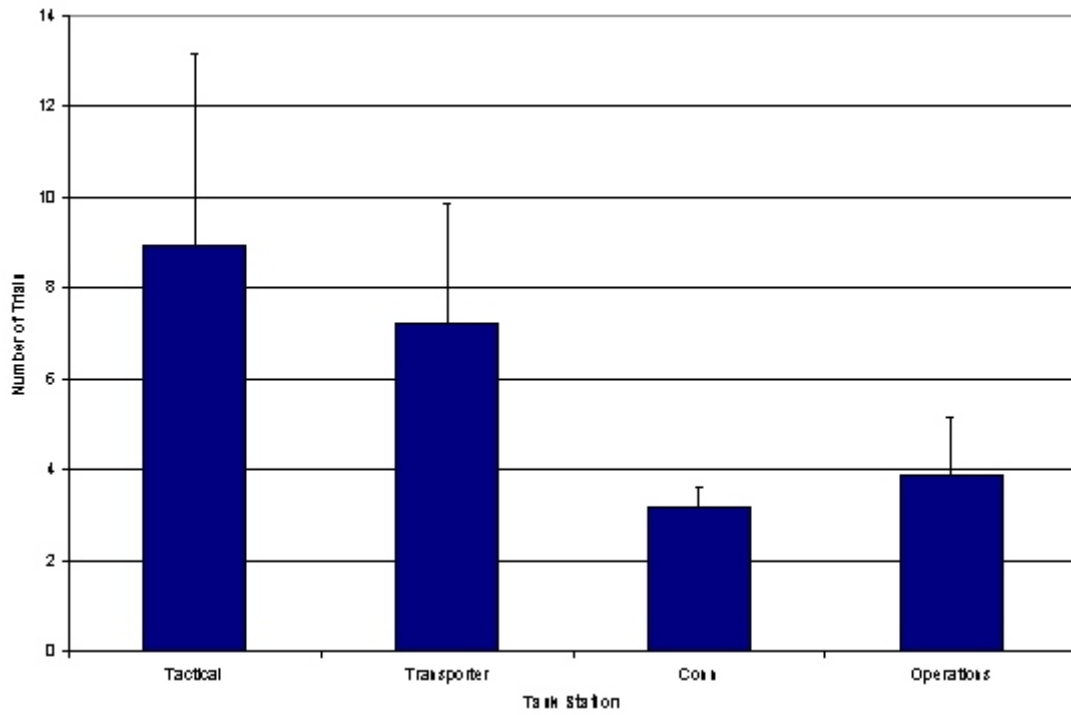


Figure 15. Experiment 2: Mean number of training trials necessary as a function of bridge task station.

Table 3

Experiment 2: Analysis of Variance for number of training trials by bridge task station as a function of condition

Task	df	F	p
Tactical	5, 54	0.55	0.73
Transporter	5, 54	1.05	0.39
Conn	5, 54	0.14	0.98
Operations	5, 54	1.154	0.34

Of primary interest in this experiment was the occurrence of error during the tactical task at the “Turn Off Tracking” step (the postcompletion step). While

participants in the Postcompletion conditions (Reprimand, Re-instruction, Praise, and Redesign) were instructed by the instruction manual for the correct action at this point in the procedure, 18 % of the participants (9 out of 50) made the postcompletion error on the initial training trial. Furthermore, over the course of training, 94 % of the Postcompletion conditions participants (47 out of 50) made the postcompletion error at some point.

Testing

The primary phenomenon of interest in the current research was task performance, and in particular the occurrence of error during task performance. One of the primary concerns therefore, was whether error occurred systematically at particular steps in a task procedure, or if it reflected some stochastic process. For the tactical and transporter tasks, which had the same task structure, a stochastic error theory would argue that any particular step in the task procedure would be responsible for $1/12$ (12 steps in the procedure), or 8.3 % of the errors made during task performance. As can be seen in Figures 16 and 17, however, certain steps in the tactical and transporter tasks reflected an error rate much greater than that proposed by the stochastic error theory.

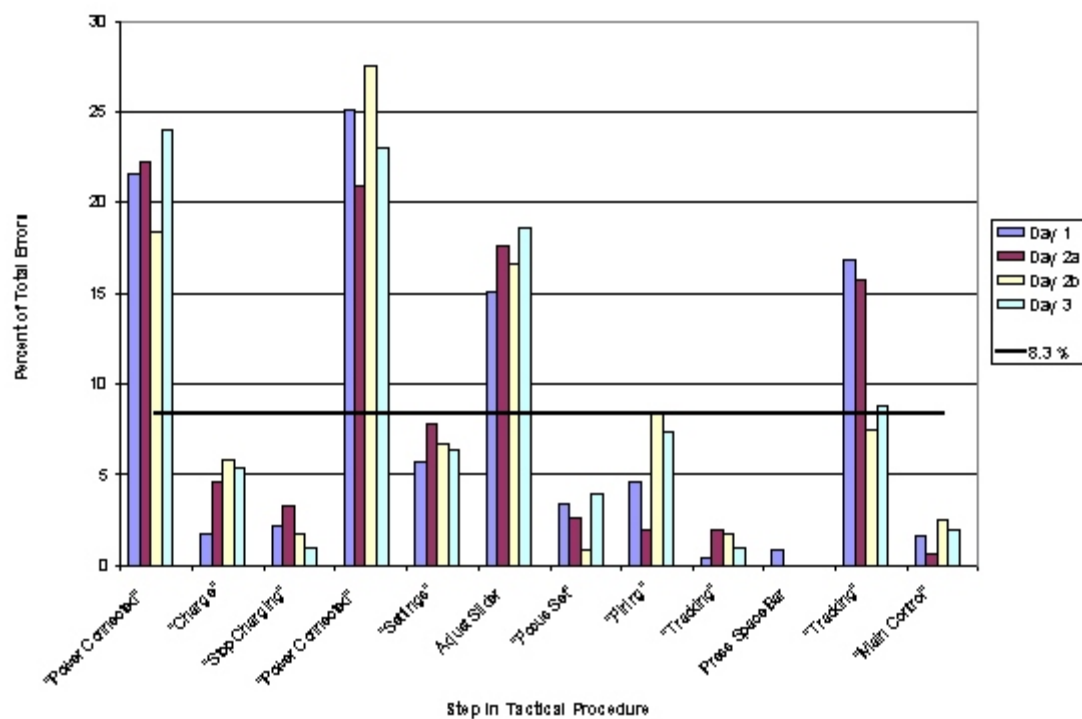


Figure 16. Experiment 2: Percent of total errors made as a function of the step in the tactical task procedure.

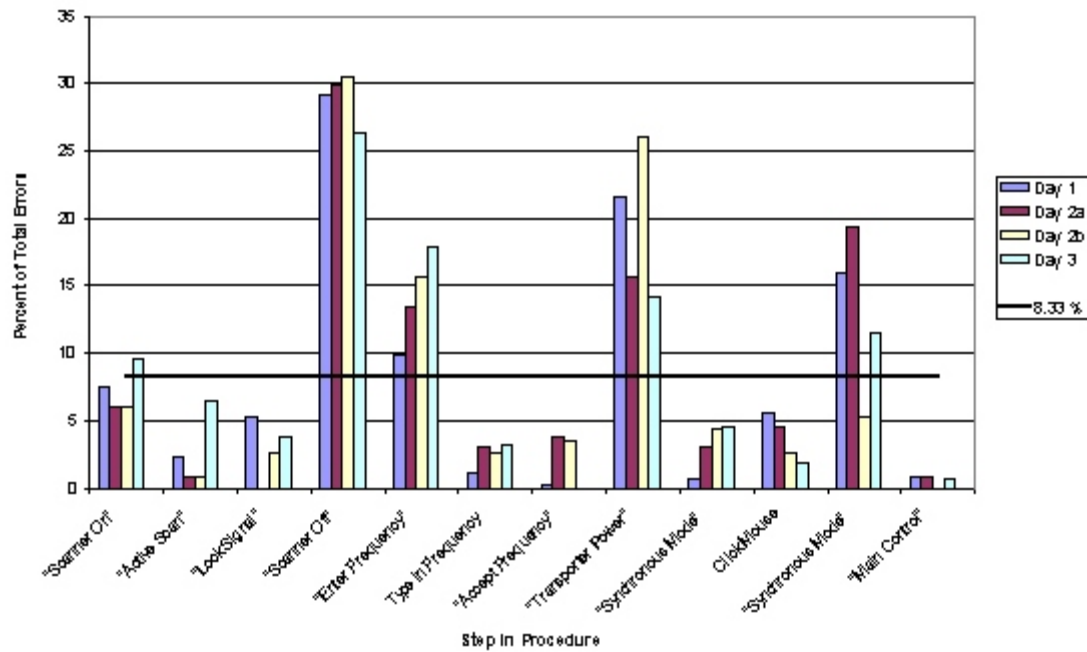
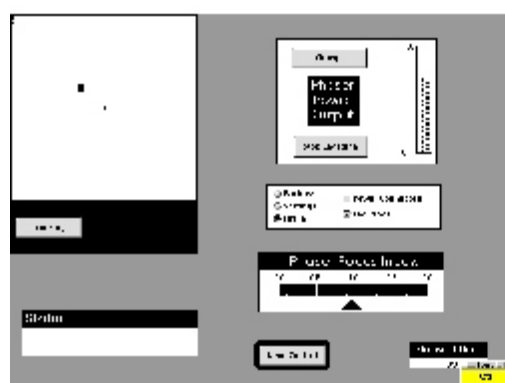


Figure 17. Experiment 2: Percent of total errors made as a function of the step in the transporter task procedure.

Examination of Figure 16 revealed that the postcompletion step, “Turn Off Tracking” in the tactical task, cannot be considered to occur stochastically. Particularly for Days 1 and 2a of testing (before the intervention), this step appeared to account for much more than 8.3 % of the errors made. Analyzed by test day, postcompletion errors at the “Turn Off Tracking” step accounted for 17 % of tactical errors on Day 1, 15.78 % of the tactical errors on Day 2a, 7.69 % of the tactical errors on Day 2b, and 8.69 % of the tactical errors on Day 3. The difference from the stochastic prediction is statistically reliable for Day 1 and Day 2a; $p < 0.001$ for Day 1 and $p < 0.001$ for Day 2a.

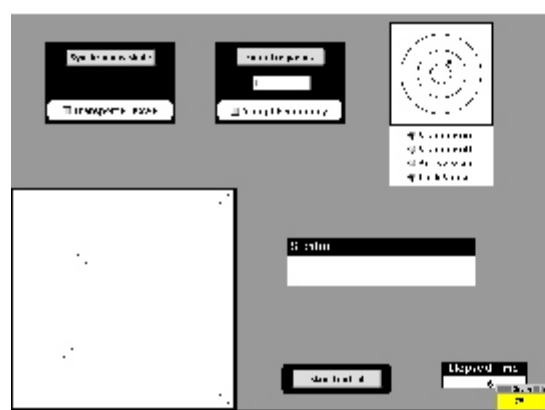
As can be seen in Figures 16 and 17, other steps in the tactical and transporter tasks exceeded the 8.3 % criterion for stochastic error. Recall that the tactical and transporter tasks share the same underlying task structure. Thus, the occurrence of error can be examined between the two tasks at various steps in the procedure. Figures 18a and 18b show the task screens and step summaries for the tactical and transporter tasks, respectively.



Summary of Phase Bank Operating Procedures (Postcompletion Version):

- Step 1: Change the phase:**
- Click "Power Cancelled"
 - Click "Change"
 - Wait until phase changes the appropriate amount
 - Click "Stop Change"
 - Click "Power Cancelled"
- Step 2: Set phase beam focus:**
- Click "Sawage"
 - Adjust location of slider to desired focus
 - Click "Focus Set"
- Step 3: Track the target:**
- Click "Firing"
 - Click "Tracking"
 - Use arrow keys to adjust location of the target indicator
- Step 4: Fire the phase:**
- Press the space bar
 - Determine if the target has been destroyed
 - If so, click "Tracking"
 - If not, return to Step 1
- Then return to Main Control

Figure 18a. Tactical task screen and step summary for the postcompletion version.



Summary of Manual Transporter System Operating Procedures:

- Step 1: Lock on to the homing signal:**
- Click "Scanner On"
 - Click "Active Scan"
 - Wait until scanner homes in on valid signal
 - Click "Lock Signal"
 - Click "Scanner Off"
- Step 2: Setting the jamming frequency:**
- Click "Enter Frequency"
 - Type in the desired scanner frequency
 - Click "Accept Frequency"
- Step 3: Synchronise the transporter and homing signal:**
- Click "Transporter Power"
 - Click "Synchronise Mode"
 - Use the mouse to track the homing signal
- Step 4: Energising the transporter:**
- Click the mouse button
 - Click "Synchronise Mode"
 - Determine if the beam has been successful
 - If not return to Step 1
- Then return to Main Control

Figure 18b. Transporter task screen and step summary.

Of particular interest were the first occurrence of the “Power Connected” step (“Power Connected 1”) and the second occurrence of the “Power Connected” step (“Power Connected 2”) in the tactical task, and the “Scanner On” and “Scanner Off” steps in the transporter task. Error at “Power Connected 1” in the tactical task and “Scanner On” in the tactical task comprise a special type of procedural error that will be referred to an “initiation” error - these steps must be completed before the primary goal each is associated with may be initiated. In the tactical task, participants must click “Power Connected” before they can begin charging the battery. In this case, connecting the power was skipped in favor of initiating the main goal of charging the battery. Correspondingly, in the transporter task, participants skipped turning on the scanner in favor of initiating the active scan for the target.

Similarly, “Power Connected 2” in the tactical task and “Scanner Off” in the transporter task both comprise what can be thought of as an “internal” postcompletion step in the tactical and transporter tasks. In the tactical task, participants must click on “Power Connected” after the battery has completed charging. However, this step occurs after the primary goal of “charge battery” is complete, leading participants to the first step (click “Settings”) of the next goal (“set focus”). In the transporter task, participants must click “Scanner Off” after locking on the target’s signal. The postcompletion step here leads participants to skip turning off the scanner after locking on the target (the primary goal). Participants jump from locking on the target to the first step (click “Enter Frequency”) of the next goal - setting the jamming frequency. Task performance

associated with each of the above steps (“Turn Off Tracking,” “Power Connected 1” / “Scanner On,” and “Power Connected 2”/ “Scanner Off”) will be discussed separately.

Overall task performance. Two measures were used to assess the occurrence of error during task performance. The first measure assessed the frequency of error at a particular step in a task given x opportunities for the error (e.g., Frequency = Number of Errors at Step X_i / Number of Opportunities for Error at Step X_i). This was calculated as a percentage value. The second measure assessed the proportion of error to the total error for a task (e.g., Proportion = Number of Errors at a Step X_i / Total Number of Errors in Task X). This was also calculated as a percentage value. This proportion measure was used in assessing the systematic occurrence of error against a stochastic error theory as was illustrated earlier in Figures 16 and 17.

Because the participants were tested across three days, one question of interest relates to the overall occurrence of error across these three days. Due to the introduction of the intervention on Day 2 of testing, this day was split into two sections, so that there were a total of four test periods. To assess the overall occurrence of error, the frequency of overall error was submitted to a repeated measures mixed within- and between-subjects ANOVA. Figure 19 displays the frequency of overall error across the four test periods. The within-subjects variable of interest was test period and the between-participants variable was the response to error condition. For the frequency of error, there was a reliable effect of test period on error occurrence, $F(3, 162) = 53.21, p < 0.001$, as well as a reliable linear trend, $F(1, 54) = 103.81, p < 0.001$. There was no reliable effect of condition on error occurrence, $F(5, 54) = .5, p = 0.77$. As evidenced by the reliable

linear component for the frequency of error, and the chart in Figure 19, there appears to be a strong learning effect for the tasks over the testing period.

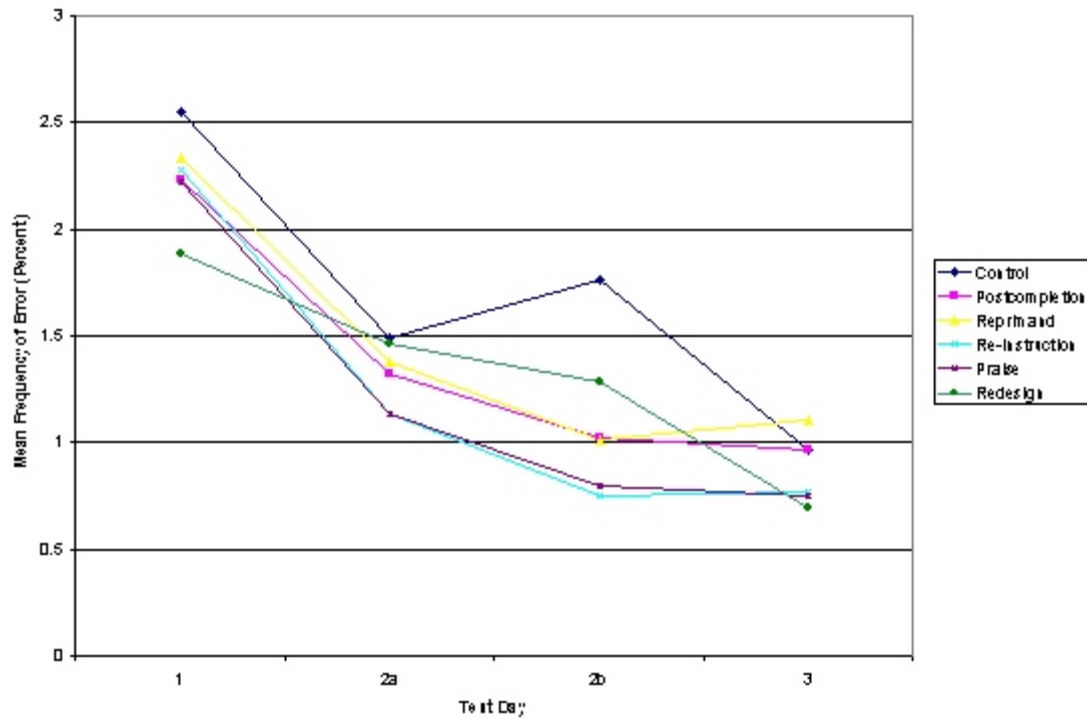


Figure 19. Experiment 2: Frequency of overall error across the four test periods as function of condition.

As described in Experiment 1, the speed-accuracy trade-off describes a phenomenon whereby it is assumed that increased accuracy on a “speeded” task comes at the expense of speed, and vice versa. Thus, another measure that might reveal differences over time between the conditions for the effect of the experimental interventions is task completion time. Given the linear trend indicating a decrease in error occurrence over time, it was assumed that the same learning effects would hold true for task completion time as well. That is, participants would get faster as they learned how to perform the

bridge station tasks. A repeated-measures analysis of variance supported this assumption, where a main effect of test period was found for overall task completion time: $F(3, 162) = 70.28, p < 0.001$. This effect is illustrated in Figure 20.

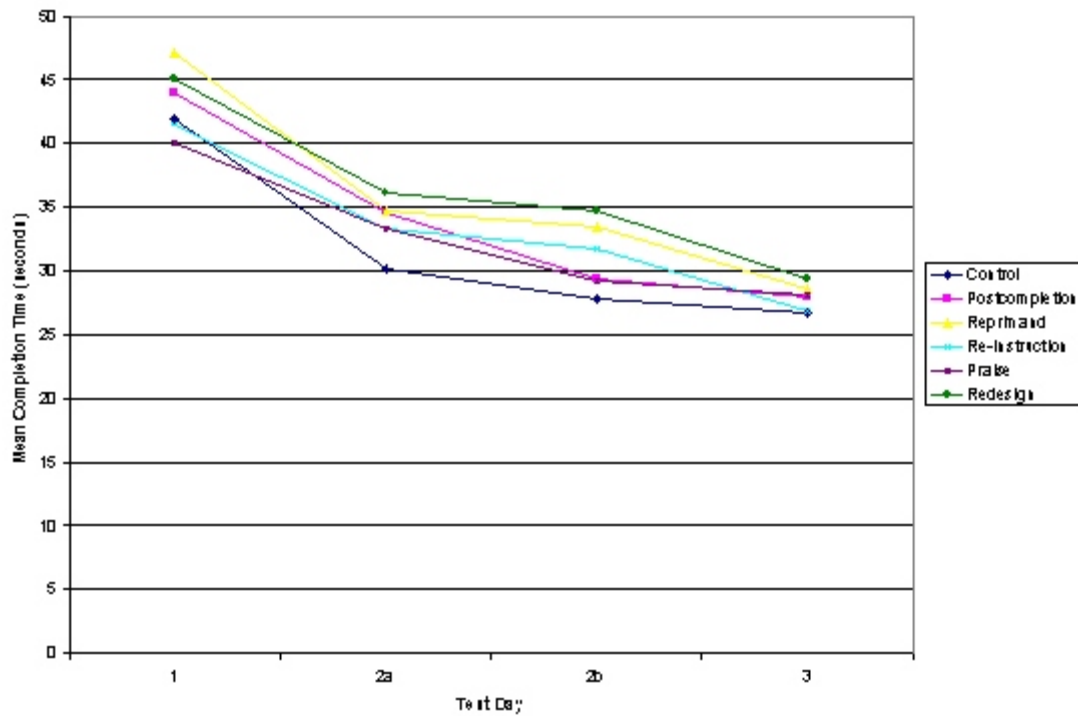


Figure 20. Experiment 2: Mean task completion time across all bridge station tasks as a function of condition.

As the tactical task was of primary interest, the overall task completion time for this task was assessed for differences between conditions. Figure 21 shows the mean task completion time for the tactical task as a function of condition. Repeated-measures ANOVA revealed a reliable main effect of test period, $F(1.72, 162) = 44.49, p < 0.001$, but not a reliable main effect of condition, $F(5, 54) = 0.33, p = 0.9$. Thus, participants became faster at completing the tactical task over the testing period, but there were not

reliable differences between the conditions on this measure. This overall assessment of task completion time will be used for comparison purposes against differences in step completion time at particular steps in the tactical task.

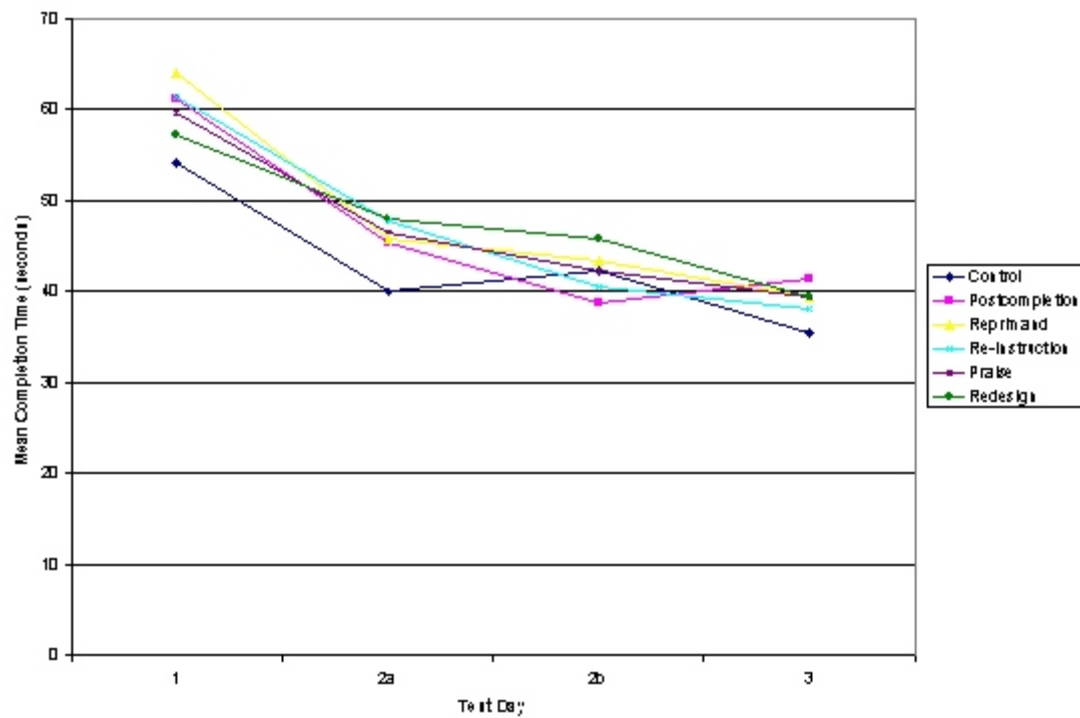


Figure 21. Experiment 2: Mean task completion time for the tactical task as a function of condition.

Performance at “Turn Off Tracking” in the tactical task. The task performance of specific interest, however, was performance at the “Turn Off Tracking” step in the tactical task: the postcompletion step. In particular, the occurrence of error at this step was of primary concern. Given evidence for the systematic occurrence of error at this step across the testing sessions, a repeated-measures ANOVA was used to analyze the effect of condition on the occurrence of error at the “Turn Off Tracking” step in the

tactical task. Figures 22 and 23 chart the two measures of error occurrence: frequency of error at the “Turn Off Tracking” step and the proportion of postcompletion errors at the “Turn Off Tracking” step to the total number of errors in the tactical task.

For both measures of error, frequency and error proportion, a reliable within-subjects effect of test period was found. For the frequency of error, $F(3, 162) = 14.30$, $p < 0.001$, and for error proportion, $F(3, 162) = 4.53$, $p = 0.004$. Both of these measures exhibited strong linear components. For the frequency of error, the linear component was $F(1, 54) = 31.35$, $p < 0.001$, and for the proportion of error, the linear component was $F(1, 54) = 8.54$, $p = 0.005$.

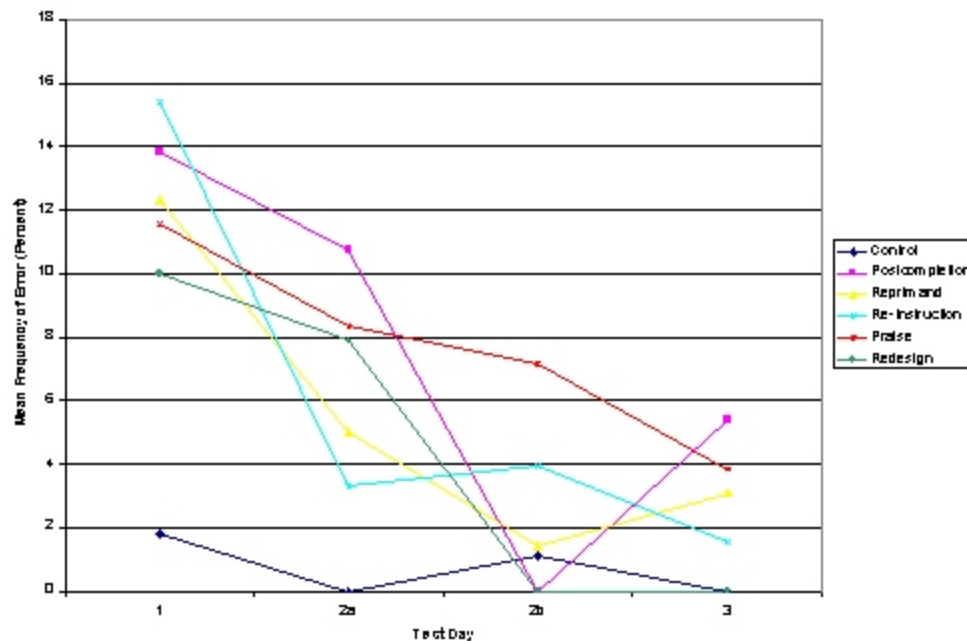


Figure 22. Experiment 2: Frequency of error at the “Turn Off Tracking” step in the tactical task.

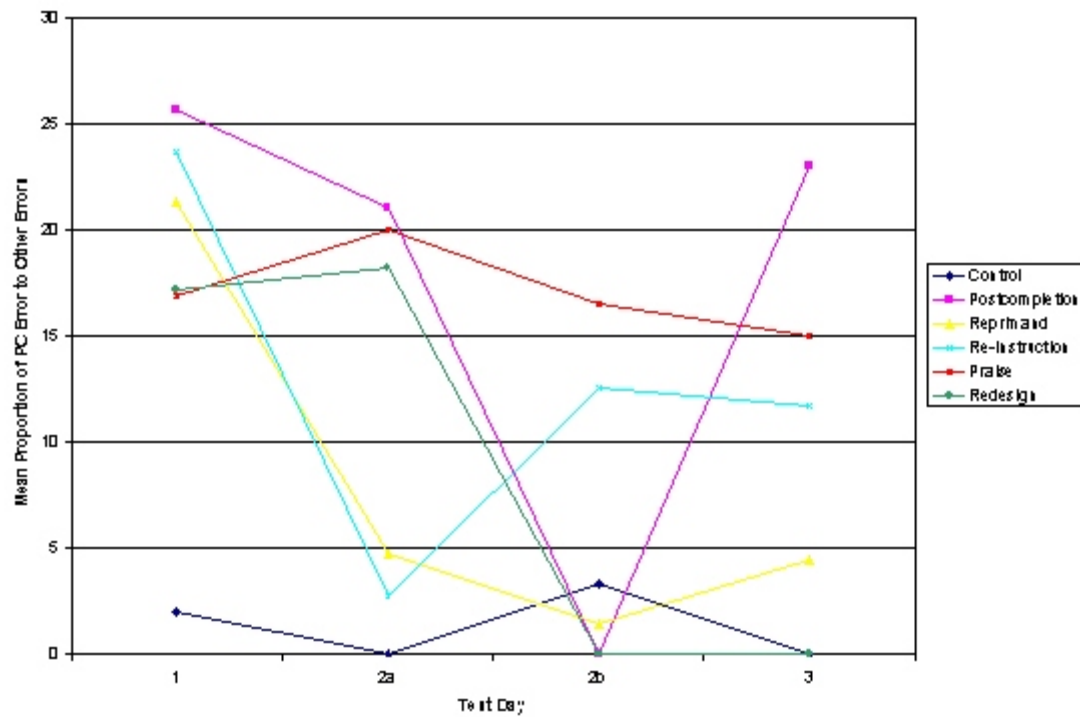


Figure 23. Experiment 2: The proportion of postcompletion errors at the “Turn Off Tracking” step to total error in the tactical task.

Between-subjects analysis revealed a reliable main effect of condition only for the proportion of postcompletion errors to total error at the “Turn Off Tracking” step; for this measure, $F(5, 54) = 2.89$, $p = 0.02$. The frequency measure approached reliability, $F(5, 54) = 2.27$, $p = 0.06$. From Figure 23, it appears that the Control condition may be reliably different from the other conditions. For both the frequency of error and the proportion of error, there was no reliable interaction of condition and test period; for frequency, $F(15, 162) = 1.18$, $p = 0.3$, and for proportion, $F(15, 162) = 1.42$, $p = 0.14$. Several planned comparisons were carried out to assess whether there were different changes in the various experimental conditions over the testing period. It was expected

that the postcompletion conditions with an intervention (Reprimand, Re-Instruction, and Praise) would not differ over time from the Postcompletion baseline condition. It was hypothesized, however, that the postcompletion group with the redesigned tactical task (the control version) would have reliably different error rates over time as a result of the intervention. As could be seen in Figures 22 and 23, however, there appears to be little difference between this Redesign group and the other Postcompletion groups, and in fact, the results did not reveal a reliable difference. Tables 4 and 5 examine the frequency and proportion of error, respectively, and contain the list of planned comparisons that examined the interaction of the test period with the experimental condition. A sequential Bonferroni correction was used to preserve the Type I error rate associated with the family wise error rate. As expected, the postcompletion conditions were not reliably different from each other. However, despite the drop to a zero error rate shown in Figures 22 and 23 after the introduction of the intervention, the Redesign condition was not reliably different from the other interventions for error occurrence.

Table 4

Experiment 2: Planned comparisons for the interaction of test period and condition on error frequency at the “Turn Off Tracking” step in the tactical task

Contrast	df	F	p
Control vs. Postcompletion	1, 54	6.14	0.02
Control vs. All Others	1, 54	5.61	0.02
Postcompletion vs. Reprimand, Re-instruction, and Praise	1, 54	1.53	0.22
Redesign vs. Other Postcompletions	1, 54	0.54	0.46
Redesign vs. Control	1, 54	5.29	0.02
Reprimand vs. Praise	1, 54	0.27	0.60
Re-instruction vs. Reprimand and Praise	1, 54	0.14	0.71
Control vs. Reprimand, Re-instruction, and Praise	1, 54	3.22	0.07

Table 5

Experiment 2: Planned comparisons for the interaction of test period and condition on the proportion of postcompletion error to total error at the “Turn Off Tracking” step in the tactical task

Contrast	df	F	p
Control vs. Postcompletion	1, 54	2.69	0.10
Control vs. All Others	1, 54	2.50	0.12
Postcompletion vs. Reprimand, Re-instruction, and Praise	1, 54	1.33	0.25
Redesign vs. Other Postcompletions	1, 54	3.46	0.06
Redesign vs. Control	1, 54	5.77	0.02
Reprimand vs. Praise	1, 54	0.94	0.33
Re-instruction vs. Reprimand and Praise	1, 54	0.65	0.42
Control vs. Reprimand, Re-instruction, and Praise	1, 54	0.72	0.39

Because of the evidence in Figures 22 and 23 that the Redesign condition reaches a level of zero error occurrence after the introduction of the intervention, it was hypothesized that perhaps the linear effect of learning over the test period was such a powerful main effect that it effectively “washed out” any other effects. To test this, the linear component was partitioned out of the data, and a new non-linear effect of day was calculated for both measures of error. For the frequency of error, partitioning out the linear component resulted in a main effect of day where $F(2, 162) = 4.60$, $p = 0.01$, and for the proportion of error, partitioning out the linear component resulted in a main effect of day where $F(2, 162) = 3.70$, $p = 0.02$. Thus, over time, there was a non-linear main effect of test period on error occurrence.

Due to the reliable main effect of the non-linear component of test day, the linear component was also partitioned out of the condition by test period interaction. For the

frequency of error, the non-linear component of the interaction approached statistical reliability, $F(10, 162) = 1.85, p = 0.055$. For the proportion of error measure, however, the non-linear component of the condition by test day interaction did reach statistical reliability, $F(10, 162) = 2.51, p = 0.007$. Thus, this non-linear component for the condition by test period interaction suggests that there was an effect of condition on error rates as a function of the test period, thereby hinting at the fact that the Redesign manipulation might indeed reduce error rates when compared to the other postcompletion interventions.

The interesting question, however, is whether participants changed position on a speed-accuracy curve for the “Turn Off Tracking” step in the tactical task after the interventions were introduced. Repeated-measures ANOVA on step completion time at the “Turn Off Tracking” step revealed reliable main effects for the within-subjects effect of test period and a reliable between-subjects effect of condition. For the effect of test period, $F(3, 162) = 31.26, p < 0.001$, and for the effect of condition, $F(5, 54) = 32.67, p < 0.001$. Figure 24 displays the mean step completion time at the “Turn Off Tracking” step as a function of condition. As with the error data for this step, several planned comparisons were conducted to assess different changes over time between the conditions. These results are presented in Table 5. A Bonferroni sequential correction was used to preserve Type I error rate and account for the effects of family-wise error. Planned comparisons that were statistically reliable are printed in bold type.

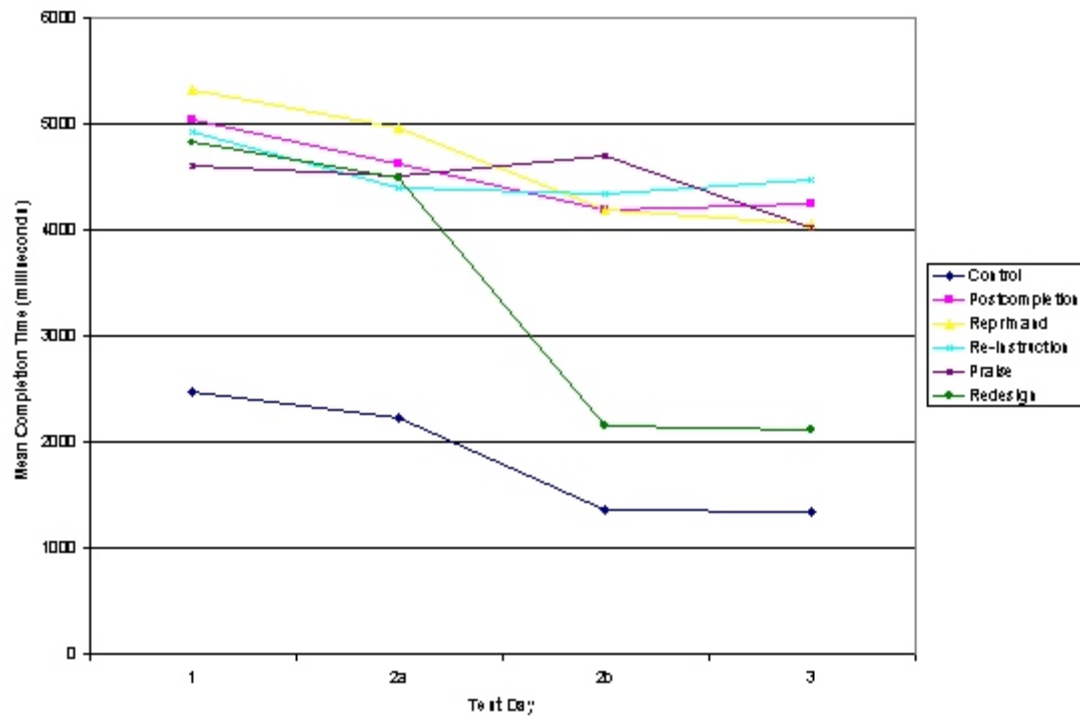


Figure 24. Experiment 2: Mean step completion time as a function of condition at the “Turn Off Tracking” step in the tactical task.

Table 6

Experiment 2: Planned comparisons for the interaction of test period and condition on step completion time at the “Turn Off Tracking” step in the tactical task

Contrast	df	F	p
Control vs. Postcompletion	1, 54	1.23	0.03
Control vs. All Others	1, 54	0.08	0.77
Postcompletion vs. Reprimand, Re-instruction, and Praise	1, 54	0.13	0.72
Redesign vs. Other Postcompletions	1, 54	48.95	< 0.001
Redesign vs. Control	1, 54	17.66	< 0.001
Reprimand vs. Praise	1, 54	5.14	0.03
Re-instruction vs. Reprimand and Praise	1, 54	1.19	0.27
Control vs. Reprimand, Re-instruction, and Praise	1, 54	2.94	0.08

As can be seen in both Figure 24 and Table 6, the postcompletion conditions (Reprimand, Re-instruction, and Praise) did not differ in step completion time over the test period, thereby indicating little effect of the interventions on decreasing step completion time. The Redesign condition, however, showed a reliable decrease in step completion time at the “Turn Off Tracking” step after the switch to the control version of the tactical task. From the chart in Figure 24, it appeared that the Redesign condition dropped to a step completion time after the introduction of the intervention that was roughly equivalent to the step completion time of the Control condition on Day 1 of testing. Over time, perhaps, the Redesign condition might reach a level equivalent to the Control condition. The interesting point in the step completion time data, however, is that despite the fact that there was not a reliable difference in error rates after the introduction of the intervention, there was a reliable difference in step completion times. Thus, the cost of reduced-error performance for the postcompletion conditions at the “Turn Off Tracking” step in the tactical task was an increased step completion time.

Performance at “Power Connected 1” and “Scanner On.” As mentioned previously, the “Power Connected 1” and “Power Connected 2” steps in the tactical task and the “Scanner On” and “Scanner Off” steps in the transporter task offer another opportunity to examine the occurrence of systematic error during task performance. Because the task structures of these two tasks were the same, errors at the identified steps can be compared for differences across the tasks. Since the two interfaces for these tasks were different, this difference essentially serves as a check that a redesign (or a different design) can influence performance. Since there were no interventions directly aimed at

influencing performance at the “Power Connected 1” and “Power Connected 2” steps in the tactical task and the “Scanner On” and “Scanner Off” steps in the transporter task, Day 2 was not split as it was for the “Turn Off Tracking” step in the tactical task.

As mentioned above, the equivalent task structures for the tactical and transporter tasks allows equivalent task steps to be assessed for differences in performance between the two tasks. Thus, a comparison was made between the “Power Connected 1” step in the tactical task and the “Scanner On” step in the transporter task. Refer to Figures 18a and 18b for the task structure of the tactical and transporter tasks, respectively. The difference in the design for these tasks lies in the location of the “Power Connected” check box for the tactical task and the “Scanner On” button for the transporter task. In the tactical task, the “Power Connected” check box is located in a physically different space than the battery meter. This check box must be selected before the goal of charging the battery can be initiated; thus, the “Power Connected” check box exists outside the physical space occupied by the main components of the goal with which it is associated. The “Scanner On” button in the transporter task, however, is grouped with other buttons that operate the scanner display and fulfill the “lock on target signal” goal.

A repeated-measures ANOVA was used to analyze the effect of task (tactical or transporter) and test day on the commission of errors at the “Power Connected 1” and “Scanner On” steps. For the frequency of error at this corresponding step for the two tasks, there was a reliable main effect of task, $F(1, 59) = 40.98$, $p < 0.001$. There was also a reliable main effect for test day, $F(2, 118) = 13.52$, $p < 0.001$. As can be seen in

Figure 25, error occurred at the “Power Connected 1” step in the tactical task much more frequently than at the “Scanner On” step in the transporter task.

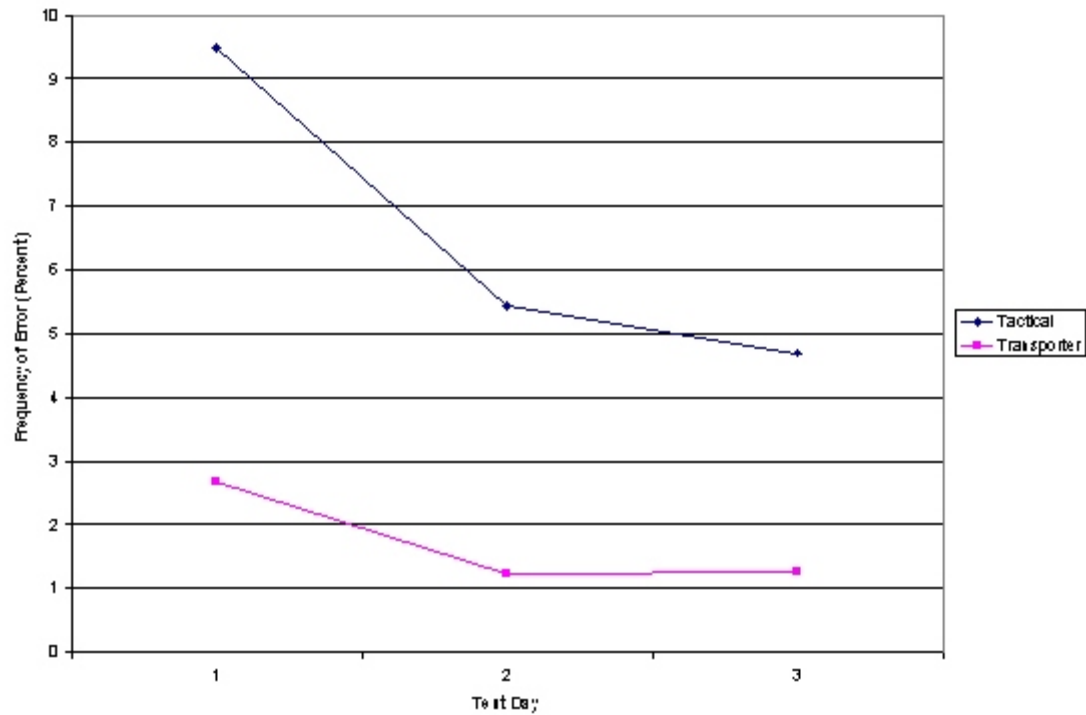


Figure 25. Experiment 2: Comparison of the frequency of error at “Power Connected 1” in the tactical task and “Scanner On” in the transporter task.

This relation held when the proportion of errors to total error at this set of corresponding steps were analyzed. For the proportion measure, there was also a reliable main effect of task, $F(1, 59) = 32.36, p < 0.001$. There was not, however, a reliable main effect of test day on the occurrence of error at these steps. Thus, while the overall error frequency decreased, the proportion of error accounted for at this set of corresponding steps did not change as a function of test day, as can be seen in Figure 26.

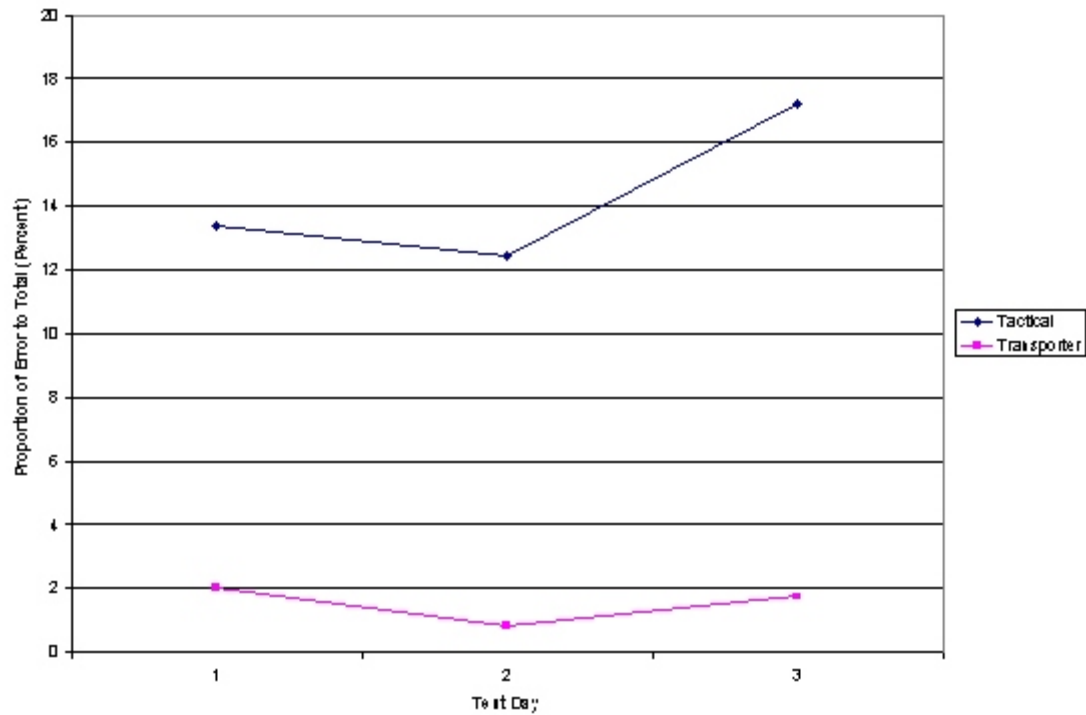


Figure 26. Experiment 2: Comparison of the proportion of error at the “Power Connected 1” step in the tactical task and the “Scanner On” step in the transporter task.

As with the measures used to assess performance at the “Turn Off Tracking” step in the tactical task, step completion time can be assessed for differences between the tactical and transporter tasks. Repeated-measures ANOVA revealed reliable main effects for task, $F(1, 54) = 65.42$, $p < 0.001$ and for test day, $F(3, 162) = 92.04$, $p < 0.001$. As can be seen in Figure 27, participants completed the “Scanner On” step in the transporter task much more quickly than the “Power Connected 1” step in the tactical task.

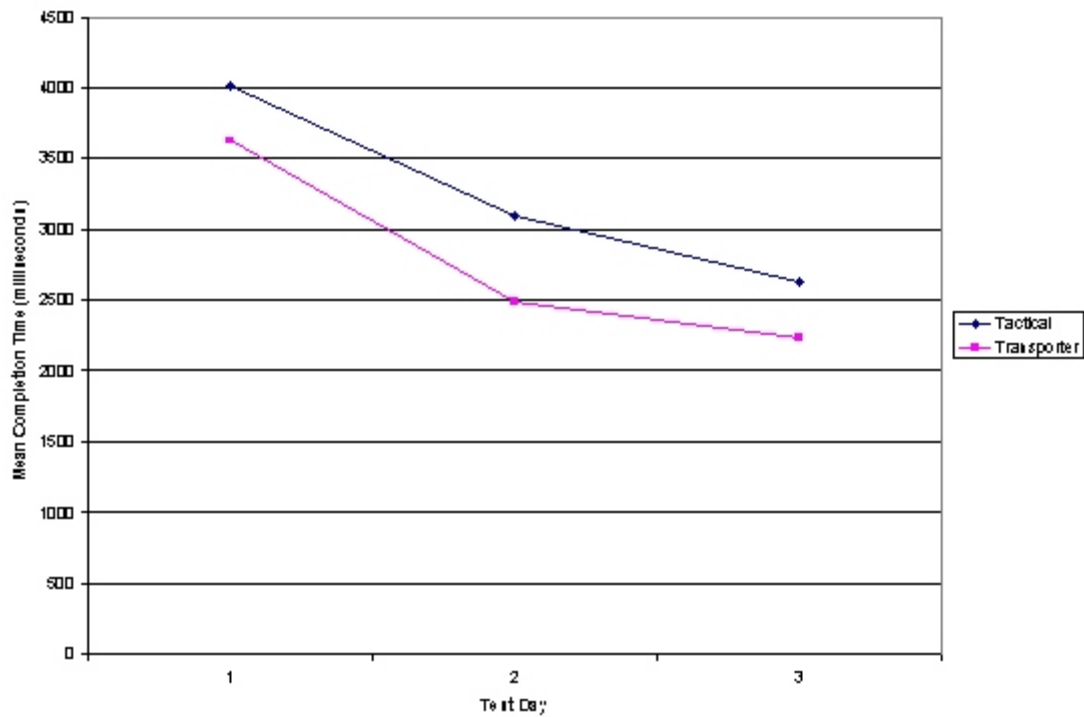


Figure 27. Experiment 2: Mean step completion time for the “Power Connected 1” step in the tactical task and the “Scanner On” step in the transporter task.

It is hypothesized that the difference in the layout of the task screens is responsible for the difference in both error rates and step completion time between the two tasks at this corresponding task step. For example, in the tactical display (refer to Figure 18a), the “Power Connected” check box is physically separated from the “charge battery” controls, and in fact, is located beneath the controls. Thus, participants may be likely to skip the “Power Connected” step and jump to the “Charge” step. The button associated with this step (“Charge”) is located near the top of the display and is physically grouped with the battery meter. In the transporter task, however, the “Scanner On” button is the first button in the set of “lock signal” buttons and is physically grouped with all of

the components of the “lock signal” display. Thus, participants may be less likely to miss executing this step in the transporter task.

Performance at “Power Connected 2” and “Scanner Off.” A further comparison was made between the tactical and transporter tasks with the “Power Connected 2” step from the tactical task and the “Scanner Off” step from the transporter task. In the tactical task, “Power Connected 2” corresponds to disconnecting the power after charging the battery (participants have to click on the “Power Connected” check box). In the transporter task, “Scanner Off” corresponds to turning off the scanner after locking in on the target signal (participants click on the “Scanner Off” button). As mentioned in the discussion for the “Power Connected 1” step in the tactical task and the “Scanner On” step in the transporter task, the physical layout of the two tasks vary on the location of the relevant controls in relation to the other controls associated with their goal. For the tactical task, the “Power Connected” check box is separated from the battery controls, and for the transporter task, the “Scanner Off” button is grouped with the other scanner controls.

A repeated-measures ANOVA was used to analyze the effect of task (tactical or transporter) and test day on the commission of errors at the “Power Connected 2” and “Scanner Off” steps. For the frequency of error at this corresponding step for the tactical and transporter tasks, there was not a reliable main effect of task, $F(1, 59) = 0.015$, $p = 0.9$. Tests of sphericity for error commission as a function of test day resulted in a Huynh-Feldt epsilon of 0.726, indicating a violation of sphericity. The corrected Huynh-Feldt values were used for the main effect of day. There was a reliable main effect for test day,

$\underline{F}(1.45, 118) = 54.98, p < 0.001$. As can be seen in Figure 28, error occurred at “Power Connected 2” in the tactical task at a nearly identical rate as at “Scanner Off” in the transporter task. This relation held when the proportion of errors to total error at this set of corresponding steps were analyzed. For the proportion measure, there still was no reliable main effect of task, $\underline{F}(1, 59) = 0.12, p = 0.73$, nor was there a reliable main effect of test day on the occurrence of error at these steps, $\underline{F}(2, 118) = 2.34, p = 0.1$. Figure 29 displays the proportion of error at the “Power Connected 2” step in the tactical task and the “Scanner Off” step in the transporter task. Thus, for this particular set of corresponding steps, there were no reliable differences between the tactical and transporter tasks for the occurrence of error.

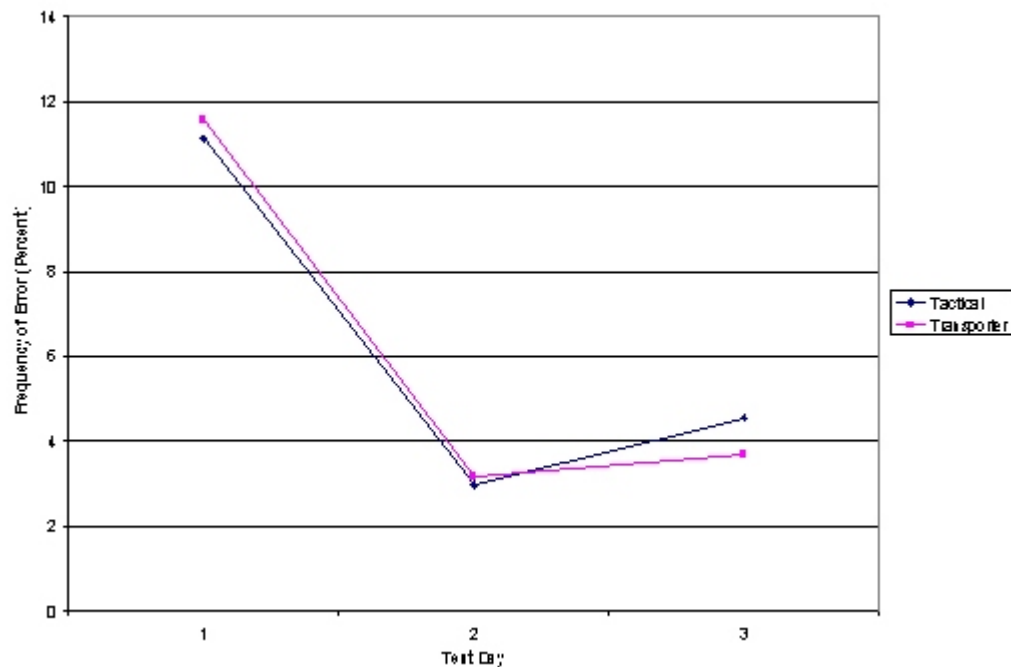


Figure 28. Experiment 2: Comparison of the frequency of error at “Power Connected 2” in the tactical task and “Scanner Off” in the transporter task.

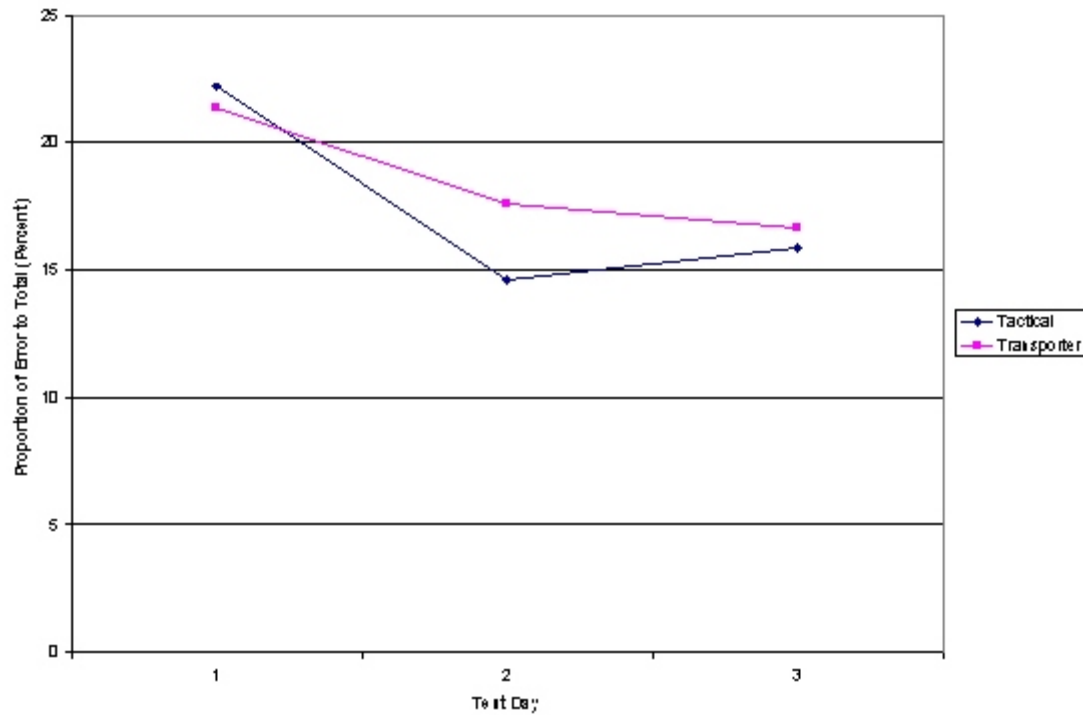


Figure 29. Experiment 2: Comparison of the proportion of error at “Power Connected 2” in the tactical task and “Scanner Off” in the transporter task.

Step completion time was also assessed for the “Power Connected 2” step in the tactical task and the “Scanner Off” step in the transporter task. Repeated-measures ANOVA revealed reliable main effects for task, $F(1, 54) = 8.47$, $p = 0.005$, and for test day, $F(3, 162) = 59.99$, $p < 0.001$. As shown in Figure 30, participants consistently completed the “Scanner Off” step in the transporter task faster than the “Power Connected 2” step in the tactical task. Thus, while there was not a difference in error rates between the tasks, it is hypothesized that a difference in task layout resulted in the difference in step completion time between the tasks. In the tactical task, there is some physical distance between the “Stop Charging” button and the “Power Connected”

button; furthermore, the “Power Connected” button does not share the same display space as the other battery controls. The “Scanner Off” button in the transporter task, however, is located in the same physical display space as the other “lock signal” controls. Thus, it may be easier to locate and recall the “Scanner Off” step when the lock signal goal has been reached.

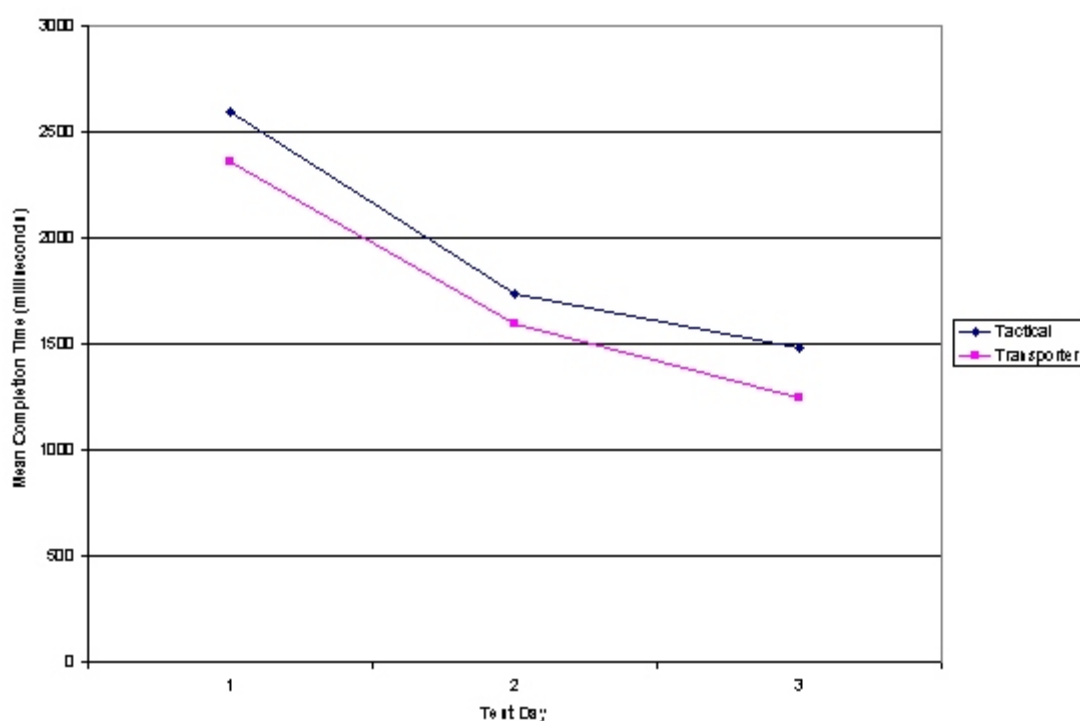


Figure 30. Experiment 2: Mean step completion time for the “Power Connected 2” step in the tactical task and “Scanner Off” in the transporter task.

Letter recall task. The concurrent letter recall task also offers insight into the differences (or lack thereof) between the conditions for error rates. The letter recall task was introduced to induce a load on working memory, thereby making it more likely that errors would occur during task performance. However, the learning exhibited by the

participants on the bridge station tasks seems to indicate that the letter recall task may not be enough to serve as a sufficient load once the tasks had been learned to a skilled level. If this were the case, then it seems likely that performance on the letter recall task would have improved over time. Figure 31 shows the percent of trials correct on the letter recall task as a function of condition.

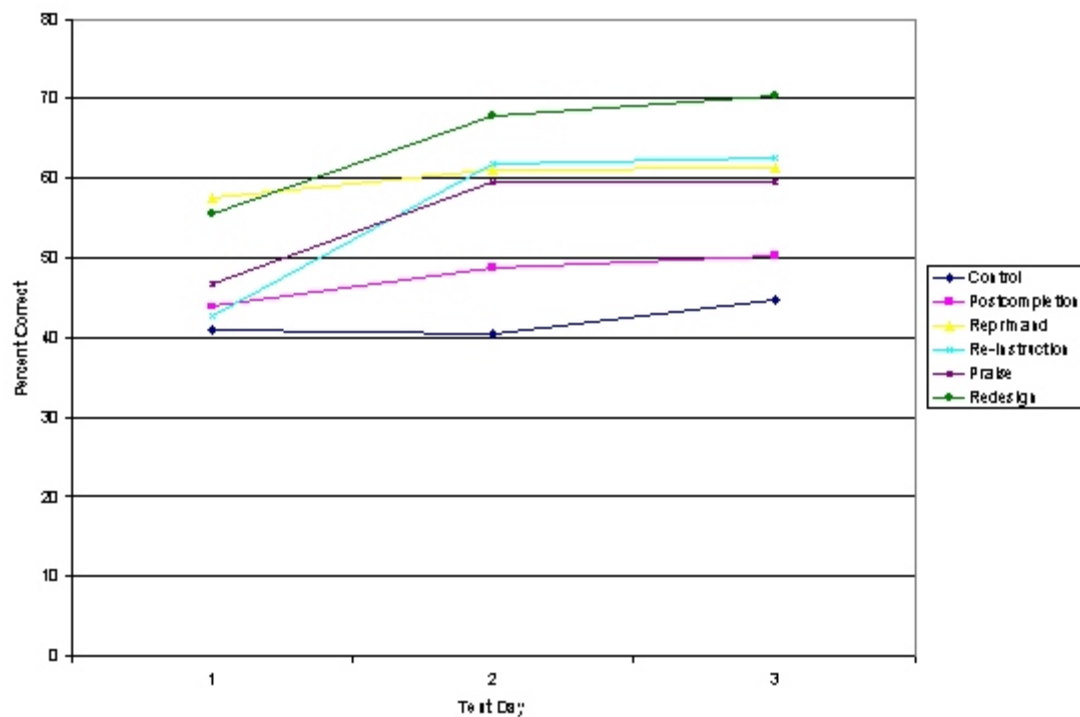


Figure 31. Experiment 2: Mean percent of trials correct on the letter recall task.

Repeated-measures ANOVA revealed that performance on the letter recall task did indeed improve across the testing period, with a main effect of day, $F(2, 108) = 16.18$, $p < 0.001$. There was not a reliable difference between the conditions for the letter recall task, indicating that the interventions did not affect performance on this secondary task.

Working memory capacity. Finally, several correlations were computed to assess the relations between working memory capacity and task performance. It was found that working memory capacity was reliably related with the percent of trials correct for the letter recall task, $r = .522$, $p < .005$, the overall mean task completion time, $r = -.340$, $p = .008$, and the overall frequency of error, $r = -.537$, $p < .0005$. Further correlations were found to exist between performance on the letter recall task and overall error frequency, $r = -.454$, $p < .0005$, and for the overall mean task completion time to the overall frequency of error, $r = .398$, $p = .002$.

Discussion

The general results of Experiment 2 were consistent with those of Experiment 1. First, the postcompletion error at the “Turn Off Tracking” step in the tactical task occurred over the three test days at a rate much higher than would be predicted by a stochastic error theory. Furthermore, participants with the control version of the tactical task made reliably fewer errors at the “Turn Off Tracking” step than participants with the postcompletion version. The results of Experiment 2 also exhibited the effects of task learning evident in Experiment 1. Over the course of the three test sessions, overall error rates decreased, as did the occurrence of the postcompletion error at the “Turn Off Tracking” step.

The results regarding the effect of the various interventions on subsequent task performance were fairly complex. The hypothesis that redesigning the problematical task (tactical) would result in reliably fewer errors than other methods (reprimands, re-instruction, and praise) was partially supported by the results. The results did not support

the claim that participants in the Redesign condition performed at a level that was reliably different from participants in other conditions. While Figures 22 and 23 show that participants in the Redesign condition did not make any errors at the “Turn Off Tracking” step after the introduction of the intervention, the error rates for all of the other conditions were not different enough to argue that the redesign intervention was systematically better at reducing the occurrence of error. The level of expertise that participants reached over the course of testing made the occurrence of any errors rare, including the postcompletion error at the “Turn Off Tracking” step in the tactical task, and it therefore became nearly impossible to find differences between the various conditions on the commission of error.

One interesting result to note, however, was the lack of difference in error commission between the Reprimand, Re-Instruction, and Praise conditions. According to the “conventional wisdom” perspective of error, errors occur because humans are inattentive, careless, negligent, and lazy, thereby tying error commission to intrinsic qualities of the individual. According to this perspective, reprimands serve the purpose of discouraging a particular behavior, while praise is intended to encourage the further occurrence of a particular behavior. Re-instruction is often mandated to discourage the occurrence of particular behaviors, with the intent that by re-instructing individuals on the “proper” way to complete a task, undesired behaviors will be abandoned. Individuals undergoing mandated re-instruction often view such remedial action as punishment - “I already know how to do x . I’ve been doing it every day here for the last 10 years. Just because I forgot to do y that one time, I have to waste my time sitting here for the next week listening to someone tell me how to do something that I could teach better than that

fool up there . . .” The underlying assumption of these responses to task performance is that the individual’s performance is under that individual’s *conscious* control. By reprimanding poor performance, it is assumed that an individual will consciously choose to change his behavior so that it matches some external criterion of acceptable behavior. The fact that these methods (reprimand, re-instruction, and praise) did not have reliable effects on the occurrence of subsequent error beyond the learning effects mentioned previously, points to support for a systems perspective of error and its underlying assumptions that the occurrence of error is seldom, if ever, under direct conscious control and is a product of the mismatch between the systems we use and the demands these systems place upon human abilities and limitations.

While there was no evidence for differences in the occurrence of error between the Reprimand, Re-Instruction, and Praise conditions, qualitative analysis of the exit questionnaire revealed interesting findings. In particular, the interventions “worked” in so far as the participants in these conditions felt reprimanded or praised, and attempted to accordingly alter their behavior. For example, in the Reprimand and Re-Instruction conditions, where a report of poor performance was part of the intervention, participants reported feeling “irritated,” “slow and inept,” “like a retard,” “very bad,” “annoyed and frustrated,” and “pressured to improve.” These responses certainly seem to reflect that these participants legitimately felt reprimanded. A similar phenomenon was found for the participants in the Praise condition, who reported feeling “good,” “proud,” “confident and relaxed,” “happy,” “encouraged,” and “motivated to achieve an even higher score” after receiving a report of excellent task performance.

In terms of the qualitative response to the intervention reports of task performance, a marked difference existed between the Praise condition and the Reprimand and Re-Instruction conditions. None of the participants in the Praise condition reported consciously changing their behavior or strategy for completing the experimental tasks (the bridge station tasks and letter recall task). Participants in the Reprimand and Re-instruction conditions, however, reported various changes in behavior and strategy. For example, some participants tried to complete the tasks more quickly, while others indicated that they tried to be more accurate, even if they had to slow down. Other participants indicated a shift in attention from the bridge station tasks to the letter recall task, while still others shifted attention the other way: from the letter recall task to the bridge station tasks.

Despite the assertions for a shift in attention, the aggregated results of the participants do not support this qualitative assessment. This was reflected in a question on the exit questionnaire where participants were asked to identify how much of their attention was directed toward the bridge station tasks, and how much toward the letter recall task. Participants were instructed to assign a number from one to ten to each of these activities so that their responses added up to ten. For participants in conditions that received an intervention, the mean amount of attention devoted to the bridge station tasks was 5.6 before the intervention and 5.65 after the intervention. For the letter recall task, the mean amount of attention devoted before the intervention was 4.4 and 4.35 after the intervention. Thus, even though participants assumed they were “switching” their attention, there seems to be little support for this assessment.

The key finding in this experiment pertains to the step completion time associated with the “Turn Off Tracking” postcompletion step in the tactical task. As shown in Figure 24, the step completion time for the Redesign condition declined precipitously from just prior to the introduction of the intervention to after its introduction. No such change in the step completion time was observed in any of the other conditions for the effect of the intervention. While it was hoped that such a marked result would also have occurred for the commission of error at the “Turn Off Tracking” step, so few errors were committed overall by the last day of testing, that it was virtually impossible to truly assess the effects of the interventions on error commission.

The strength of the change in step completion time for the Redesign participants becomes even more convincing when Figure 21 is considered, which reveals that there were no reliable differences between the conditions for the overall time it took to complete the tactical task. The fact that a difference exists at the “Turn Off Tracking” step marks it as important. The introduction of the control version of the tactical task precipitates a move on the speed-accuracy curve for these participants; they were moved to what could be argued was a more optimal position on the curve. It is hypothesized that the decrease in step completion time is due to the fact that participants had to click the “Turn Off Tracking” button to learn the outcome of firing on the target. Because participants with the postcompletion version of the tactical task did not need to execute the “Turn Off Tracking” step to learn the outcome of the firing action, there is less incentive to recall that this step must still be executed before exiting the task. The “incentive” of learning the outcome of the firing action creates a forcing function that

automatically moves the Redesign participants to an optimal position on the speed accuracy curve after they switch to the control version of the tactical task from the postcompletion version. Because the “Turn Off Tracking” step becomes linked to learning the status of the system, participants are quick to execute the step, and seldom make errors regarding the execution of this step.

The participants with the postcompletion version of the tactical task (Reprimand, Re-instruction, and Praise) operate from a different location on the speed-accuracy curve. Over time, these participants exhibited improved performance - they seldom made the error at the “Turn Off Tracking” step. The cost of this accurate performance, however, was a longer step execution time. This movement on the speed-accuracy curve was consistent with some of the reported strategies adopted by participants in response to the “motivational” interventions. Several participants reported trying to be more accurate on the bridge station tasks, even if they slowed down their task performance to do so. Thus, to a certain extent, participants could change their own behavior in response to information about their performance. The performance associated with these conscious changes, however, cannot compare to the performance associated with the automatic change along the speed-accuracy curve forced by the introduction of the control version of the tactical task to Redesign participants.

GENERAL DISCUSSION

The experiments reported here were intended to address the role organizational responses to performance have on subsequent task performance. In particular, the effects of various responses to error during task performance were examined. There are generally two approaches that are taken in response to error. The first approach adopts a “conventional wisdom” perspective to error commission. According to this perspective, errors occur because humans are lazy, careless, negligent, and inattentive. These traits are viewed as integral to an individual’s nature, and are outwardly reflected as low motivation and error-prone performance. Thus, individuals with low motivation for task performance are viewed as being more lazy, careless, negligent, and inattentive, and therefore more error prone, than more motivated-appearing individuals. Because this perspective holds that error commission is intrinsic to the individual, responses to error are directed at the individual (i.e., reprimands), with the intention that these “motivational” responses will decrease the commission of errors in the future as reflected by a *choice* of the *individual* to behave in a different manner during task performance.

A second approach that can be taken in response to error is a systems approach. This perspective acknowledges that individuals do indeed make errors, and that these errors often arise as a result of human interaction with technology; error is a consequence of the use of technology, its environment, and the conventions, conditions, and procedures for its use, and the way these factors interact with and challenge the nature of human abilities and limitations. As such, the systems approach does not target single individuals as the guilty perpetrators of error, rather, it looks for patterns of error and

works to decrease the mismatch between the system's requirements and human abilities and limitations by changing the system to accommodate those abilities and limitations.

As discussed previously, there does appear to be some evidence that a systems perspective of error occurrence and its corresponding response to such error results in improved task performance. The application of the systems approach in this experiment took the form of redesigning one of the experimental tasks (tactical) to address a design flaw that allowed participants to make a postcompletion error at a particular step ("Turn Off Tracking") in the procedure. Other participants received "organizational" responses following the "conventional wisdom" perspective of error, these participants were either reprimanded for poor performance, reprimanded for poor performance and retrained on the task procedure, or praised for excellent performance. While a reliable difference was not found between the group with the redesigned task and the groups receiving an organizational intervention for the occurrence of error at the "Turn Off Tracking" step in the tactical task, the fact remains that once the procedure was redesigned, the participants in this group made *no errors* at this step on the last day of testing. The groups receiving organizational interventions were still making the occasional error at this step on the last day of testing.

The greatest difference between the Redesign group and the organizational intervention groups was the amount of time it took to complete the "Turn Off Tracking" step. Before the introduction of the redesigned tactical task, the participants in the Redesign group took the same amount of time to complete this step as the participants in the organizational intervention groups. After the redesigned task was introduced,

however, these participants had a reliably faster step completion time than the participants receiving organizational interventions. The participants in the Redesign group made a move on the speed-accuracy curve that was different than their organizational intervention counterparts: the Redesign participants became reliably quicker at the postcompletion step (“Turn Off Tracking”), as well as committing fewer errors (but not in a statistically reliable fashion). The participants in the organizational intervention conditions, however, did not change their step completion time after the introduction of their interventions. Thus, these participants “paid” for their decrease in error rates over time with an increased step completion time; they had to pay more attention (and therefore more time) at the “Turn Off Tracking” step than the redesign participants, in order to avoid making the error.

General Model of Task Performance over Time

Beyond the influence of the “conventional wisdom” and systems approaches to error, the current research exhibited a general model of task performance over time. This model is best described as an integration and synthesis of several cognitive disciplines; in particular, skill acquisition and learning, working memory, and error.

Cognitive Components of Task Performance

Skill Acquisition in Learning

Fitts and Posner (1967) provide a model of skill acquisition that forms the familiar basis of both Rasmussen’s (1987) and Reason’s (1990) systems theories of error. According to the model proposed by Fitts and Posner, there are three phases of skill learning: the early or cognitive phase, the intermediate or associative phase, and the final

or autonomous phase. In the early, or cognitive, phase, individuals are working to understand the task and what it demands of their abilities. At this point, individuals are attending to cues, events, and responses that will later go unnoticed during task performance. In other words, performance is extremely goal-oriented and focused on the individual components (cues, events, and responses) that will produce the desired goal.

Performance at the intermediate, or associative, phase reflects the synthesis of the individual components from the cognitive phase into coherent patterns of action. The patterns that were learned as individual components from the early phase of skill learning are “tried out” and formed together into new patterns of action. While the individual components are being synthesized into new patterns of action, and individuals reflect a relatively competent level of performance, they may not yet be aware of all the nuances involved in task performance. At this point in skill acquisition, “errors (grossly inappropriate subroutines, wrong sequences of acts, and responses to the wrong cues), which are often frequent at first, are gradually eliminated” (Fitts & Posner, 1967, p. 12).

Lastly, at the final, or autonomous, phase of skill learning, the component processes become increasingly autonomous - they are less subject to cognitive control and to the interference from other ongoing activities or environmental distractions. This was demonstrated in a study by Bahrick, Noble, and Fitts (1954; as cited in Fitts & Posner, 1967), where participants were tested on two concurrent tasks. The first task was a key-press task in response to a light. For one group, the light appeared at regular intervals, while for a second group, the light appeared at random intervals. The second task was the performance of arithmetic operations on orally presented numbers. The initial results

showed comparable results on the arithmetic task for the two groups (regular and random light appearance interval). With further practice over time, however, the group with regular intervals of light appearance scored higher on the arithmetic task than the group with the random interval of light appearance. Fitts and Posner (1967) suggest that these “results specifically support the idea that continued practice on a predictable activity not only renders that activity less susceptible to interference from a second task but permits the subject to allocate more of his capacity to the second task, thus indirectly enhancing performance on that task as well” (p. 15).

Working Memory

While there are several theories and models of working memory, one that seems appropriate in this research is a model presented by Baddeley and Logie (1999). According to this model, working memory is comprised of multiple specialized components of cognition that allow humans to comprehend and naturally represent their immediate environment, retain information about past experience, support the acquisition of new knowledge, solve problems, and formulate, relate, and act on current goals. In this framework, working memory consists of specialized components that include a supervisory system (the central executive), and temporary memory systems. These temporary memory systems are used to actively maintain memory traces that overlap with those that are involved in perception through the utilization of rehearsal mechanisms. There are two of these temporary memory systems: a phonologically based system (the phonological loop), and a visuospatial store (the visuospatial sketchpad). The central executive is involved in the control and regulation of the working memory system

through various executive functions, such as coordinating the slave systems (phonological loop and visuospatial sketchpad), focusing and switching attention, and activating memory representations in long-term memory.

According to the Baddeley and Logie (1999) model, the specialized temporary memory systems are thought to have constraints on their capacity that are commensurate with the special function that each provides. For example, in the phonological loop, it is assumed that individual differences in phonological loop capacity reflect the amount of memory activation available. These individual differences are assumed to exist in rehearsal capacity, which occurs in “real time.” In the visuospatial sketchpad, a dissociation exists between the capacity for retaining sequences of movements and the capacity for retaining visual patterns. For capacity limits on the sequence of movements, retention of movement or of paths between objects and locations does not necessarily depend on visual perceptual input. For example, input can come in the form of touch (i.e., braille). Thus, it is not the modality that influences capacity limits, but the complexity of the sequence; more complex sequences are more likely to be incompletely processed. For visual patterns, capacity limits are tied to the similarity of the items to be remembered to each other, and the number of items.

The role of the central executive is to control and coordinate the two slave systems, as well as to focus attention, switch attention, and activate memory representations within long-term memory. Individual performance, however, is not merely a product of the processing capacity of the central executive. Rather, the temporary storage embodied in the slave systems and the controlled attention provided by

the central executive each make semi-independent contributions to individual variations in performance. Thus, the underlying assumption is that as memory demands increase, taxing the capacity of a slave system, the central executive devotes more controlled attention to the task, leaving less capacity available for further and additional processing.

The limits defined by this model provide a clear role for working memory in a complex cognitive activity such as perceptuomotor control. According to this model, the general cognitive load will be very high in the early stages of learning the task, and motor control over the actions will be very poorly deployed. As a result of this high general cognitive load, any secondary cognitive load will be sufficient to disrupt performance on the primary task. With the acquisition of expertise, however, the general purpose cognitive resources are replaced by more specialized resources that require less directive attention for execution.

Error

Gray (2000) examined the nature of errors in a simple rule-based task - programming a VCR. This examination was guided by a least-effort principle, where it was assumed that, all else being equal, the behavior that individuals adopt to interact with an object arises from an implicit attempt to minimize cost while maximizing benefits. This approach resulted in a control structure for the “program VCR” task that can be characterized by the principle of display-based difference-reduction. This principle is associated with what Gray calls the strategy of least effort in place-keeping, where “place-keeping entails knowing what parts of the task have been completed and what parts remain to be accomplished” (p. 221). Place-keeping can occur at two levels of scope

during task performance: global and local. Global place-keeping keeps track of what goals have been accomplished and what goals remain to be accomplished, while local place-keeping keeps track of progress on the current goal.

For Gray's (2000) task of programming the VCR, the cognitive burden of place-keeping was largely alleviated through what he calls display-based difference-reduction. According to the general principle of difference-reduction, an individual proceeds through a task by gradually reducing the difference between the current state of the world and the goal state. When difference-reduction is display-based, as in the VCR programming task, the device in question displays information about the various goal states. Thus, as each step makes the current state more similar to the goal state, the past states do not need to be retained, and planning more than the next step ahead is not required. Display-based difference-reduction usually takes form as a partial implementation: in terms of global place-keeping, the individual needs to remember the task-based and device-specific goals, but does not have to remember the states of these goals. Thus, local place-keeping is transformed from a predominantly cognitive task to a predominantly perceptual one.

During task performance, errors can fall into two broad categories: push and pop. Push errors are those that occur while setting goals and subgoals. Pop errors are errors in completing goals and subgoals. Both categories of error have specific cases that relate to display-based difference-reduction. For push errors, a particular device might have a mode that indicates that an error occurred during previous task performance (i.e., a "verify program" mode on a VCR). The display-based difference-reduction paradigm

suggests that a certain goal needs to be accomplished (i.e., correct the incorrect program information), but the device design prohibits this from being accomplished without changing mode first. Errors occur when individuals try to correct the past behavior in what amounts to an “information-only” mode that does not allow corrections.

For pop errors, one type of error was referred to as “premature pops.” These premature pops can occur during task performance, when the correct goal is interrupted before it is completed. Gray (2000) hypothesized that if goals are simply assumed to be chunks in declarative memory, then it is possible that goals may compete with each other. In a display-based difference-reduction environment, goal competition may be particularly problematic, as reminders of alternative goals are readily apparent on the visual display.

Integrated Model of General Task Performance over Time

The models described here for skill learning, working memory, and error in terms of display-based difference reduction all offer tantalizing bits of a puzzle that forms a picture of task performance. In the experiments reported here, task performance followed a general pattern: participants made the most errors (both errors in general, and postcompletion errors) on the first day of testing, and the occurrence of error decreased across the course of testing. The postcompletion error at the “Turn Off Tracking” step in the tactical task, while not systematic by Day 3 of testing, still occurred. The step completion time for the “Turn Off Tracking” step did not reliably change over the three test days for the participants in the Reprimand, Re-Instruction, and Praise conditions. However, participants in the Redesign condition had a reliably faster step completion

time for the “Turn Off Tracking” step than the other conditions after the introduction of the intervention. In terms of the secondary letter recall task, performance also reliably improved across the three test days for all of the experimental conditions.

By piecing together the different models described above, a unified picture of task performance emerges. The strong linear effects for the decrease in the occurrence of error in both Experiment 1 and Experiment 2 suggest that participants were moving along a skill acquisition curve over the three days of testing. Despite a day of “Training” and acknowledged performance at some pre-determined standard of proficiency (three error-free trials), it could be argued that on Day 1 of testing, participants were still learning the tasks and fell into the early or cognitive phase of skill learning as proposed by Fitts and Posner (1967). Thus, even though participants were familiar with the four bridge station tasks and their goals, they were hardly experts, and their attention was focused accordingly on the execution of the bridge station tasks, rather than the secondary letter recall task. Evidence from the Exit Questionnaire as discussed previously tends to support this “balance of attention.” As reported earlier, participants were asked to identify how much of their attention was directed toward the bridge station tasks, and how much toward the letter recall task by assigning a number from one to ten to each of these activities so their responses added up to ten. Across all participants, for the period before the interventions on Day 2, participants reported devoting a mean 5.95 % of their attentional resources to the bridge station tasks, and only 4.05 % to the letter recall task.

These findings are consistent with Baddeley and Logie’s (1999) working memory model, where it would be hypothesized that all components of the system should be taxed

on Day 1 of testing. According to this model, the general cognitive load is very high during the early stages of skill learning, and motor control is poorly executed.

Additionally, any secondary cognitive load is sufficient to disrupt performance. Within the characteristics established by this model, it would seem reasonable that the two performance activities in the experiment (the bridge station tasks and letter recall) tap both the phonological loop and visuospatial sketchpad slave systems. Performance on the letter recall task falls subject to the rehearsal mechanisms of the phonological loop, and the bridge station tasks fall subject to the component of the visuospatial sketchpad responsible for retaining sequences of movements. Retaining and executing the correct sequence of movements for each of the bridge station tasks pushes the limits of the visuospatial sketchpad, where the central executive then devotes more controlled attention to executing these tasks. This leaves little directive capacity in the central executive for overseeing the rehearsal mechanisms of the phonological loop on the letter recall task. The reliable correlations of working memory capacity to the overall frequency of error during task performance, the percent of trials correct for the letter recall task, and overall mean task completion time supports this hypothesis. All else being equal, those participants with a greater working memory capacity were able to complete the tasks with fewer errors, devote more attention to the letter recall task for a correspondingly higher percentage of correct responses, and complete the tasks more quickly than participants with shorter working memory capacity spans.

Given Gray's (2000) model for display-based difference-reduction, it could be further argued that on Day 1, the participants were learning to recognize the different

visual states of the task displays, and their relation to the task goal. At this early stage of learning, Gray noted that push errors, errors in goal-setting, would be highly probable. In particular, these errors could be exacerbated by the difference-reduction display that revealed where erroneous actions had occurred, but could not be corrected from the current device state. Indeed, while not discussed in the results, this phenomenon was observed by the experimenter during participant task performance. For example, in the tactical task, participants were instructed to charge the battery until it was within certain tolerance limits. Occasionally, a participant would stop charging the battery before it had reached the lower control limit. Upon discovering this error, as it was readily visible on the task display, participants would try to start charging the battery again, an action that could not be taken once the “Stop Charging” action had been executed. This model also supports the idea that the visuospatial sketchpad in working memory was pushing its capacity limits. As introduced earlier, display-based difference-reduction transfers the burden of local place-keeping (progress toward a goal) from a “cognitive” level to a perceptual level. However, participants must learn what actions influence the task display and how these relate to the task goal.

This is related to Rasmussen’s (1987) “skill-rule-knowledge” classification of error and the discussion of knowledge-based behavior discussed earlier. Recall that Rasmussen described performance at the knowledge-based level in the following manner: “a useful plan is developed - by selection, such that different plans are considered and their effect tested against the goal, physically by trial and error, or conceptually by means of understanding the functional properties of the environment and prediction of the

effects of the plan considered” (p. 55). In essence, this describes the process by which the participants were transforming their cognitive knowledge of the task in the visuospatial sketchpad to a perceptual level through display-based difference-reduction. Thus, it could be proposed that participants were undergoing this transformation process on Day 1, as they progressed through the early stage of skill acquisition.

As performance improved on Day 2, participants moved to the intermediate/associative phase of skill learning. At this point, they were familiar with the task structure for all of the tasks, but had not quite discovered all of the nuances associated with the tasks. For example, in the tactical task, a probability function made it more unlikely that a target would be destroyed if it was lined up directly with the tracking cross-hairs. There was a higher probability of destroying the target if it was in a zone that surrounded the cross-hairs. At this associative level, participants were fairly competent in executing their knowledge of the bridge station task structures, but they may not have yet discovered the “best” (least-effort) method of obtaining the task goals.

The working memory model would argue that by Day 2, where participants were relatively secure in their knowledge of task execution, memory capacity was not taxed, allowing participants to commit fewer errors, both on the bridge stations tasks and the letter recall task. This assertion is supported by the linear trends in the experiment data, that reveal fewer error commissions on the bridge station tasks over time, as well an increase in the percent of trials correct for the letter recall task.

By Day 2 of testing, the participants should have developed a relatively strong model of display-based difference-reduction. This claim was demonstrated by the

participants in the Re-Instruction condition. During Day 2, these participants received an intervention where they were informed that their performance was poor and that they should re-read the instruction manuals for the tactical and transporter tasks. After reading each manual, these participants completed a pencil-and-paper test where they were instructed to take a randomly ordered list of the task steps and put them in the correct order for achieving the task goal. While most of these participants turned in “perfect” test papers, it took them several minutes to complete the task, and several test papers had steps crossed out and re-ordered. The apparent difficulty of this task seems to suggest that the bridge station tasks were completed by employing a display-based difference-reduction. Without the information provided by the changing screen states, participants had a difficult time ordering the steps for the procedure. Their difficulty would also argue for the storage of the bridge task execution sequences in the visuospatial sketchpad. By forcing the participants to order the written statements of task steps, they had to recall these sequences from the visuospatial sketchpad and convert them to declarations that could be manipulated in the phonological loop.

At the beginning of Day 3, it could be argued that participants were either very advanced in the intermediate/associative phase, or just entering the final/autonomous phase of skill learning. At this point, the component processes of the bridge station tasks were largely autonomous, due in large part to the display-based difference-reduction. The data from Day 3 indicate that participants were making fewer errors on the bridge station tasks and the letter recall task. This relates back to the Bahrick, Noble, and Fitts (1954) study cited by Fitts and Posner (1967), where practice on a predictable task improved

performance on a secondary task. Other than occurring in random order, the individual bridge station tasks were relatively predictable in the sense that their task structures did not change (except in the tactical task for the Redesign condition). Because performance on these tasks had reached a relatively practiced level, and required little capacity from working memory to complete, more attention could be directed to the letter recall task, resulting in a correspondingly higher percentage of correct trials.

The display-based difference-reduction model offers an explanation for the occurrence of the postcompletion error at the “Turn Off Tracking” step in the tactical task and the difference in step completion time for the Redesign condition on Day 3. According to Gray (2000), premature pop errors occur when a correct goal is interrupted before it is completed. The “Turn Off Tracking” postcompletion error is essentially a premature pop error: participants interrupt the correct goal of “Turn Off Tracking” in the postcompletion version of the tactical task before it is completed in favor of exiting the task. Changing the task structure to the control version, however, virtually eliminates this error, as evidenced by participant performance in the Redesign condition. In the postcompletion conditions (Postcompletion, Reprimand, Re-Instruction, and Praise), the display-state after destroying the target lures the participants into making the premature pop error. Upon firing, the participants are informed of the outcome of this action before they execute the “Turn Off Tracking” step. Thus, because the display-state shows a completed goal, the difference-reduction at this point is essentially zero. The cost of not committing the “Turn Off Tracking” error as often as might be expected is revealed in the

longer step completion times (up to 2 or 3 seconds longer than the Redesign group) of the postcompletion groups.

The display-based difference-reduction model of task performance helps the participants in the Redesign condition, who switch during Day 2 from the postcompletion version of the tactical task to the control version. In the control version of the task, the display-state reveals no information about the firing action until the “Turn Off Tracking” step is executed. Thus, after firing, the difference-reduction between the current display-state and the desired goal state is one - participants must execute one more step to match the display to the goal (determining the status of the target). Linking the “Turn Off Tracking” step to system status decreases the time it takes participants to execute this step, as it is the next logical step in moving the display-state to the goal-state.

Conclusion

This research has several implications from an applied perspective. In applied settings, the systems approach to error presented throughout this paper would argue that errors could be predicted under certain circumstances. The high error rates produced by participants on Day 1 of testing seem to indicate that new, or unpracticed, tasks are likely to have higher error rates associated with them until a certain level of skill is achieved. The distinction between new and unpracticed tasks is a subtle but important one. For example, it is relatively likely that a new employee may have high error rates and demonstrate a “learning curve” in her performance as she learns the new procedures and policies that she must practice in her new position. Unpracticed tasks, however, are slightly different. In this case, it is assumed that an individual has a certain level of

proficiency on a task but does not have to perform it often. Consider the fact that most organizations require their employees to receive yearly certification in “fire safety.”

While the employees have satisfied a criterion for proficiency, their “fire safety skills” are generally applied so seldom that one could expect performance to be error-prone.

Thus, the value of a systems approach lies in identifying where individuals may make errors and redesigning the system to accommodate and recover from error when it does occur. For example, in the case of a new employee, an established employee might be assigned to verify the quality and accuracy of the new employee’s work over a predefined probation period. In the example of seldom performed tasks, such as fire safety, the role of the systems approach is to identify the task as one that is seldom performed and redesign the task so that it can be executed effectively when the need arises.

In terms of tasks which operators perform often, other applied implications become important. For example, if operators are far enough out on the skill learning curve, even under a working memory load, it is unlikely that they will commit errors. This, however, assumes a perfect world where time pressures on task performance do not exist. The results reported in the experiments seemed to indicate that participants’ resistance to error on the postcompletion step (“Turn Off Tracking” in the tactical task) came at the cost of time - participants took longer to execute this step than those participants in a control version (no postcompletion step). While the points system established in Experiment 2 penalized participants for errors, it did not penalize them in terms of time. Thus, a possible trade-off made by participants may have been to slow

down at the “Turn Off Tracking” step to ensure correct execution. In applied settings, however, there is seldom the luxury of unlimited time to complete tasks. Imagine an operating suite where, under “normal” circumstances, a surgeon seldom (if ever) makes a particular error that is predicated by a postcompletion step. Now imagine the same surgeon with a patient that is going into arrest, and the surgeon must complete the procedure he is executing (that has a postcompletion step) with haste. Denied the luxury of the extra few seconds it takes to remember the correct procedure, an error with adverse outcomes could occur.

The experimental results may have exhibited a different pattern had the participants been operating under external time pressure (i.e., points were taken off for not executing a given step within a pre-defined time period). Participants may have changed their criterion on the speed-accuracy trade-off and moved towards faster task execution, and possibly, correspondingly higher error. Gray (2000) indicated that as expertise on a routine task increases, the occurrence of premature pop errors may increase. This could certainly be true if time constraints were an issue. If an individual is under pressure to complete a task as quickly as possible, once the display-based difference-reduction was at a value of zero (the display-state matched the goal-state), despite the fact that a step remained in the procedure, the individual may very well be more likely to commit the premature pop error.

While the data seem to indicate that as task performance becomes more practiced, errors become exceedingly rare, the caveat of human error remains: It only takes one error in the wrong place at the wrong time (or is it the right place at the right time?) to create a

catastrophic outcome. Thus, lest the occasional occurrence of error seem trivial, and beyond the need to examine from a systems approach, Leape (1994) quotes a personal communication from W. E. Deming, where even 99.9 % error free performance is not good enough:

If we had to live with 99.9 %, we would have: 2 unsafe plane landings per day at O'Hare, 16,000 pieces of lost mail every hour, and 32,000 bank checks deducted from the wrong bank account every hour. (p. 1852)

Can error still be characterized as the external manifestation of inattentiveness, laziness, negligence, and carelessness? Given the presented evidence for the complex interaction of the effects of technology and its environment on the limitations of the human cognitive system, it hardly seems that the occurrence of error can be attributed to such “conventional wisdom” causes. By adopting a systems perspective on the commission of error, an approach that recognizes the need for altering the external environment to accommodate the abilities and limitations of the human cognitive system during task performance must come to replace the organizational approach that views error as an intrinsic trait of the individual that must be “motivated out” to ensure future error-free behavior.

REFERENCES

- Arvey, R. D., & Jones, A. P. (1985). The use of discipline in organizational settings: A framework for future research. Research in Organizational Behavior, 7, 367 - 408.
- Baars, B. J. (1992). The many uses of error: Twelve steps to a unified framework. In B. J. Baars (Ed.), Experimental slips and human error: Exploring the architecture of volition (pp. 3 - 34). New York: Plenum.
- Baddeley, A. D., & Logie, R. H. (1999). The multiple component model. In A. Miyake and P. Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control (pp. 28 - 61). Cambridge: Cambridge University Press.
- Byrne, M. D. (1998). Taking a computational approach to aging: The SPAN theory of working memory. Psychology and Aging, 13 (2), 309 - 322.
- Byrne, M. D., & Bovair, S. (1997). A working memory model of a common procedural error. Cognitive Science, 21 (1), 31 - 61.
- Day, P. (2000a). The usual suspects (P. Day, Director). In S. Klebel (Producer), Why doctors make mistakes. TLC.
- Day, P. (2000b). Why doctors make mistakes (P. Day, Director). In S. Klebel (Producer), Why doctors make mistakes. TLC.
- Edmondson, A. C. (1996). Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error. Journal of Applied Behavioral Science, 32 (1), 5 - 28.

- Fitts, P. M, & Posner, M. I. (1967). Human error. Belmont, CA: Brooks/Cole Publishing Company.
- Gray, W. D. (2000). The nature and processing of errors in interactive behavior. Cognitive Science, 24 (2), 205 - 248.
- Hobgood, C. D., Ma, O. J., & Swart, G. L. (2000). Emergency medicine resident error: Identification and educational utilization. Academic Emergency Medicine, 7 (11), 1317 - 1320.
- Institute for Safe Medication Practices (1999, July 28). Evidence builds: Lack of focus on human factors allows error-prone devices. ISMP Medication Safety Alert [Online]. Available: <http://www.ismp.org/MSAarticles/HumanFactors.html>
- Leape, L. L. (1994). Error in medicine. The Journal of the American Medical Association, 272 (23), 1851 - 1857.
- Rasmussen, J. (1987). The definition of human error and a taxonomy for technical system design. In J. Rasmussen, K. Duncan, & J. Leplat (Eds.), New Technology and Human Error (p. 53 - 62). Chichester: John Wiley & Sons.
- Reason, J. (1984). Lapses of attention in everyday life. In R. Parasuraman & D. R. Davies (Eds.), Varieties of attention (pp. 515 - 549). New York: Academic Press.
- Reason, J. (1987). Generic error modelling system (GEMS): A cognitive framework for locating common human error forms. In J. Rasmussen, K. Duncan, & J. Leplat (Eds.), New Technology and Human Error (p. 63 - 83). Chichester: John Wiley & Sons.
- Reason, J. (1990). Human Error. New York: Cambridge University Press.

- Reason, J. (1994). Foreword. In M. S. Bogner (Ed.), Human Error in Medicine (pp. vii - xv). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Senders, J. W., & Moray, N. P. (1991). Human Error: Cause, Prediction, and Reduction. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Serig, D. I. (1994). Radiopharmaceutical misadministrations: What's wrong? In M. S. Bogner (Ed.), Human Error in Medicine (pp. 179 - 195). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Van Cott, H. (1994). Human errors: Their causes and reduction. In M. S. Bogner (Ed.), Human Error in Medicine (pp. 53 - 65). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Wickens, C. D. (1992). Engineering psychology and human performance. HarperCollins Publishers.
- Young, D., & Lebrun, D. (2001, January 24). Blondie. Houston Chronicle, p. 15D.

APPENDIX A

OPERATIONS OFFICER QUALIFYING EXAMS: CADET MANUAL



Operations Officer Qualifying Exam: Technical Requirements

Cadet:

Congratulations on completing the Academy's General Studies program. You are now prepared to continue your studies at a more advanced level. By electing to major in Operations, you have selected a rigorous course of training leading to a possible command line commission. An Operations Officer provides an interface between the day-to-day normal activities of the crew on board ship and the command personnel. As an Operations Officer, you must be able to perform at all duty stations, including Ops, Conn, Tactical, and other auxiliary bridge stations. You are asked to take this qualifying exam to demonstrate your knowledge and aptitude for leadership before you are admitted as a candidate for Operations Management.

You will be tested on your performance at several bridge stations. As the safety of the ship and crew depend on your performance at these tasks, speed and accuracy are of the essence. Your performance will be scored at each station; you will be timed and assigned a score reflecting your accuracy at the assigned task. As only the top cadets of each class are admitted to candidacy for Operations, your scores will be compared to those of other cadets participating in qualifications. Should your performance fall below a certain criterion, you may be asked to discontinue your qualification effort and leave the testing center.

The following describes the responsibilities at each station:

Ops

The Operations Management Office (Ops) is responsible for the coordination and scheduling of resources and hardware to the various missions being performed aboard a starship. The Ops panel on the bridge displays a list of all current major shipboard activities and the status of major shipboard system components (including core power, phaser reserves, deflector strength, and shield strength).

This information is used to evaluate the current state of the ship's activities, so that priority decisions can be made regarding resource allocation.

Conn

The responsibility for actual piloting and navigation of a starship lies with the Flight Control Officer (Conn). Receiving instructions directly from the commanding officer, the Conn's duties include: navigational references and course plotting, supervision of automatic flight operations, manual flight operations, positive verification, and acting as Bridge Liaison to Main Engineering. While these functions are heavily automated, their importance to the safety of the ship and the missions at hand demands that an officer be assigned to oversee the Conn at all time.

Tactical

Defensive Systems Control and starship internal security are the duties of the Tactical Station (Tactical). Tactical security coverage ranges from low-level crew safety to full counter-intelligence measures against sabotage or terrorism. External security systems (defensive shields, phasers, photon torpedoes) are controlled from Tactical, as well as sensor arrays, probes, buoys, and tractor beam systems.

Transporter

While all starships have several Transporter Rooms, one bridge station is devoted to Transporter functions that may be used in emergency situations. Transporters are matter-energy conversion devices that take an object or being and transform it into a pattern of phased energy that can be transmitted as a complex trans-barrier signal through the first level of subspace to a set of desired coordinates. At the desired coordinates, it is reintegrated into its original structure.

This portion of the qualifying exam will test your technical knowledge and performance at the various bridge stations. Other portions of the exam will be administered separately and test your aptitude for organization, management, and communications skills. Good luck on your Operations Officer Qualifying Exam.

Admiral Michael D. Byrne
Commandant, Starfleet Academy

Admiral Elizabeth M. Serig
Commander, Operations Command

APPENDIX B

TACTICAL STATION TRAINING MANUAL:

POSTCOMPLETION VERSION

Star Fleet Operations Manual

Model MB-X15.1(pc) Phaser Control Bank

This manual describes how to operate the MB-X15.1 Star Fleet standard Phaser Control Bank, the primary weapon on current Star Fleet vessels. Understanding how this control system operates is critical in ensuring the security of Star Fleet vessels. Figure 1 is a picture of the MB-X15.1 Phaser Control Bank.

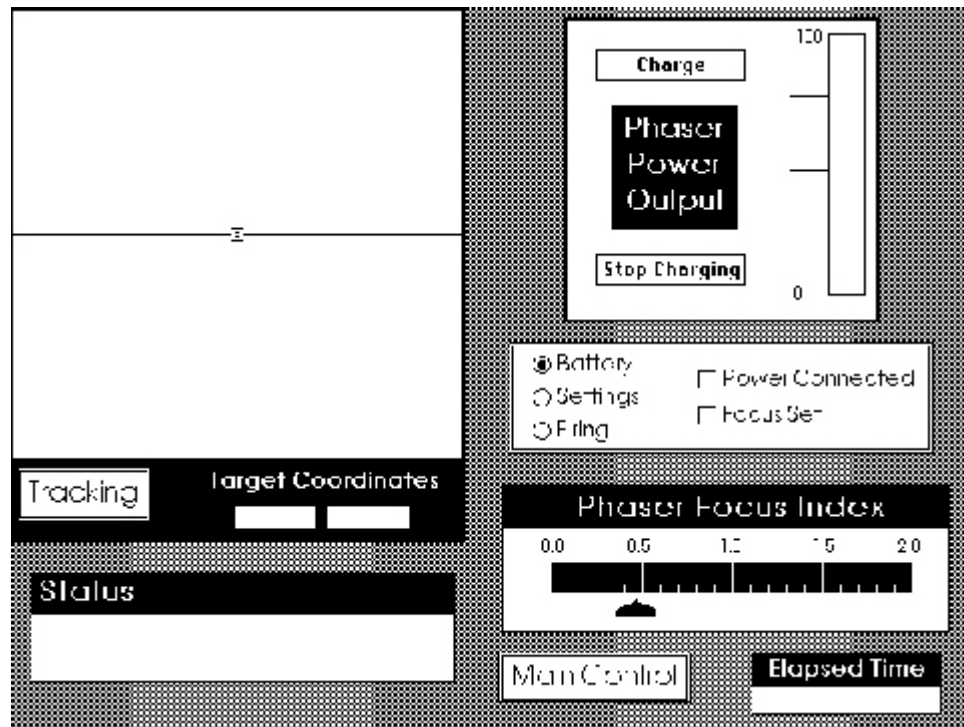


Figure 1. The MB-X15.1 Phaser Control Bank

There are four essential steps involved in firing the class X15 phaser:

- Charging the phaser
- Setting the focus of the phaser beam
- Tracking the target
- Firing the phaser.

Each step will be described in detail in the remainder of this manual.

Step 1. Charging the Phaser

Summary of steps to charge the phaser:

- Click "Power Connected"
- Click "Charge"
- Wait until phaser charges the appropriate amount
- Click "Stop Charging"
- Click "Power Connected"

The X15 class phaser requires more instantaneous power than can be generated by the standard power plant. This problem is solved by charging a virtual "battery" which yields the high instantaneous power output, but it must be re-charged after each firing of the phaser.

There are several steps in charging the battery. The first step is to connect the battery to the power source. This is done by clicking the "Power Connected" button on the control panel, as shown in Figures 2 and 3.

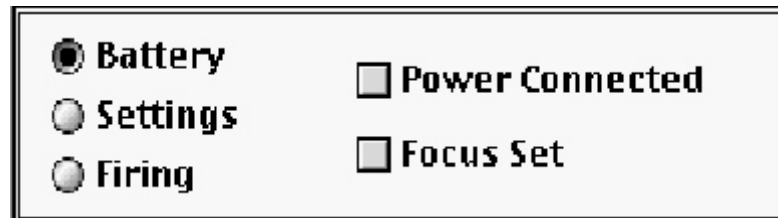


Figure 2. Before clicking "Power Connected"

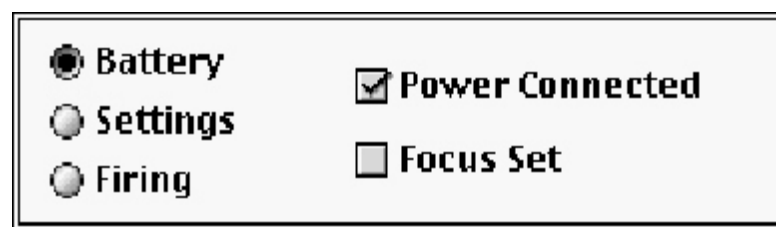


Figure 3. After clicking "Power Connected"

Once the power source is connected to the battery, it is possible to charge the phaser. This is done by clicking on the "Charge" button and waiting for the meter to indicate that the phaser has charged beyond the minimum and less than the maximum allowed safe value. These levels are marked on the meter with horizontal lines. When the meter level reaches the allowable range, the "Stop Charging" button should be clicked to halt the charging. Figures 4 - 6 demonstrate the process.

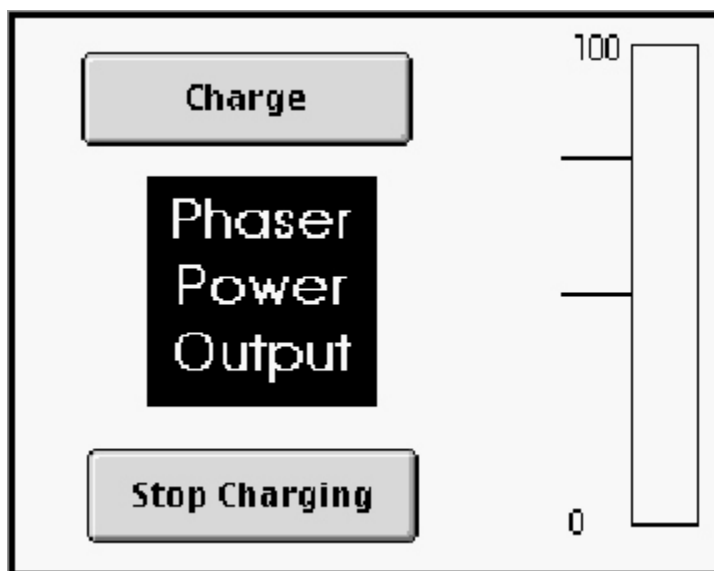


Figure 4. Before clicking "Charge"

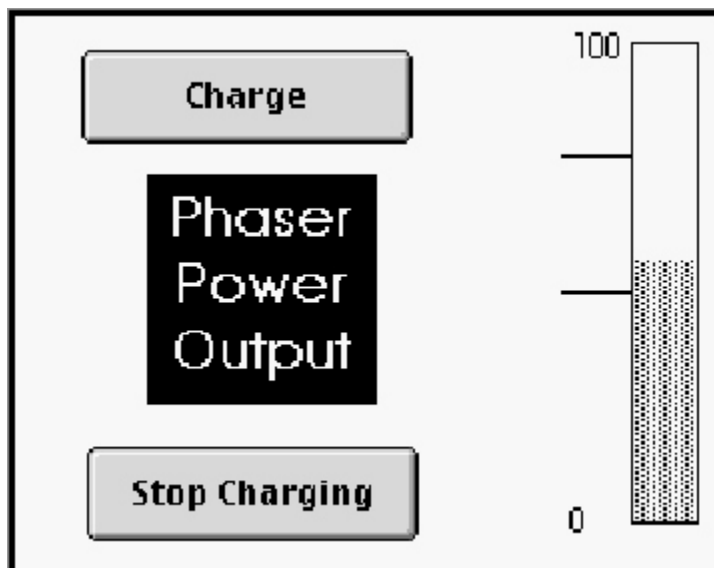


Figure 5. Phaser while charging

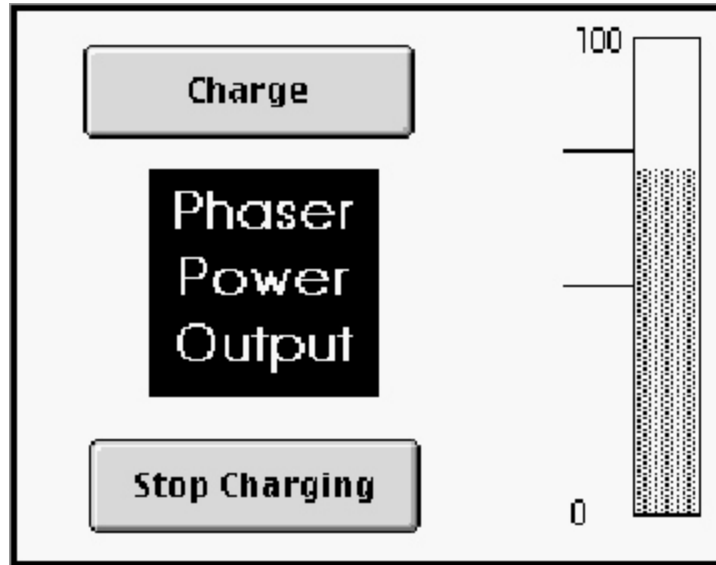


Figure 6. After clicking "Stop Charging"

It is critical that the phaser be charged within the allowable range (between the horizontal lines). Undercharging the phaser will make it unable to fire, and overcharging the phaser may cause a power feedback that will damage the phaser unit.

Once the phaser has charged, it is necessary to disconnect the battery from the power source. This is done by once again clicking the "Power Connected" button, which will cause the marking "X" to disappear. It is not possible to operate the other phaser controls unless the power has been disconnected.

Step. 2 Setting phaser beam focus

Summary of steps to set phaser beam focus:

- Click "Settings"
- Adjust location of slider to desired focus
- Click "Focus Set"

The opti-magnetic focusing apparatus of the X15 class phaser makes it possible to control the amount of dispersion of the phaser beam itself. The higher the dispersion, the larger is the perpendicular cross-section of the beam. Therefore, the higher the dispersion, the easier it is to hit the target. However, the more dispersed the beam, the less damaging it is. Dispersion is controlled by the Phaser Focus Index, which is set on the Phaser Control Bank.

The first step in setting the Focus Index is to enable the alteration of current settings. This is done by clicking on the "Settings" button, as is demonstrated in Figures 7 and 8.

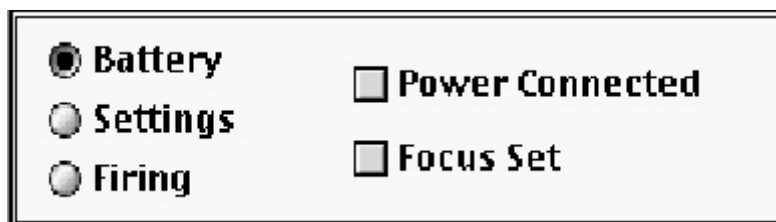


Figure 7. Before clicking "Settings"

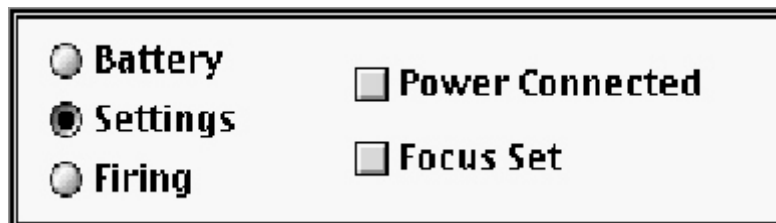


Figure 8. After clicking "Settings"

Once this is done, the Focus Index *must* be set by means of a slider-style control. The slider is adjusted by clicking on the indicator and dragging it to the desired setting. This is illustrated in Figures 9 and 10.

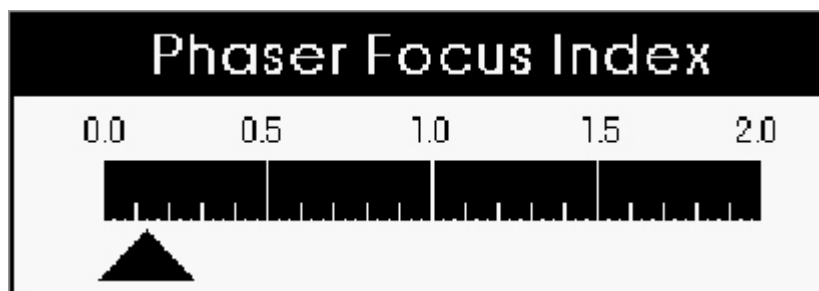


Figure 9. The mouse button is clicked down and held

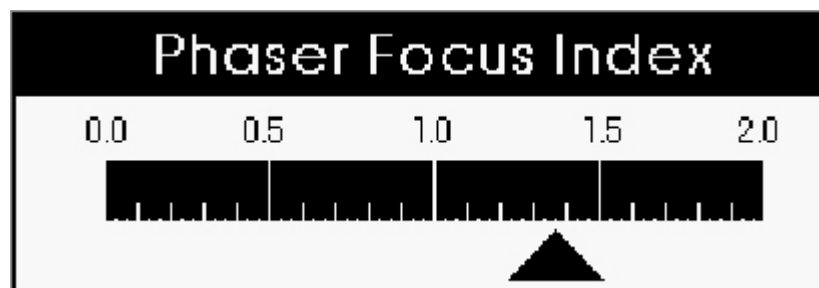


Figure 10. The mouse button is released at the desired index

Once the desired focus is set, it is necessary to inform the X15 firing system that the current focus is the desired one. That is, the focus setting must be "locked in." This is done by clicking the "Focus Set" button, as demonstrated in Figures 11 and 12.

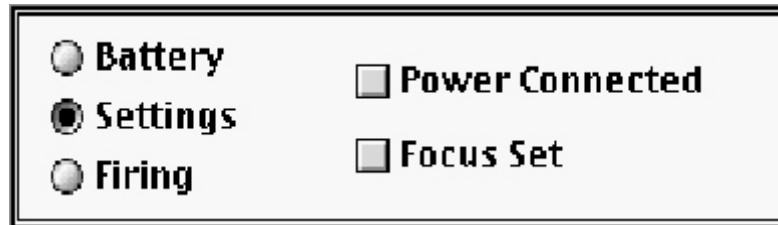


Figure 11. Before clicking "Focus Set"

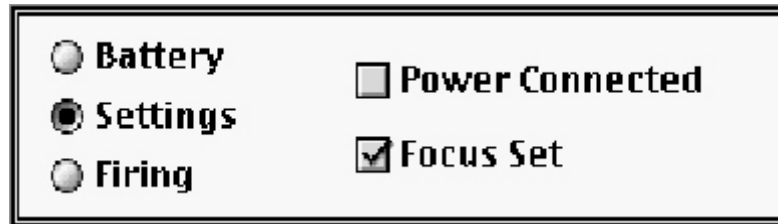


Figure 12. After clicking "Focus Set"

Once this has been completed, the Phaser Focus Index has been set. As with charging the phaser battery, it is impossible to operate the other controls until this has been done.

Step 3. Tracking the target

Summary of steps to track the target:

- Click "Firing"
- Click "Tracking"
- Use number keys to adjust location of the target indicator

The X15 class phaser system contains a sophisticated tracking system that has the capability to track a target up to near-relativistic speeds. However, there is a price that must be paid for this impressive performance. The tracking system can only get close to the target, but will almost never hit it automatically; there must be some operator intervention in the process. The nature of that intervention is described below.

The first step in tracking the target is to enable firing of the phaser, since it is somewhat senseless to track a target at which one cannot fire. Enabling firing is done by clicking the "Firing" button, as shown in Figures 13 and 14.

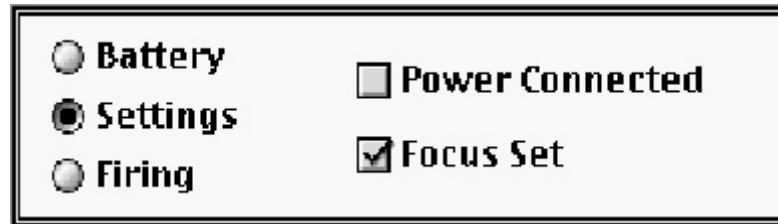


Figure 13. Before clicking "Firing"

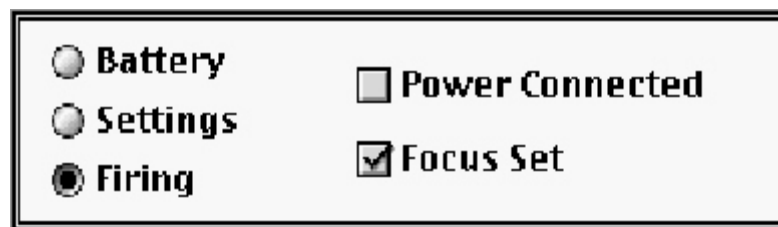


Figure 14. After clicking "Firing"

Once the system is ready to fire, the target can be tracked. The tracking system is turned on by clicking on the "Tracking" button, as shown in Figures 15 and 16. It is possible to tell that the tracking system is active by noting the presence of the target indicator (refer to the Figures).

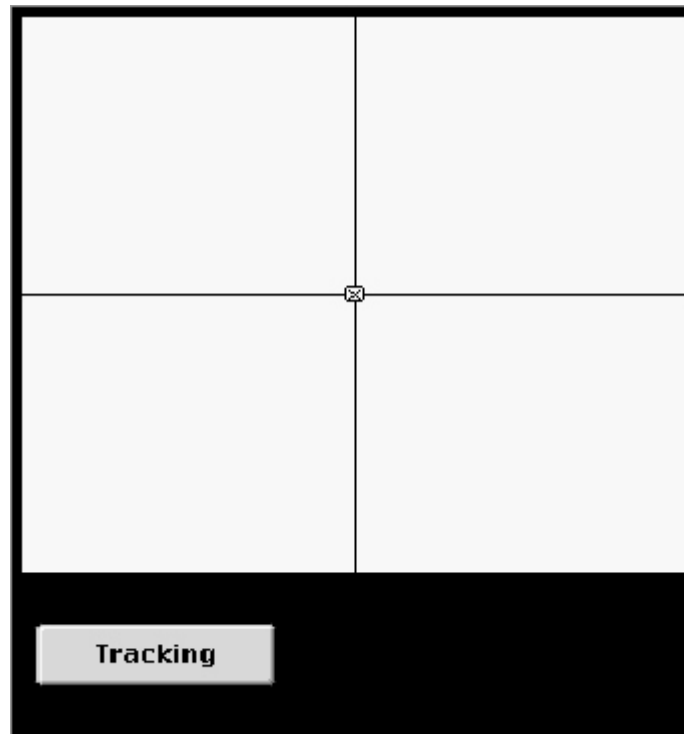


Figure 15. Before clicking "Tracking"

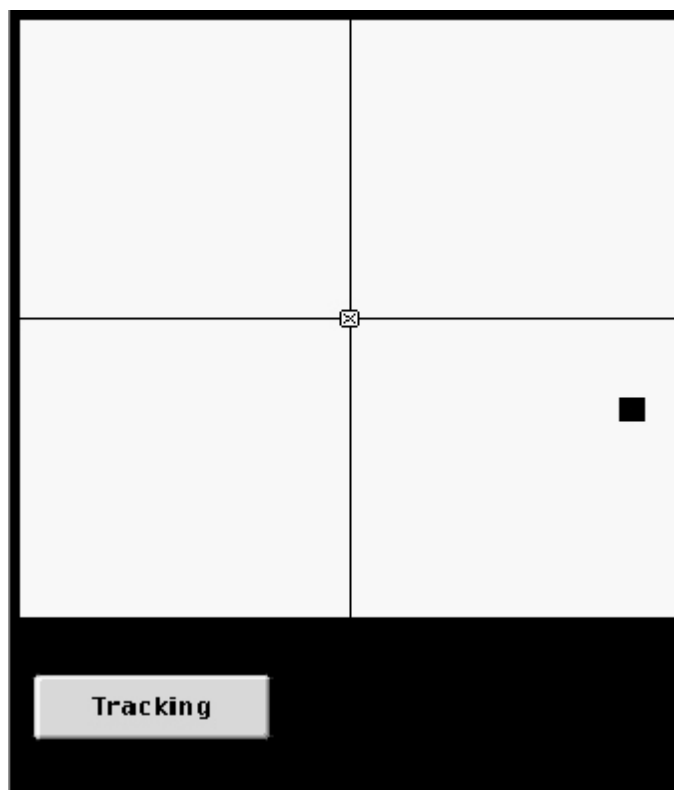


Figure 16. After clicking "Tracking"

Actual adjustments to tracking are made using the four keys on the numeric pad of the keyboard, as illustrated in Figure 17. These keys will adjust the tracking system in the direction indicated by the position of the key.



Figure 17. Keyboard with numeric keys highlighted

Pressing these keys has the effect of causing the target indicator to move in the opposite direction of the key that is pressed. Thus, to fire at the target in Figure 16, one would press the "2" key several times, and the "6" key several times. The tracking system does not operate as fast as you will be able to hit the numeric keys, so use caution. It can be

disturbing to switch from pressing one of the keys to pressing another and have no response to the latter key for a moment or two.

Use moderation in the rate of key presses. For the same reason, it is not advised to hold down the arrow keys, but rather to simply press them repeatedly.

Because of the difficulty in tracking high-speed objects such as threatening vessels, it is not guaranteed that the target will be hit, no matter how well the tracking is adjusted. The probability of hitting the target is a function of the distance of the target indicator from the center point of the tracking meter, and it is quite possible to hit the target when the indicator is merely in the vicinity of the center of the meter. Conversely, it is possible to miss the target even when it is dead in your sights.

Step 4. Firing the phaser

Summary of steps to fire the phaser:

- Press the space bar
- Determine if the target has been destroyed
- If so, click "Tracking"

Once the tracking has been adjusted and it is judged that the target is close enough to be fired upon, the phaser should be fired immediately. The X15 class phaser is fired by pressing the space bar on the keyboard.

Firing the phaser will have several effects. First, you will hear the sound of the power discharge. You will then notice that several things on the control panel have returned to their "rest" state. The charge will have dropped to zero, the "Focus Set" indicator will be off, and the toggle will have returned to "Battery." Figures 17 and 18 illustrate these changes.

The Status section of the panel will indicate the results of the firing. There are three possibilities:

- the phaser will miss the target
- the phaser will hit the target, but not destroy it
- the phaser will destroy the target

In either of the first two cases, it will be necessary to return to Step 1 (Charging the Phaser) in order to fire the phaser again. In the third case, the task is complete and you must turn off the tracking system, which will still be on. This can be done by clicking on the "Tracking" button again. You should only turn off the tracking if the target has been destroyed. Once you have done this, you may return to the main control screen by clicking on the "Main Control" button. This button should be obvious and will not be displayed in a figure.

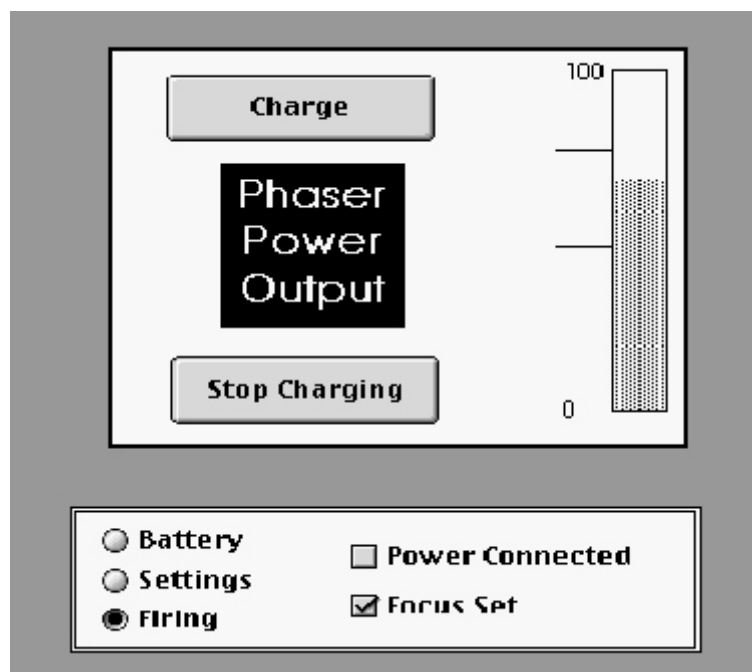


Figure 18. Panel section before firing phaser

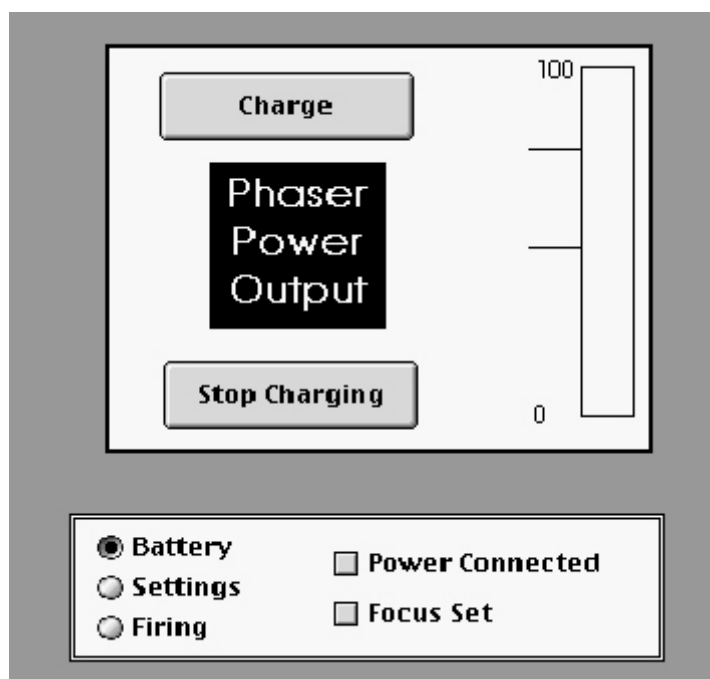


Figure 19. Panel section after firing phaser

Summary of Phaser Bank Operating Procedures:

Step 1. Charge the phaser:

- Click "Power Connected"
- Click "Charge"
- Wait until phaser charges the appropriate amount
- Click "Stop Charging"
- Click "Power Connected"

Step 2. Set phaser beam focus:

- Click "Settings"
- Adjust location of slider to desired focus
- Click "Focus Set"

Step 3. Track the target:

- Click "Firing"
- Click "Tracking"
- Use arrow keys to adjust location of the target indicator

Step 4. Fire the phaser:

- Press the space bar
- Determine if the target has been destroyed
- If so, click "Tracking"
- If not, return to Step 1

Then return to Main Control.

APPENDIX C
TACTICAL STATION TRAINING MANUAL:
CONTROL VERSION

Star Fleet Operations Manual

Model MB-X15.2(n) Phaser Control Bank

This manual describes how to operate the MB-X15.2 Star Fleet standard Phaser Control Bank, the primary weapon on current Star Fleet vessels. Understanding how this control system operates is critical in ensuring the security of Star Fleet vessels. Figure 1 is a picture of the MB-X15.2 Phaser Control Bank.

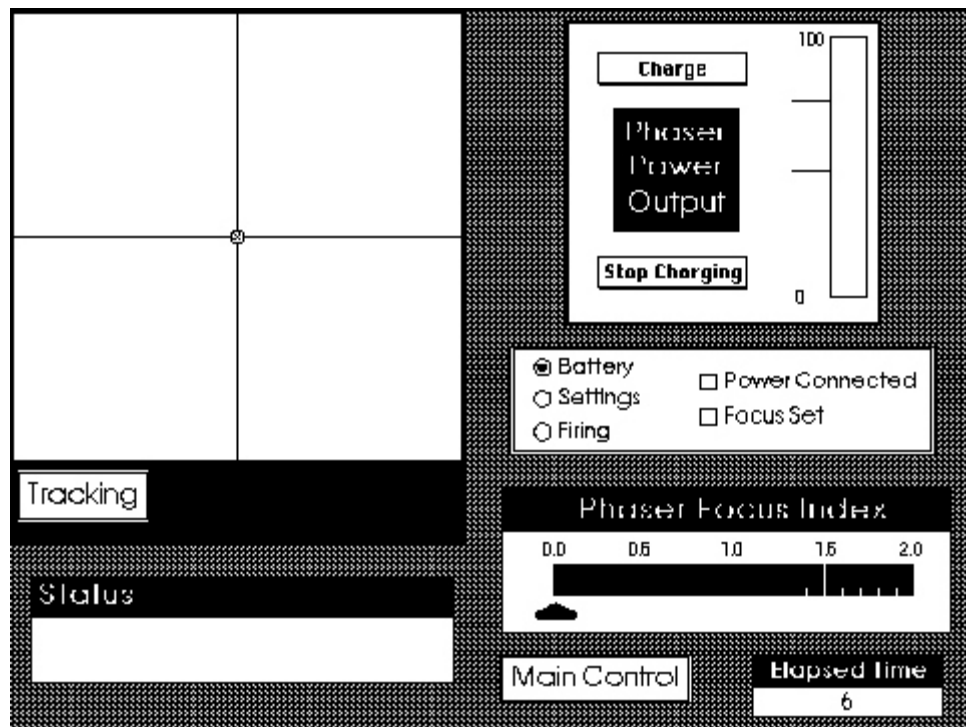


Figure 1. The MB-X15.2 Phaser Control Bank

There are four essential steps involved in firing the class X15 phaser:

- Charging the phaser
- Setting the focus of the phaser beam
- Tracking the target
- Firing the phaser.

Each step will be described in detail in the remainder of this manual.

Step 1. Charging the Phaser

Summary of steps to charge the phaser:

- Click "Power Connected"
- Click "Charge"
- Wait until phaser charges the appropriate amount
- Click "Stop Charging"
- Click "Power Connected"

The X15 class phaser requires more instantaneous power than can be generated by the standard power plant. This problem is solved by charging a virtual "battery" which yields the high instantaneous power output, but it must be re-charged after each firing of the phaser.

There are several steps in charging the battery. The first step is to connect the battery to the power source. This is done by clicking the "Power Connected" button on the control panel, as shown in Figures 2 and 3.

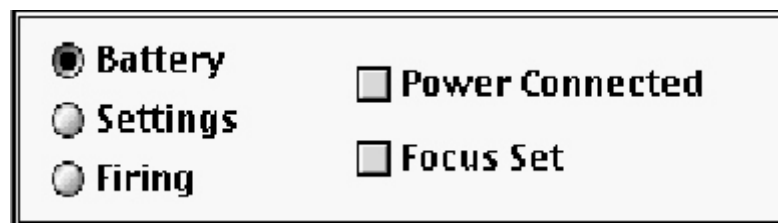


Figure 2. Before clicking "Power Connected"

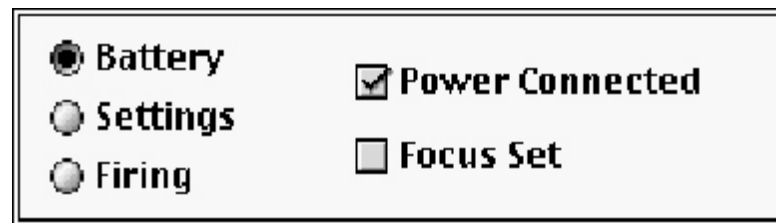


Figure 3. After clicking "Power Connected"

Once the power source is connected to the battery, it is possible to charge the phaser. This is done by clicking on the "Charge" button and waiting for the meter to indicate that the phaser has charged beyond the minimum and less than the maximum allowed safe value. These levels are marked on the meter with horizontal lines. When the meter level reaches the allowable range, the "Stop Charging" button should be clicked to halt the charging. Figures 4 - 6 demonstrate the process.

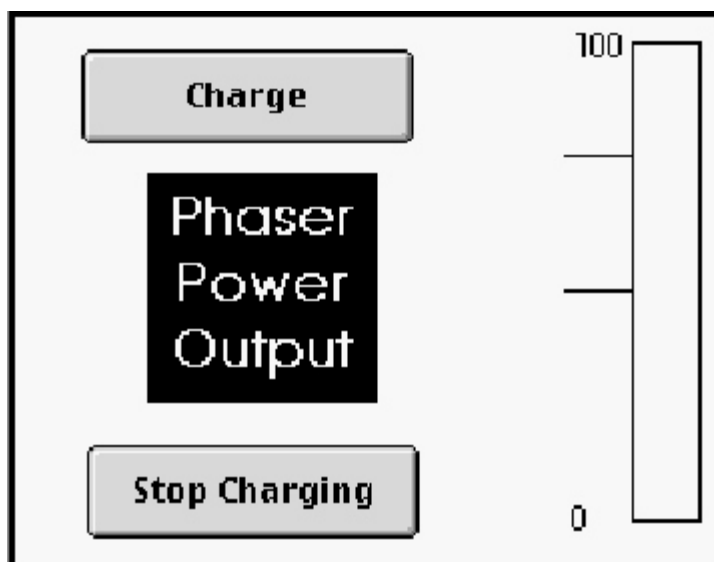


Figure 4. Before clicking "Charge"

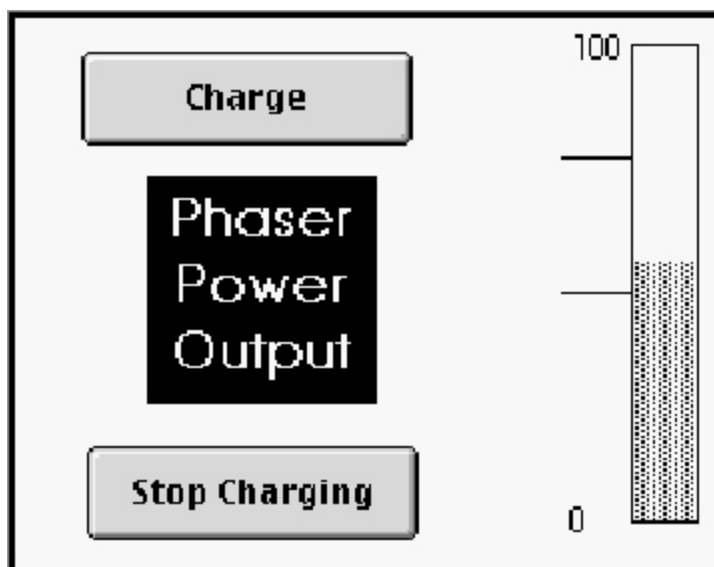


Figure 5. Phaser while charging

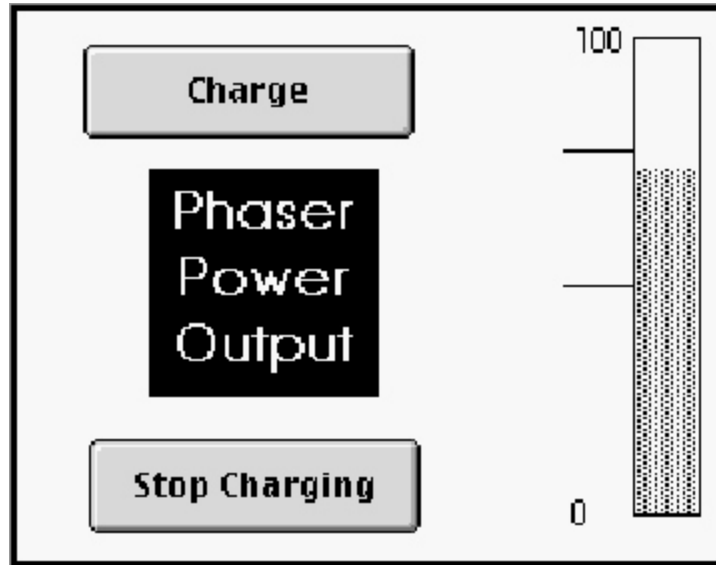


Figure 6. After clicking "Stop Charging"

It is critical that the phaser be charged within the allowable range (between the horizontal lines). Undercharging the phaser will make it unable to fire, and overcharging the phaser may cause a power feedback that will damage the phaser unit.

Once the phaser has charged, it is necessary to disconnect the battery from the power source. This is done by once again clicking the "Power Connected" button, which will cause the marking "X" to disappear. It is not possible to operate the other phaser controls unless the power has been disconnected.

Step. 2 Setting phaser beam focus

Summary of steps to set phaser beam focus:

- Click "Settings"
- Adjust location of slider to desired focus
- Click "Focus Set"

The opti-magnetic focusing apparatus of the X15 class phaser makes it possible to control the amount of dispersion of the phaser beam itself. The higher the dispersion, the larger is the perpendicular cross-section of the beam. Therefore, the higher the dispersion, the easier it is to hit the target. However, the more dispersed the beam, the less damaging it is. Dispersion is controlled by the Phaser Focus Index, which is set on the Phaser Control Bank.

The first step in setting the Focus Index is to enable the alteration of current settings. This is done by clicking on the "Settings" button, as is demonstrated in Figures 7 and 8.

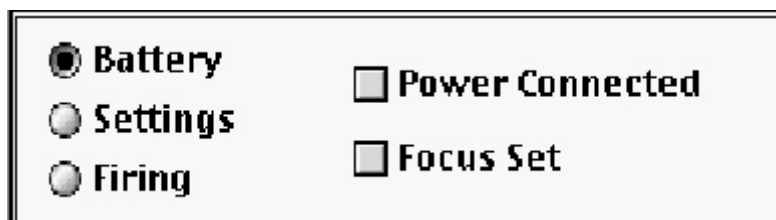


Figure 7. Before clicking "Settings"

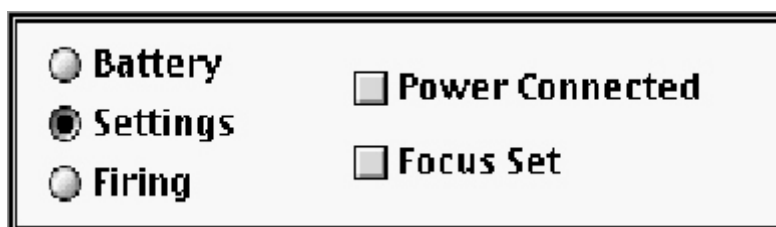


Figure 8. After clicking "Settings"

Once this is done, the Focus Index *must* be set by means of a slider-style control. The slider is adjusted by clicking on the indicator and dragging it to the desired setting. This is illustrated in Figures 9 and 10.

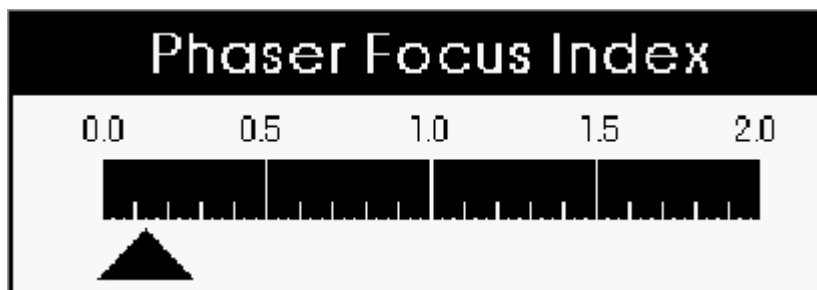


Figure 9. The mouse button is clicked down and held

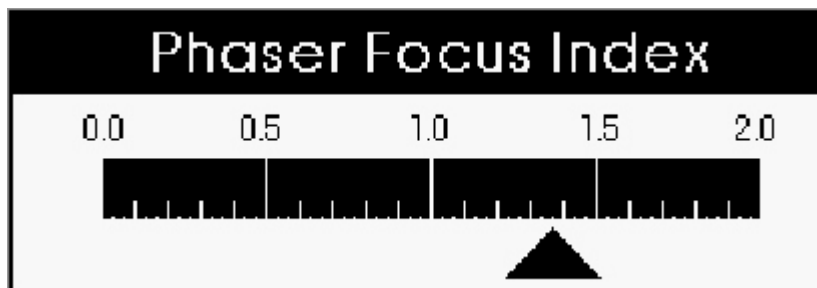


Figure 10. The mouse button is released at the desired index

Once the desired focus is set, it is necessary to inform the X15 firing system that the current focus is the desired one. That is, the focus setting must be "locked in." This is done by clicking the "Focus Set" button, as demonstrated in Figures 11 and 12.

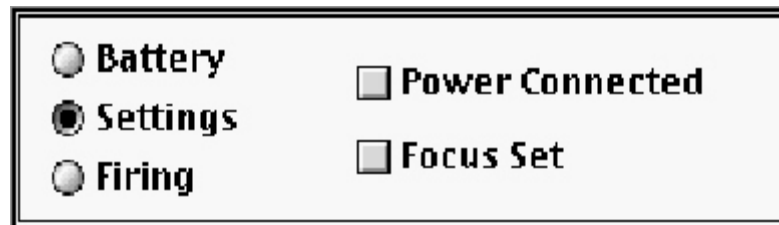


Figure 11. Before clicking "Focus Set"

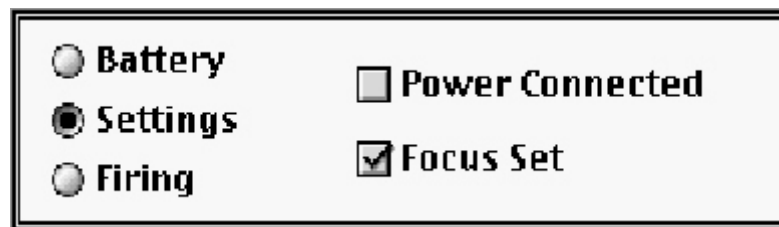


Figure 12. After clicking "Focus Set"

Once this has been completed, the Phaser Focus Index has been set. As with charging the phaser battery, it is impossible to operate the other controls until this has been done.

Step 3. Tracking the target

Summary of steps to track the target:

- Click "Firing"
- Click "Tracking"
- Use number keys to adjust location of the target indicator

The X15 class phaser system contains a sophisticated tracking system that has the capability to track a target up to near-relativistic speeds. However, there is a price that must be paid for this impressive performance. The tracking system can only get close to the target, but will almost never hit it automatically; there must be some operator intervention in the process. The nature of that intervention is described below.

The first step in tracking the target is to enable firing of the phaser, since it is somewhat senseless to track a target at which one cannot fire. Enabling firing is done by clicking the "Firing" button, as shown in Figures 13 and 14.

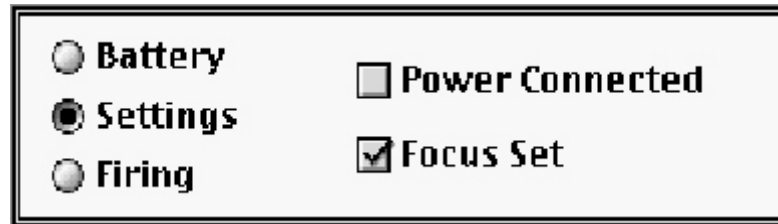


Figure 13. Before clicking "Firing"

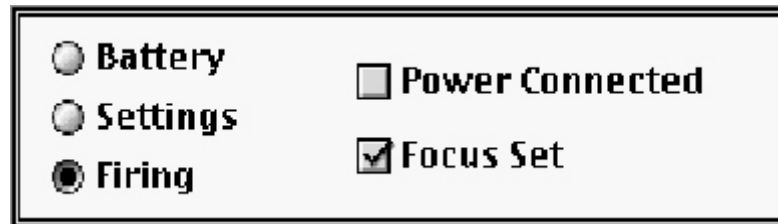


Figure 14. After clicking "Firing"

Once the system is ready to fire, the target can be tracked. The tracking system is turned on by clicking on the "Tracking" button, as shown in Figures 15 and 16. It is possible to tell that the tracking system is active by noting the presence of the target indicator (refer to the Figures).

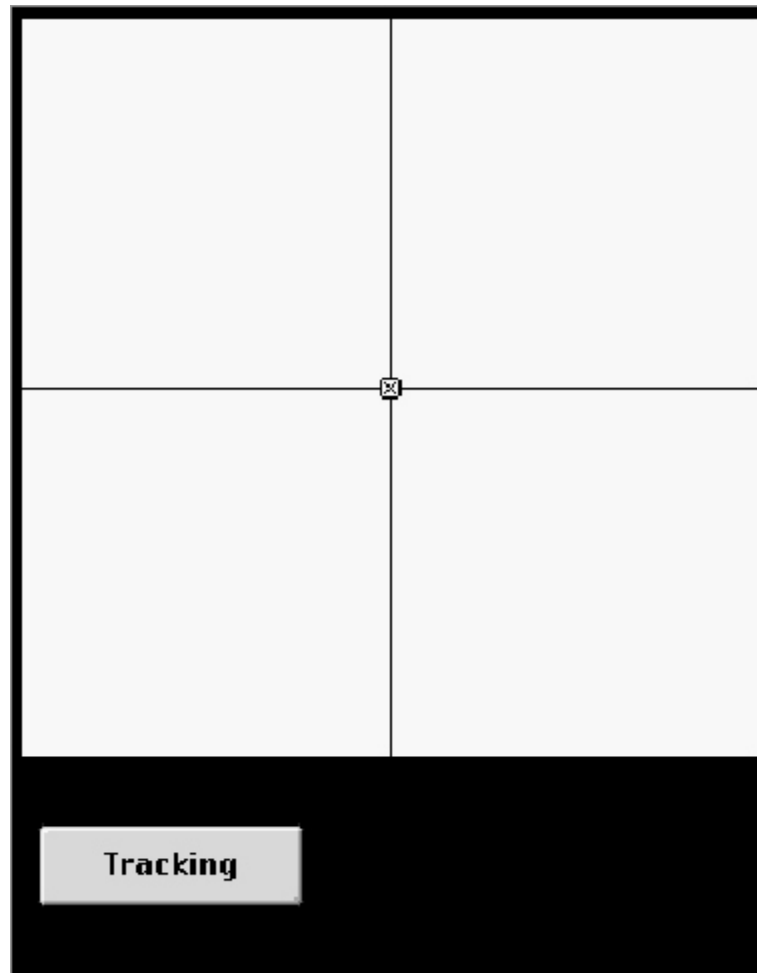


Figure 15. Before clicking "Tracking"

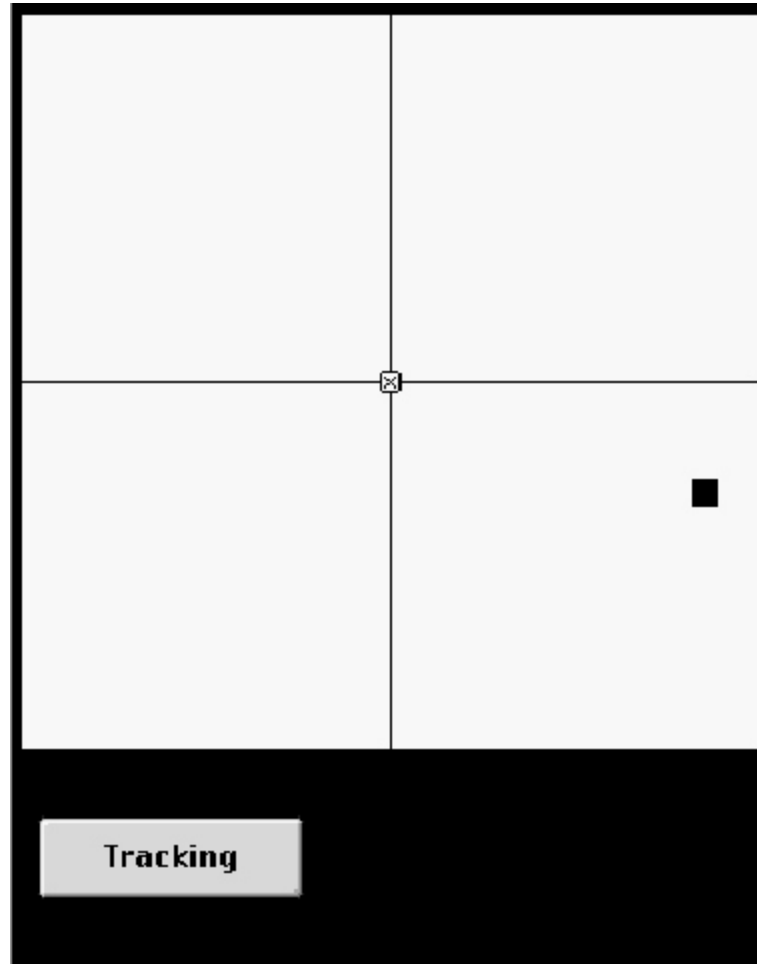


Figure 16. After clicking "Tracking"

Actual adjustments to tracking are made using the four keys on the numeric pad of the keyboard, as illustrated in Figure 17. These keys will adjust the tracking system in the direction indicated by the position of the key.

Figure 17. Keyboard with numeric keys highlighted



Pressing these keys has the effect of causing the target indicator to move in the opposite direction of the key that is pressed. Thus, to fire at the target in Figure 16, one would press the "2" key several times, and the "6" key several times. The tracking system does not operate as fast as you will be able to hit the numeric keys, so use caution. It can be disturbing to switch from pressing one of the keys to pressing another and have no response to the latter key for a moment or two. Use moderation in the rate of key presses. For the same reason, it is not advised to hold down the arrow keys, but rather to simply press them repeatedly.

Because of the difficulty in tracking high-speed objects such as threatening vessels, it is not guaranteed that the target will be hit, no matter how well the tracking is adjusted. The probability of hitting the target is a function of the distance of the target indicator from the center point of the tracking meter, and it is quite possible to hit the target when the indicator is merely in the vicinity of the center of the meter. Conversely, it is possible to miss the target even when it is dead in your sights

Step 4. Firing the phaser

Summary of steps to fire the phaser:

- Press the space bar
- Click "Tracking"
- Determine if the target has been destroyed

Once the tracking has been adjusted and it is judged that the target is close enough to be fired upon, the phaser should be fired immediately. The X15 class phaser is fired by pressing the space bar on the keyboard.

Firing the phaser will have several effects. First, you will hear the sound of the power discharge. You will then notice that several things on the control panel have returned to their "rest" state. The charge will have dropped to zero, the "Focus Set" indicator will be off, and the toggle will have returned to "Battery." Figures 17 and 18 illustrate these changes.

In order to determine the outcome of the phaser shot, it is necessary to re-engage the tracking system (actions other than tacking turn the system off). This is done by clicking on the "Tracking" button. The Status section of the panel will then indicate the results of the firing. There are three possibilities:

- the phaser will miss the target
- the phaser will hit the target, but not destroy it
- the phaser will destroy the target

In either of the first two cases, it will be necessary to return to Step 1 (Charging the Phaser) in order to fire the phaser again. In the third case, the task is complete and you

should return to the main control screen by clicking on the "Main Control" button. This button should be obvious and will not be displayed in a figure.

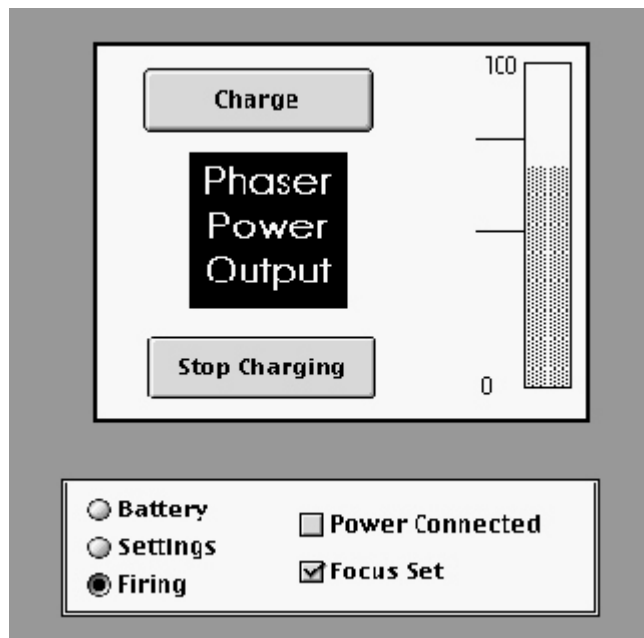


Figure 18. Panel section before firing phaser

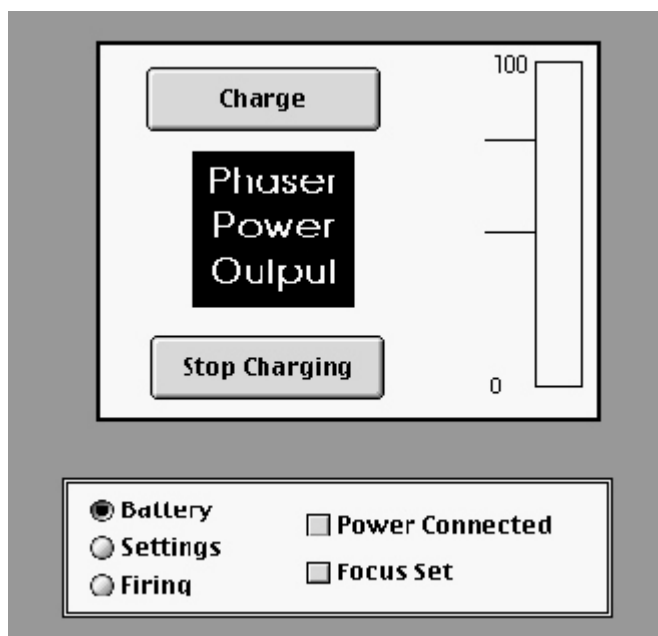


Figure 19. Panel section after firing phaser

Summary of Phaser Bank Operating Procedures:

Step 1. Charge the phaser:

- Click "Power Connected"
- Click "Charge"
- Wait until phaser charges the appropriate amount
- Click "Stop Charging"
- Click "Power Connected"

Step 2. Set phaser beam focus:

- Click "Settings"
- Adjust location of slider to desired focus
- Click "Focus Set"

Step 3. Track the target:

- Click "Firing"
- Click "Tracking"
- Use arrow keys to adjust location of the target indicator

Step 4. Fire the phaser:

- Press the space bar
- Click "Tracking"
- Determine if the target has been destroyed
- If not, return to Step 1

Then return to Main Control.

APPENDIX D
TRANSPORTER TASK TRAINING MANUAL

Star Fleet Operations Manual

Model MB-X15.1(n) Manual Transporter System (MTS)

This manual describes how to operate the MB-X15.1 Star Fleet standard Manual Transporter System (MTS), the primary method of bringing aboard crewmembers in hostile circumstances when automatic transporters are being jammed. Figure 1 is a picture of the MB-X15.1 Manual Transporter System.

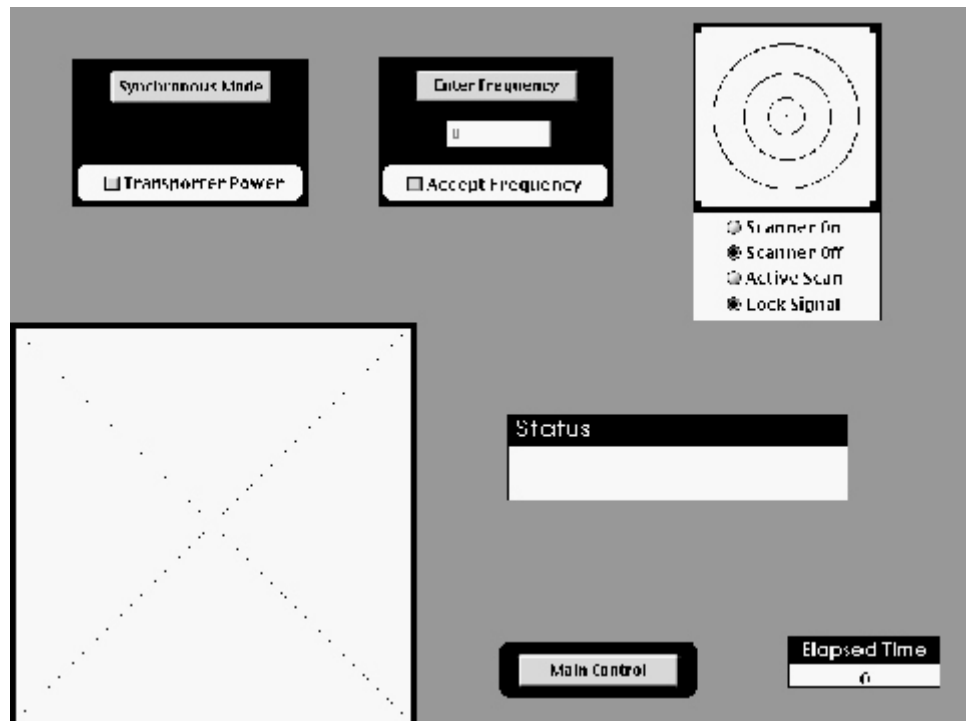


Figure 1. The MB-X15.1 Manual Transporter System

There are four essential steps involved in operating the MTS:

- Lock on to the homing signal
- Setting the jamming frequency
- Synchronizing the transporter and homing signal
- Energizing the transporter.

Each step will be described in detail in the remainder of this manual.

Step 1. Lock on to the homing signal

Summary of steps to lock on to the homing signal:

- Click "Scanner On"
- Click "Active Scan"
- Wait until scanner homes in on remaining signal
- Click "Lock Signal"
- Click "Scanner Off"

When crewmembers need to be beamed aboard in hazardous situations, one or more of the crewmembers will use their communicator to send out a signal broadcasting their location. It is necessary to scan for this signal and then lock the MTS onto it.

There are several steps involved in locking onto the homing signal. The first step is to turn on the scanner. This is done by clicking on the "Scanner On" button, as depicted in Figures 2 and 3.

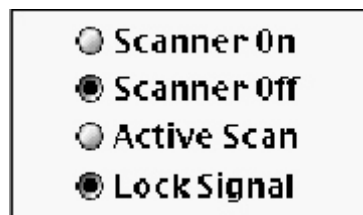


Figure 2. Before clicking "Scanner On"

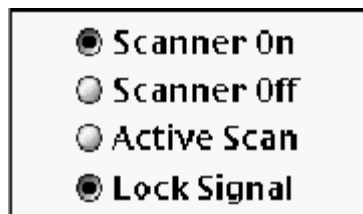


Figure 3. After clicking "Scanner On"

Once the scanner has been turned on, it is possible to activate the scanning system and track the signal. This is done by clicking on the "Active Scan" button and waiting until the scanner has found a unique signal and has phase-aligned with that signal. When the "Active Scan" button is pressed, the scanner will find a number of signals, gradually eliminate the false signals, and home in on the last remaining signal. When the signal indicator for the unique signal is inside the inner phase circle, the "Lock Signal" button should be pressed to lock on that homing signal. This is illustrated in Figures 4 through 6.

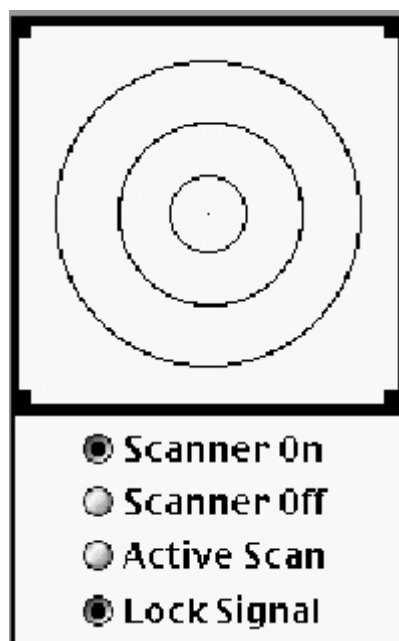


Figure 4. Before clicking "Active Scan"

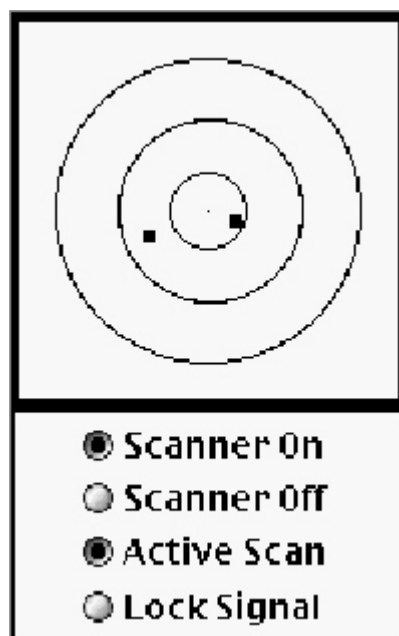


Figure 5. Display while scanning

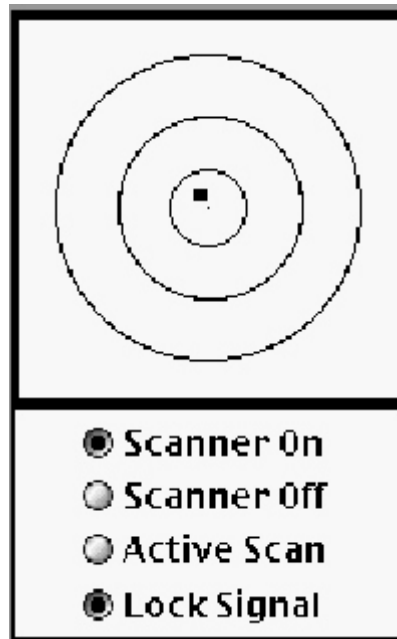


Figure 6. After clicking "Lock Signal"

It is critical that there be only one signal active in the scanner, because if not, something other than the crewmembers may be beamed aboard. It is equally important that the last signal be in-phase (the indicator inside the small circle) so that the transporter has an accurate signal and therefore a reasonable chance to beam successfully.

Once the homing signal has been locked in, it is necessary to turn off the signal scanner. This is done by clicking on the "Scanner Off" button, much like clicking the "Scanner On" button.

Step. 2 Setting the jamming frequency

Summary of steps to set the jamming frequency:

- Click "Enter Frequency"
- Type in the desired scanner frequency
- Click "Accept Frequency"

When operating the MTS system, one has to decide what jamming frequency to use. Jamming frequency is the operator's best guess as to the frequency of the signal being used by the hostile forces in jamming the transporter beam. Note that it may be possible to successfully beam crewmembers aboard even if the jamming frequency selected is incorrect. Conversely, it is possible to fail even when the jamming frequency is correct.

The jamming frequency entered on the MTS display may range from 1-100. Higher frequencies increase the probability of successfully jamming the enemy signal and thereby getting an accurate read on the crewmembers' location.

However, a higher frequency lowers the strength of the transporter beam, thus making it less likely to actually bring the crew aboard once the beam has been activated. It is the operator's responsibility to decide how to trade these things off against each other.

The first step in setting the jamming frequency is enabling the keyboard entry of a signal value. This is done by clicking on the "Enter Frequency" button. When this button has been clicked, a blinking cursor will appear in the frequency field, as show in Figures 7 and 8.

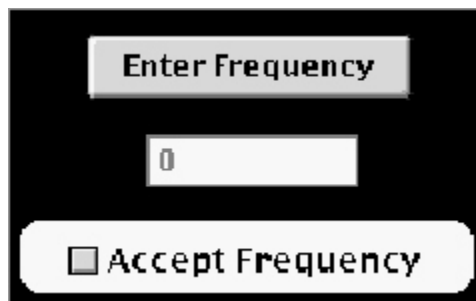


Figure 7. Before clicking "Enter Frequency"

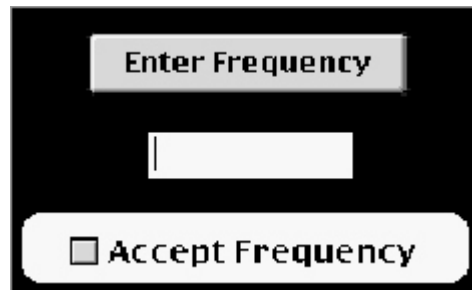


Figure 8. After clicking "Enter Frequency"

After text entry is enabled, the frequency value should be entered on the keyboard. This number should be between 1 and 100. Once this number is entered, it is necessary to tell the system to commit to that frequency. This is done by clicking on the "Accept Frequency" button, as shown in Figures 9 and 10.

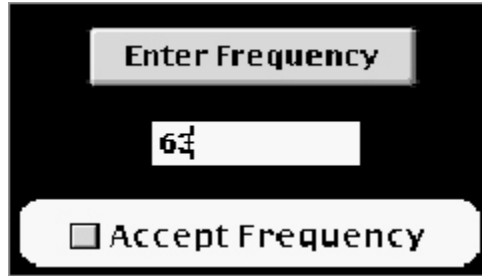


Figure 9. Before clicking "Accept Frequency"

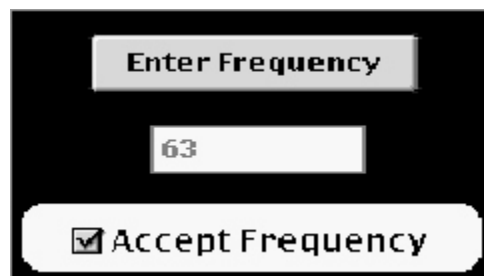


Figure 10. After clicking

"Accept Frequency"

Step 3. Synchronizing the transporter and homing signal

Summary of steps to track the target:

- Click "Transporter Power"
- Click "Synchronous Mode"
- Use the mouse to track the homing signal

The X15 class transporter system must be manually synchronized with the homing signal. This is done by turning on the transporter system itself, telling the system to allow operator synchronous tracking, and then manually tracking the signal. Turning on the main transporter power is done by clicking on the "Transporter Power" button, as shown in Figures 11 and 12.

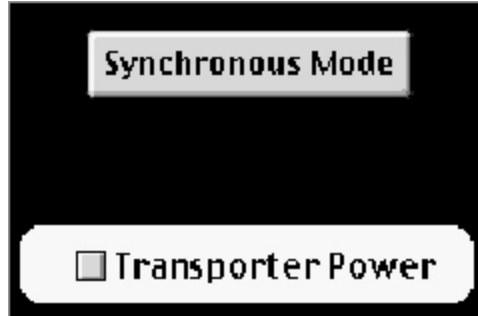


Figure 11. Before clicking "Transporter Power"

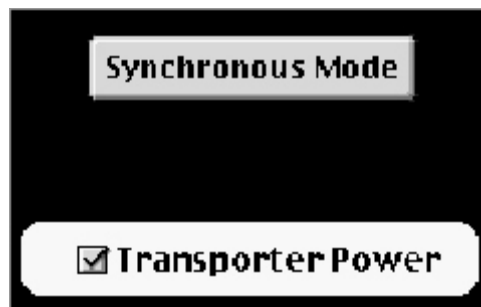


Figure 12. After clicking "Transporter Power"

Second, it is necessary to switch the system into synchronous tracking mode. This is done by clicking on the "Synchronous Mode" button, which is also displayed in Figures 11 and 12. When this button has been clicked, the cursor will turn into a targeting circle with cross hairs, and the square target signal will appear in the tracking area of the screen. This is illustrated in Figure 13, with the cursor in the upper portion of the figure, and the target in the lower right.

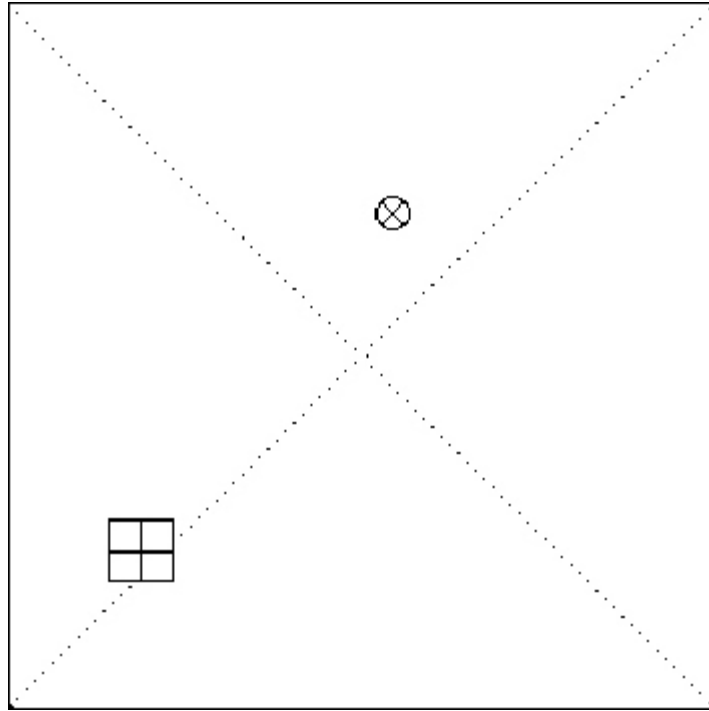


Figure 13. Tracking area after clicking "Synchronous Mode"

The target will be in constant motion, reflecting the instability of the homing signal. The transporter and the homing signal are considered synchronized to the degree that the target and the cursor overlap, so the operator of the MTS system should move the mouse as close as possible to the target before energizing the transporter.

Step 4. Energizing the transporter

Summary of steps to energize the transporter:

- Click the mouse button
- Click "Synchronous Mode"
- Determine if the beam has been successful

When the MTS is in synchronous mode, the transporter will energize when the mouse button is clicked. This should be done as soon as the cursor is close to the target. Energizing the transporter will rest MTS display. Next, it is necessary to switch off synchronous mode and find out the outcome of the beam attempt. This is done by clicking on the "Synchronous Mode" button. The "Status" box will then indicate the outcome of the transportation attempt. There are only two outcomes here: either the beam was a success, or it was not.

If the beam was unsuccessful, it is necessary to return to Step 1 (Lock on to the homing signal.) to try again. If the beam was successful, you may return to the main control screen by clicking on the "Main Control" button. This button should be obvious and will not be displayed in a figure.

Summary of Manual Transporter System Operating Procedures:

Step 1. Lock on to the homing signal

- Click "Scanner On"
- Click "Active Scan"
- Wait until scanner homes in on remaining signal
- Click "Lock Signal"
- Click "Scanner Off"

Step 2. Setting the jamming frequency:

- Click "Enter Frequency"
- Type in the desired scanner frequency
- Click "Accept Frequency"

Step 3. Synchronize the transporter and homing signal:

- Click "Transporter Power"
- Click "Synchronous Mode"
- Use the mouse to track the homing signal

Step 4. Energizing the transporter:

- Click the mouse button
- Click "Synchronous Mode"
- Determine if the beam has been successful
- If not, return to Step 1

Return to Main Control.

APPENDIX E
CONN TASK TRAINING MANUAL

Star Fleet Operations Manual

Model RD-X15.1 Navigation Console

This manual describes how to operate the RD-X15.1 Star Fleet standard Navigation Console. Understanding how this console system operates is critical to ensuring safe travel through space to an intended destination. Figure 1 is a picture of the RD-X15.1 Navigation Console.

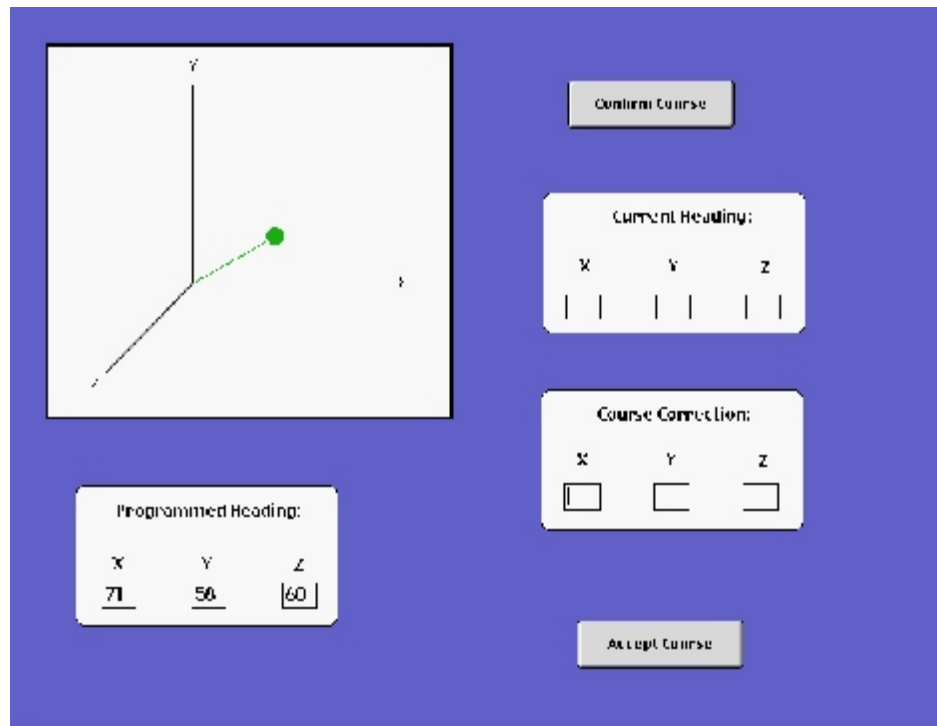


Figure 1. The RD-X15.1 Navigation Console

There are three essential steps involved in taking readings from the RD-X class navigation console:

- Determine ship's course heading relative to programmed heading
- Compute course difference if current heading deviates from programmed heading
- Enter course correction into navigation system

Each step will be described in detail in the remainder of this manual.

Step 1. Determine Ship's Course Heading

Summary of steps to determine ship's current course heading relative to programmed heading:

Click "Confirm Course"

During space flight, variations in space (such as solar winds, nebulae, and stellar dust fields) can influence a starship's projected course, essentially knocking it off course. Thus, it is necessary to engage in frequent course correction.

The first step of correcting the ship's course, is to determine its current heading. This is done by clicking on the "Confirm Course" button on the control panel, as shown in Figure 2.



Figure 2. Click on "Confirm Course"

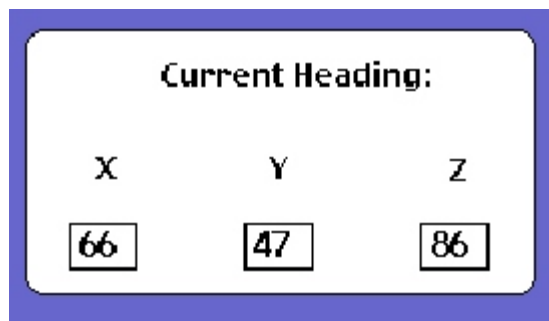
Step 2. Compute Course Difference from Programmed Heading

Summary of steps to compute course difference if current heading deviated from programmed heading:

Compare course heading in "Programmed Heading" with course heading in "Current Heading"

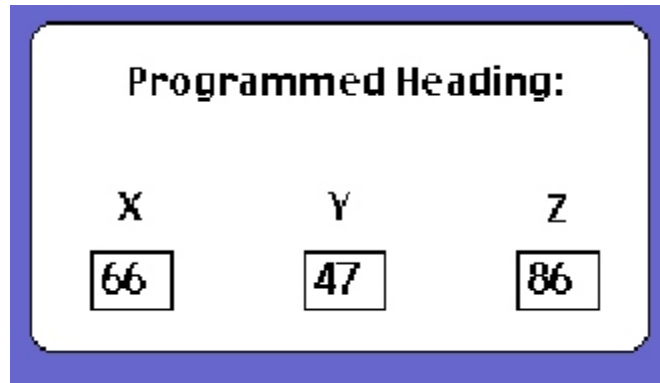
Enter course correction in "Course Correction" text boxes

If the course identified in the "Current Heading" matches the "Programmed Heading," enter 0 in the X, Y, and X fields by using the numeric keypad. This is illustrated in Figures 3 - 5.



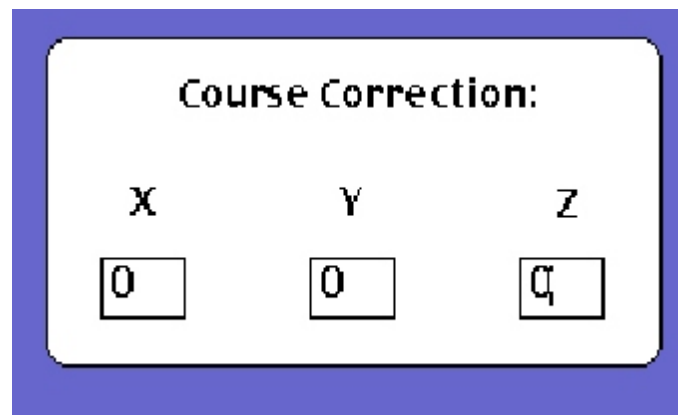
Current Heading:		
X	Y	Z
66	47	86

Figure 3. Current Heading



X	Y	Z
66	47	86

Figure 4. Programmed Heading



X	Y	Z
0	0	0

Figure 5. Course Correction

If the course identified in the "Current Heading" does NOT match the "Programmed Heading," you must compute the difference between the intended (programmed) course and the actual (current) course. It is necessary to compute the difference between these courses to enter the course correction. To compute the difference, subtract the "Current Heading" values from the "Programmed Heading" values. For example, if the Programmed Heading is [40, 25, 64], and the Current Heading is [60, 20, 60], you would calculate the course correction as follows:

Axis: Programmed - Current

X: $40 - 60 = -20$

Y: $25 - 20 = 5$

Z: $64 - 60 = 4$

You would enter [-20, 5, 4] in the "Course Correction" area. Note that if the Programmed Heading for a value is less than the Current Heading, the Course Correction value will be

negative (you need to move "down" in space to reach the correct heading). Figures 6-8 illustrate this process.

Current Heading:		
X	Y	Z
68	59	93

Figure 6. Current Heading

Programmed Heading:		
X	Y	Z
45	49	40

Figure 7. Programmed Heading

Course Correction:		
X	Y	Z
-23	-10	-53

Figure 8. Course Correction

Note: The Current Heading + the Course Correction should equal the Programmed Heading.

Step 3. Enter course correction into navigation system

Summary of steps to enter course correction into navigation system

Click on "Accept Course"

In order to accept the course correction into the navigation computer, click the "Accept Course" button, as shown in Figure 9.



Figure 9. Click on "Accept Course"

Once this is done, the entered correction is entered into the ship's log and the navigation computer makes the correction to the ship's course.

Summary of Navigation Console Operating Procedures:**Step 1. Determine Ship's Course Heading**

- Click "Confirm Course"

Step 2. Compute Course Difference from Programmed Heading

- Compare course heading in "Programmed Heading" with course heading in "Current Heading"
- Calculate difference between Programmed Heading and Current Heading by subtracting Current Heading from Programmed Heading (Correction = Programmed - Current)
- Enter course correction in "Course Correction" text boxes

Step 3. Enter course correction into navigation system

- Click "Accept Course"

APPENDIX F
OPERATIONS TASK TRAINING MANUAL

Star Fleet Operations Manual

Model ES-X15.1 Status Console

This manual describes how to operate the ES-X15.1 Star Fleet standard Status Console. Understanding how this console system operates is critical to ensuring safe operation of Star Fleet vessels. Figure 1 is a picture of the ES-X15.1 Status Console.

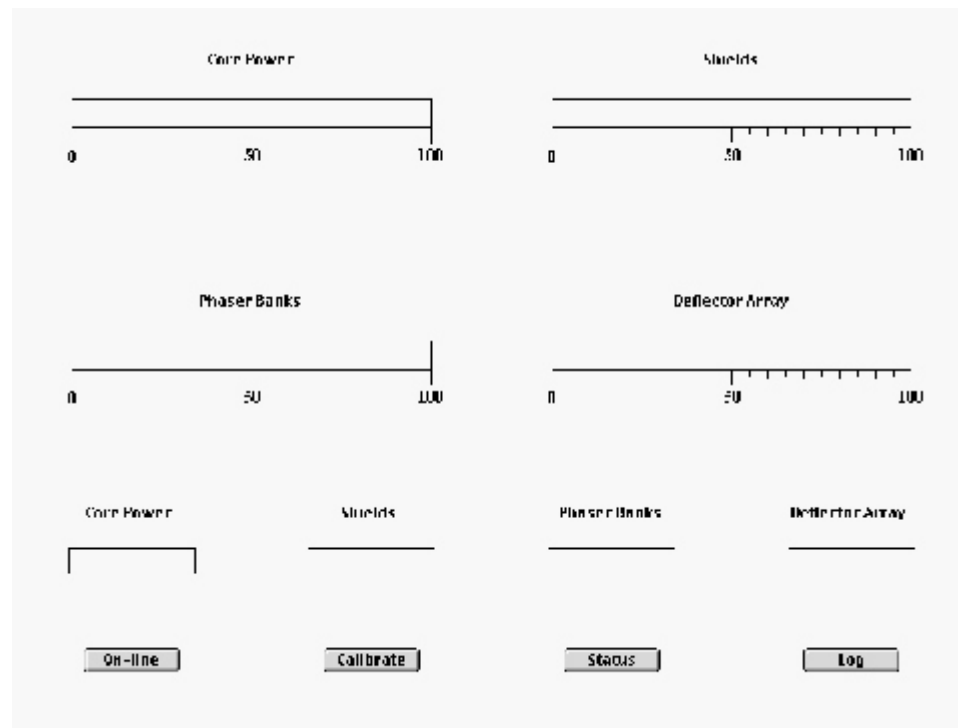


Figure 1. The ES-X15.1 Status Console

There are three essential steps involved in taking readings from the ES-X class status console:

- Calibrating internal sensors
- Taking readings from internal sensors
- Entering status readings into the ship's log

Each step will be described in detail in the remainder of this manual.

Step 1. Calibrate Internal Sensors

Summary of steps to calibrate internal sensors:

- Click "On-Line"

- Click "Calibrate"

- Wait until charge levels in sensor meters read at zero

The ES-X class status console registers readings from numerous internal sensors. Before readings are taken for the ship's log, the internal sensors must be calibrated to ensure that local deviations and fluctuations do not influence the readings.

There are several steps in calibrating the internal sensors. The first step is to take the sensors off-line. This is done by clicking the "On-Line" button on the control panel, as shown in Figures 2 and 3.



Figure 2. Before clicking "On-Line"



Figure 3. After clicking "On-Line"

Once the internal sensors are off-line, it is possible to calibrate the sensors. This is done by clicking on the "Calibrate" button and waiting for the levels in the sensor meters to indicate readings of zero. When sensor meter levels reach zero, the "Off-Line" button should be clicked to bring the sensors back on-line. Figures 4 - 6 demonstrate the process.



Figure 4. Before clicking "Calibrate"

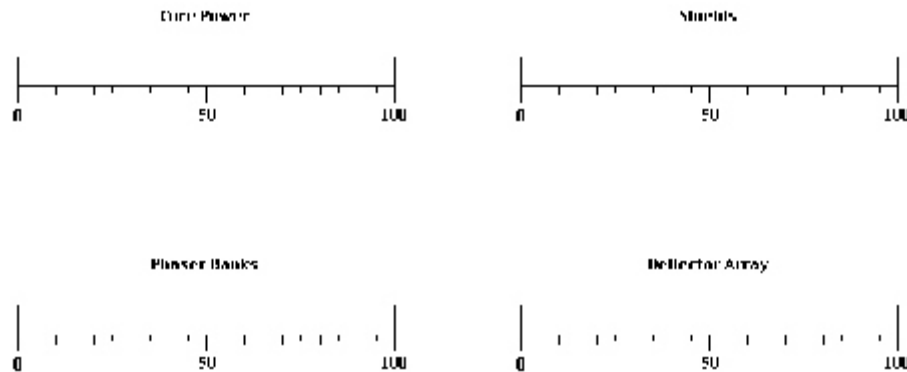


Figure 5. Sensor meter levels at zero



Figure 6. After clicking "Off-Line"

Step 2. Take Readings from Internal Sensors

Summary of steps to take readings from internal sensors

- Click "Status" button

- Wait for sensor levels to settle in sensor meters

- Enter readings from each sensor into appropriate text box

The first step in taking readings from the internal sensors is to initiate current sensor readings. This is done by clicking on the "Status" button, as is demonstrated in Figure 7.



Figure 7. Click on "Status"

Once this is done, the readings in the sensor meters will begin to fluctuate. Wait until all the sensor reading meters are stable (levels stop fluctuating). This is illustrated in Figure 8.

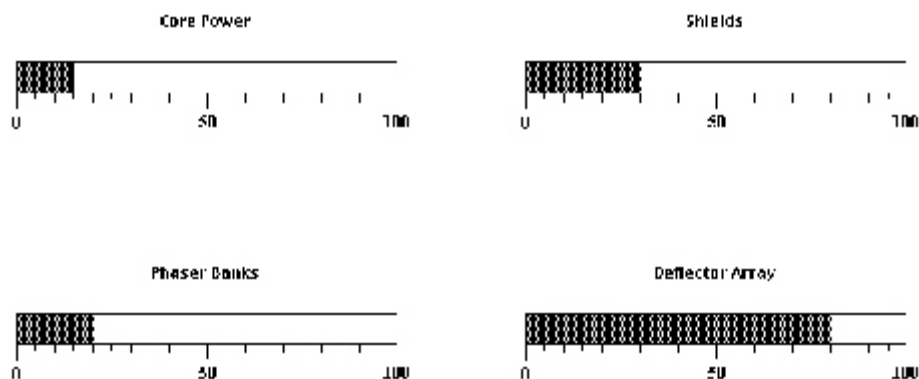


Figure 8. Sensor meters in stable state

Once the sensor readings have stabilized, it is necessary to enter the readings in the appropriate text box for each status sensor. Use the mouse to click in the "Core Power" box. Use the numeric keypad to enter the value from the Core Power sensor meter. Estimate the value as closely as possible. See Figures 9 - 11. Repeat this step for the remaining sensors: Shields, Phaser Banks, and Deflector Array, as illustrated in Figure 12.

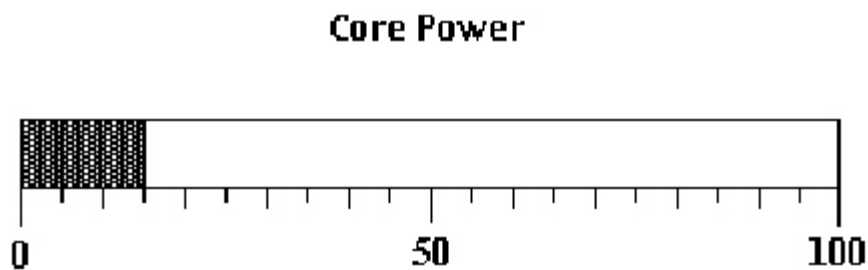


Figure 9. Core Power sensor meter reading

Core Power

Figure 10. Click in the "Core Power" text box

Core Power

Figure 11. Core Power values in Core Power text box

Shields	Phaser Banks	Deflector Array
<input type="text" value="30"/>	<input type="text" value="20"/>	<input type="text" value="80"/>

Figure 12. Sensor Readings in Shields, Phaser Banks, and Deflector Array text boxes.

Step 3. Enter Status Readings into Ship's Log

Summary of steps to enter status readings into ship's log

Click "Log" button

In order to enter the recorded readings into the ship's log, click the "Log" button, as shown in Figure 13.



Figure 13. Click on "Log"

Once this is done, the recorded readings are logged into the ship's computer with a time stamp as determined by the computer system's internal chronometer.

Summary of Status Console Operating Procedures:

Step 1. Calibrate Internal Sensors

- Click "On-Line"
- Click "Calibrate"
- Wait until charge levels in sensor meters read at zero
- Click "Off-Line"

Step 2. Take Readings from Internal Sensors

- Click "Status" button
- Wait for sensor levels to settle in sensor meters
- Enter readings from each sensor into appropriate text box

Step 3. Enter Status Readings into Ship's Log

- Click "Log"

APPENDIX G
EXIT QUESTIONNAIRE

APPENDIX G-1
EXIT QUESTIONNAIRE FOR
CONTROL AND POSTCOMPLETION CONDITIONS

Experiment 28: Bridge Officer Qualifications Exit Questionnaire

1. How heavy was the workload while you were playing the games? Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | |
 Light Moderate Heavy

2. How much attention did you devote to playing the games? How much attention did you devote to doing the letter recall task? Assign a number from one to ten to each of these tasks; your responses must add up to ten.

Games _____

Letter Recall _____

Total _____

3. How do you think you performed overall on the games and the recall task? Rate your performance on the following scale. Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | |
 Poor Average Excellent

4. How well do you think you performed overall on the games? Rate your performance on the following scale. Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | |
 Poor Average Excellent

5. How well do you think you performed overall on the recall task? Rate your performance on the following scale. Please circle the best answer.

A horizontal scale from 0 to 7 with dashed lines. Below the scale are three labels: "Poor" under 0, "Average" between 3 and 4, and "Excellent" under 7.

6. What approaches did you use to deal with the letter recall task? For example, did you repeat the letters out loud? Did you “drop” the first letter in a string to “make room” for the next letter to be spoken? Please describe in as much detail as possible.
7. Did your approach to the letter recall task change at any point during the three days of testing? Please describe in as much detail as possible.
8. Did you make any changes to the way you played the games at any point during the three days of testing? Please describe in as much detail as possible.
9. Overall, what was the hardest of all things to do?
10. Overall, what was the hardest of all things to remember to do?
11. What things do you think you did wrong the most often?
12. What things made it easier to perform your tasks?

APPENDIX G-2

EXIT QUESTIONNAIRE FOR REPRIMAND CONDITION

Experiment 28: Bridge Officer Qualifications Exit Questionnaire

1. Use a one or two word phrase to describe how you felt after you were informed that your performance was falling below the 25th percentile in comparison to the other participants.
2. Did you feel this was an accurate reflection of your overall performance on the games and letter recall task? Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | | |
 Not Moderately Very
 Accurate Accurate Accurate

3. How heavy was the workload while you were playing the games? Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | | |
 Light Moderate Heavy

4. At the start of the day that you were told your performance was falling below the 25th percentile, how much attention did you devote to playing the games? How much attention did you devote to doing the letter recall task? Assign a number from one to ten to each of these tasks; your responses must add up to ten.

Games _____

Letter Recall _____

Total _____

5. At the end of the same day (when you were told that your performance was falling below the 25th percentile), how much attention did you devote to playing the games? How much did you devote to doing the letter recall task? Assign a number from one to ten to each of these tasks; your responses must add up to ten.

Games _____

Letter Recall _____

Total _____

6. How do you think you performed today? Rate your performance on the following scale. Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | | |
 Poor Average Excellent

7. What approaches did you use to deal with the letter recall task? For example, did you repeat the letters out loud? Did you “drop” the first letter in a string to “make room” for the next letter to be spoken? Please describe in as much detail as possible.
8. Did your approach to the letter recall task change after you were informed that your performance was falling below the 25th percentile? Please describe in as much detail as possible.
9. After you were told that you were falling below the 25th percentile, what changes did you make to the way you played the games? Please describe in as much detail as possible.
10. Overall, what was the hardest of all things to do?

11. Overall, what was the hardest of all things to remember to do?

12. What things do you think you did wrong the most often?

13. What things made it easier to perform your tasks?

APPENDIX G-3

EXIT QUESTIONNAIRE FOR RE-INSTRUCTION CONDITION

Experiment 28: Bridge Officer Qualifications Exit Questionnaire

1. Use a one or two word phrase to describe how you felt after you were informed that your performance was falling below the 25th percentile in comparison to the other participants and that you must undergo retraining.
2. Did you feel this was an accurate reflection of your overall performance on the games and letter recall task? Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | |
 Not Moderately Very
 Accurate Accurate Accurate

3. How heavy was the workload while you were playing the games? Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | |
 Light Moderate Heavy

4. At the start of the day that you were told your performance was falling below the 25th percentile and that you must undergo retraining, how much attention did you devote to playing the games? How much attention did you devote to doing the letter recall task? Assign a number from one to ten to each of these tasks; your responses must add up to ten.

Games _____

Letter Recall _____

Total _____

5. At the end of the same day (when you were told that your performance was falling below the 25th percentile and that you must undergo retraining), how much attention did you devote to playing the games? How much did you devote to doing the letter recall task? Assign a number from one to ten to each of these tasks; your responses must add up to ten.

Games _____

Letter Recall _____

Total _____

6. How do you think you performed today? Rate your performance on the following scale. Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | | |
 Poor Average Excellent

7. What approaches did you use to deal with the letter recall task? For example, did you repeat the letters out loud? Did you “drop” the first letter in a string to “make room” for the next letter to be spoken? Please describe in as much detail as possible.
8. Did your approach to the letter recall task change after you were informed that your performance was falling below the 25th percentile and that you needed to undergo retraining? Please describe in as much detail as possible.
9. After you were told that you were falling below the 25th percentile and that you must undergo retraining, what changes did you make to the way you played the games? Please describe in as much detail as possible.
10. Overall, what was the hardest of all things to do?

11. Overall, what was the hardest of all things to remember to do?

12. What things do you think you did wrong the most often?

13. What things made it easier to perform your tasks?

APPENDIX G-4

EXIT QUESTIONNAIRE FOR PRAISE CONDITION

Experiment 28: Bridge Officer Qualifications Exit Questionnaire

1. Use a one or two word phrase to describe how you felt after you were informed that your performance was above the 90th percentile in comparison to the other participants.
2. Did you feel this was an accurate reflection of your overall performance on the games and letter recall task? Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | | |
 Not Moderately Very
 Accurate Accurate Accurate

3. How heavy was the workload while you were playing the games? Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | | |
 Light Moderate Heavy

4. At the start of the day that you were told your performance was above the 90th percentile, how much attention did you devote to playing the games? How much attention did you devote to doing the letter recall task? Assign a number from one to ten to each of these tasks; your responses must add up to ten.

Games _____

Letter Recall _____

Total _____

5. At the end of the same day (when you were told that your performance above the 90th percentile), how much attention did you devote to playing the games? How much did you devote to doing the letter recall task? Assign a number from one to ten to each of these tasks; your responses must add up to ten.

Games _____

Letter Recall _____

Total _____

6. How do you think you performed today? Rate your performance on the following scale. Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | | |
 Poor Average Excellent

7. What approaches did you use to deal with the letter recall task? For example, did you repeat the letters out loud? Did you “drop” the first letter in a string to “make room” for the next letter to be spoken? Please describe in as much detail as possible.
8. Did your approach to the letter recall task change after you were informed that your performance was above the 90th percentile? Please describe in as much detail as possible.
9. After you were told that you were performing above the 90th percentile, what changes did you make to the way you played the games? Please describe in as much detail as possible.
10. Overall, what was the hardest of all things to do?

11. Overall, what was the hardest of all things to remember to do?

12. What things do you think you did wrong the most often?

13. What things made it easier to perform your tasks?

APPENDIX G-5

EXIT QUESTIONNAIRE FOR REDESIGN CONDITION

Experiment 28: Bridge Officer Qualifications Exit Questionnaire

1. Use a one or two word phrase to describe how you felt after you were informed that your performance was falling below the 25th percentile in comparison to the other participants and that you must switch to a new version of the phaser bank.
2. Did you feel this was an accurate reflection of your overall performance on the games and letter recall task? Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | |
 Not Moderately Very
 Accurate Accurate Accurate

3. How heavy was the workload while you were playing the games? Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | |
 Light Moderate Heavy

4. At the start of the day that you were told your performance was falling below the 25th percentile and that you must switch to a new version of the phaser bank, how much attention did you devote to playing the games? How much attention did you devote to doing the letter recall task? Assign a number from one to ten to each of these tasks; your responses must add up to ten.

Games _____

Letter Recall _____

Total _____

5. At the end of the same day (when you were told that your performance was falling below the 25th percentile and that you must switch to a new version of the phaser bank), how much attention did you devote to playing the games? How much did you devote to doing the letter recall task? Assign a number from one to ten to each of these tasks; your responses must add up to ten.

Games _____

Letter Recall _____

Total _____

6. How do you think you performed today? Rate your performance on the following scale. Please circle the best answer.

0-----1-----2-----3-----4-----5-----6-----7
 | | | | | | | |
 Poor Average Excellent

7. What approaches did you use to deal with the letter recall task? For example, did you repeat the letters out loud? Did you “drop” the first letter in a string to “make room” for the next letter to be spoken? Please describe in as much detail as possible.
8. Did your approach to the letter recall task change after you were informed that your performance was falling below the 25th percentile and that you needed to switch to a new version of the phaser bank? Please describe in as much detail as possible.
9. After you were told that you were falling below the 25th percentile and that you must switch to a new version of the phaser bank, what changes did you make to the way you played the games? Please describe in as much detail as possible.

10. Overall, what was the hardest of all things to do?
11. Overall, what was the hardest of all things to remember to do?
12. What things do you think you did wrong the most often?
13. What things made it easier to perform your tasks?