REMEDIATION OF SUSTAINED ATTENTION FOLLOWING TRAUMATIC BRAIN INJURY: Vigilance Task Training and the Generalization of its Effects

by

Jon Jay Van Doren
B.S., University of Florida, 1980
M.Sc., University of Victoria, 1982

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

ACCEPTED
FACULTY OF GRADUATE STUDIES

DEAN

DATE

We accept this thesis as conforming
to the required standard

Dr. Bram Goldwater, Supervisor (Department of Psychology)

Dr. Loren Acker, Departmental Member (Department of Psychology)

Dr. Frank Spellacy, Departmental Member (Department of Psychology)

Dr. Joseph Parsons, Outside Member (Counselling Services)

Dr. Robert Kohlenberg, External Examiner
(Department of Psychology, University of Washington)

University of Victoria

Supervisor: Dr. Bram Goldwater

ABSTRACT

Studies of sustained-attention retraining following brain injury are reviewed, and found to have produced inconclusive results. The reason for this, it is suggested, is that a standard operational analysis of attention has not been applied, as evidenced by considerable inconsistency in the dependent measures and treatment methods used from study to study. The present study addresses this concern by applying well established principles of operant conditioning to the analysis and remediation of attention deficits. After briefly reviewing the variety of task parameters in the attention literature, noting ambiguities inherent in the various conceptualizations of attention, it is decided to train vigilance task performance, a relatively unambiguous and uncontroversial operational definition of sustained attention. Both the principle of immediacy of reinforcement (feedback of correct and incorrect on each trial) and shaping (gradual increase of speed demands contingent on increased performance accuracy) are employed. The issue of generalization is deemed central to concerns of treatment efficacy, and is explored by administration of alternate versions of the same basic vigilance task. Results show that training with immediate reinforcement and speed-shaping produced better acquisition of the trained task than delayed feedback and invariant speed of stimulus presentation. Furthermore, gains resulting from training were essentially limited to the task on which training was conducted, with little evidence for generalization to like tasks employing different stimuli. These results are discussed in terms of the applicability of the construct of sustained attention to head injury rehabilitation.

Examiners:

Dr. Bram Goldwater, Supervisor
(Department of Psychology)

Dr. Loren Acker, Departmental Member
(Department of Psychology)

Dr. Frank Spellacy, Departmental Member
(Department of Psychology)

Dr. Joseph Parsons, Outside Member
(Counselling Services)

Dr. Robert Kohlenberg, External Examiner
(Dept. of Psychology, Univ. of Washington)

CONTENTS

SECTION 1: INTRODUCTION

SECTION 2: METHOD

SECTION 3: RESULTS and DISCUSSION

Subject 1

Skill Acquisition
Generalization Across Tasks
Specific vs. Nonspecific Treatment Effects

Subject 2

Subject 3

Subject 4

Subject 5

Subject 6

Subject 7

Subject 8

Subject 9

SECTION 4. GENERAL DISCUSSION

APPENDICES

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

This monograph would never have been completed without the support and assistance of many people. Most especially, I would like to acknowledge the efforts of the subjects of this study; they have taught me much more than is conveyed within the confines of these pages.

Furthermore, I must acknowledge the guidance and nurturance of my supervisor, Dr. Bram Goldwater. His quiet patience must be boundless, else I surely would have found its limit.

Finally, thanks to my wife, Ann Van Doren, for her understanding and patience. To say that her support contributed immensely to the completion of this project would be an understatement.

This dissertation is dedicated to my loving parents,
Charles P. and Julia E. Van Doren.
Through them, I have been blessed.

SECTION 1: INTRODUCTION

Rationale

Accidental trauma is the leading cause of death for persons under 45 years of age, and an estimated 65% of such injuries involve head trauma (Klauber, Barrett-Conner, Marshall, & Bowers, 1981; Naugle, 1990). Fortunately the likelihood of surviving serious head injury has been increasing. Thanks to improved timeliness of emergency medical response and evacuation, as well as to technical advances in acute treatment of trauma victims, the chances of surviving such injury are now thought to be in the range of 7% - 16% (Naugle, 1990), producing an estimated incidence of nonlethal traumatic brain injury of 168-186 per 100,000 citizens. With a U.S. census of about 250 million, approximately 450,000 new cases are expected nationally each year, continually adding to the numbers of this chronically disabled population.

Survivors can be faced with a broad range of lasting physical, emotional, and cognitive challenges. As their numbers have swelled there has been mounting activism addressing quality of life issues, and an attendant rise in demand for rehabilitative services. The present monograph focuses on treatment of one of the more ubiquitous difficulties plaguing such patients: disorders of attention. Such disorders can be considered as limitations in speed or accuracy under certain conditions, potentially impeding performance of virtually any activity undertaken under those conditions. As we will see, the fact that attention has been quite broadly (and loosely) defined poses serious challenges to its objective study. Before proceeding, therefore, let us confront the task of identifying various definitions of attention for the purpose of distilling their common essence. Doing so at this point will more clearly establish the subject matter of the present paper and will ultimately permit our settling on an operational definition, and, thus, a particular behavior as the focus of treatment.

Defining Attention

The diversity with which the term "attention" has been used reflects considerable ambiguity in its definition, ambiguity borne largely of the mentalistic conceptualizations currently in vogue in the field of attention research. As will be discussed below, this has led to failure to develop a universally accepted operational definition of attention, and has been especially noticeable in studies of attention retraining. Remedial studies have used apparently arbitrary, and widely varying, measures to reflect the construct of attention. The tacit assumption reflected by these studies is that "attention" is a broad response class that includes behaviors addressed by the various treatments as well as by the various dependent measures employed.

However, there has been an emerging trend toward a more operational approach as efforts to study attention retraining have increased. This is no doubt a natural consequence of the demands inherent in treatment studies, since these tend to be quite practical in nature and are therefore less tolerant of ambiguity regarding the behaviors being trained and require clearly defined indices of successful treatment. Unfortunately, despite this trend, the conceptualizations of attention remain somewhat ambiguous and continue to hamper progress in the field. In reviewing the definitions of attention proposed by two groups of influential authors, we will see the beginning of a trend toward a more operational approach. From there, we will establish an operational definition to be used in the current study.

Acknowledging the inconsistency with which the construct of attention has been applied, Posner and Rafal (1987) identified three main contexts in which the term has been used: alertness, selection, and vigilance. By their definition, alertness refers to both tonic and phasic arousal, to "preparedness to respond" which corresponds to overall responsiveness to "any and all signals." Selection, on the other hand, refers to "responsiveness to selected information," to "the processing of specific information, while tuning out other available signals" (p.184). Finally, in the words of the authors, vigilance "refers to the ability to sustain conscious attention over long periods" (p.186), and might therefore be interpreted as consistent maintenance of alertness or selective attention over time.

A similar taxonomy, partly stemming from the work of Posner and his associates, has been proposed by Sohlberg and Mateer (1988). The conceptualization proposed by these authors differs from the above in that it includes two additional performance characteristics, and is described in more operational terms. "Based on the experimental attention literature, clinical observation, and patients' subjective complaints," Sohlberg and Mateer identified five "levels" of attention: focused, sustained, selective, alternating, and divided.

Focused attention refers to responding "discretely to specific visual, auditory, or tactile stimuli," and is comparable to "alertness" of Posner and Rafal (1987). Both refer to performing a task consisting of individual trials with clear forewarning. A focused attention task could, for example, require subjects to press a button each time green appears among a series of displayed colors. Each trial would be preceded by a warning, such as a countdown, indicating when the next color was to be displayed; the interval between trials would not necessarily be predetermined, but could be varied according to subject preparedness. Deficit in focused attention would be exhibited as inconsistent responding to individual trials, even without the requirement of sustaining performance over time. Such patients are rarely referred for intensive outpatient services, not necessarily because of low incidence, but perhaps because admitting physicians and/or third party payors see them as so severely impaired that significant benefit from intensive outpatient therapy is deemed unlikely.

Next, sustained attention refers to maintenance of "a consistent behavioral response during continuous or repetitive activity." Both groups of authors see sustained attention as persistence on a given task over a prolonged period, with responding to discriminative stimuli that are presented without warning. Such a task could be identical to the focused task already described, except that it would consist of ongoing presentation of colors, requiring performance over time and response to targets presented at varying intervals. Patients deficient in this respect might reliably perform discrete trials but show inconsistent performance when many trials are presented in series.

Selective attention refers to maintenance of "a cognitive set which requires activation and inhibition of responses dependent upon discrimination of stimuli." Both groups describe this as performance of a requisite task in the face of irrelevant stimuli. Although to some degree inherent in all tasks, the demand for stimulus discrimination is more explicit and more difficult in selective attention tasks. In such a task, extraneous stimuli would be administered, such as simultaneously playing a tape-recording of spoken colors while the subject performs either of the two previously mentioned tasks. Among patients with difficulties in this respect, introduction of stimuli not relevant to the task at hand leads to exaggerated diminution of performance.

Alternating attention refers to "mental flexibility which allows for moving between tasks having different cognitive requirements." This requires shifting from one task to another, and may therefore be described as requiring the subject to respond to discriminative stimuli indicating a change in contingencies such that formerly irrelevant stimuli become relevant, and vice versa. In keeping with the above examples, alternating attention could involve the experimenter's occasionally announcing a shift in target, back and forth from the color green to the spoken word "red."

Finally, divided attention refers to accurate responding to multiple simultaneous tasks, and may therefore be considered performance of tasks consisting of multiple contingencies, such as pressing a button when green is seen or when "red" is spoken. These two contingencies would be operating simultaneously, with no requirement to alternate.

In light of the variety of senses in which the term attention has been used, the feat of deriving an overall definition becomes one of isolating the factor(s) common to all. To assist in identifying such common elements, let us consider the operational equivalents of each aspect of attention, those aspects of a task which determine the type of attention to which it corresponds (see Table 1.1).

Note that each attention subtype is defined not by performance of a specific task, but refers to characteristics of a potentially broad range of tasks. A focused attention task, for example, has only to consist of discrete and forewarned trials. The nature of the stimuli provided or responses required need not be specified. Furthermore, by varying the task characteristics described above one could develop multiple versions of what would otherwise be the same task, producing a focused attention version, a sustained attention version, and so on.

Table 1.1
Operational Definitions of Attention Subtypes

-SUBTYPE-	-DEFINING CHARACTERISTICS-
Focused	warnings provided, individual trials
Sustained	no warnings, multiple trials
Selective	presentation of irrelevant stimuli
Alternating	multiple tasks performed consecutively
Divided	multiple tasks performed concurrently

Clearly, attention is not presently viewed as a monolithic construct; rather, various subtypes of attention have been emerging in the literature. In general, however, it seems that the basic notion underlying them all is one of consistency of task performance across conditions. That is, a patient would be described as suffering an attention deficit only if performance of a given task was adequate under some conditions and inadequate under others; were performance consistently inadequate across conditions, the difficulty would not be considered one of attention but might instead be related to basic sensory or motor impairment. Therefore when inferring either deficits or improvements in attention, one must demonstrate that elements of the task other than those listed above are not determining performance, being sure to attribute deficits to the degree of control exerted by discriminative stimuli under certain conditions rather than to wholly ineffective discriminative stimuli or lack of available reinforcers. An extreme example of such confounding would be to infer focused attention deficit in a deaf subject who has been asked to perform a task involving auditory warning stimuli. A similar mistake would be to infer improved focused attention in a non-Anglophone subject who had, over time, learned to distinguish the warning stimulus "get ready" from the stimulus "not yet."

The process of defining subtypes of attention, it seems, is simply one of identifying those conditions that are associated with variations in "strength" of the stimulus-response relationship. Strength of stimulus control is reflected in either speed or accuracy of response. For example, were performance adequate (with respect to speed and accuracy) with invariant contingencies but inadequate (with respect to speed or accuracy) with shifting contingencies, then the deficit would be described in terms of alternating attention. Similarly, should performance that had been good without distractors suddenly fail with introduction of extraneous stimuli, a deficit in selective attention would be invoked. Unfortunately it is by no means certain that these subtypes of attention have any existence of their own, other than as heuristic constructs. That is, the present author knows of no empirical evidence indicating that correlations across tasks within subtype are high while correlations across subtypes are low.

Assuming that such subtypes do exist, it seems unlikely that training to improve one would be expected to generalize to another, given that each subtype consists of very different behavioral demands. Training a patient to maintain performance in the face of distraction, for example, would not be expected to improve his shifting from one set of contingencies to another when signalled. The question faced by the current monograph will be whether, even within subtype, training generalizes across tasks (e.g., whether training on one task with distractors will improve resistance to distraction on other tasks, or on the same task with different distractors). Indeed, the degree to which generalization occurs within, and is limited to, individual subtypes would provide empirical support for the existence of such subtypes. However the scope of the present study will be limited to exploration of the extent of generalization within a single subtype, and will not directly evaluate the specificity of generalization to a single subtype. Therefore we must choose one subtype of attention and devise a research protocol for studying its remediation.

Narrowing the Field of Study

When considering the five types of attention, it was apparent that the first two, focused and sustained, are mutually exclusive. The former refers to tasks consisting of individual trials, the latter to multiple-trial tasks. Note that this dichotomy is comprehensive; any task can be classified as one type or the other. We may therefore consider focused and sustained attention to comprise the entire continuum of task durations, "focused" referring to the minimum endpoint of the continuum and "sustained" referring to all tasks consisting of durations greater than a single trial.

The remaining three types of attention are, it seemed, independent of one another. A task's demand for selective attention is manipulated by varying the discriminability or ratio of relevant to non-relevant stimuli, that for divided attention by varying number of simultaneously active contingencies, and alternating attention by varying frequency of shifts between contingencies. Each of these task characteristics can be independently combined in various permutations, but no task may simultaneously include a demand for both focused (single trials) and sustained (multiple trials) attention.

The latter observation, combined with a desire to study a single attention subtype, immediately narrowed the field of study to either focused or sustained attention. A task fitting the defining characteristics of any of the other three attentions (e.g., requiring alternation between sets of contingencies) would also fit those of either focused or sustained attention (consisting of either individual or multiple trials). On the other hand, when designing a task to meet the defining characteristics of focused or sustained attention, one could more readily limit the degree to which other attention-behaviors are involved (i.e., only one stimulus-response relationship, no change in contingencies, and elimination of extraneous stimuli to the extent possible).

Having narrowed the focus of the current investigation to either focused or sustained attention, the field of study was further narrowed by excluding focused attention. This decision was based on the fact, already mentioned, that subjects with deficiencies in focused attention are rarely referred for intensive outpatient therapy. They were therefore not expected to be available for study. Furthermore, the focused attention paradigm is rather artificial in that only discrete, forewarned trials are presented, a circumstance rarely encountered in daily experience and therefore inherently less interesting than the sustained attention paradigm.

The object of the preceding discussion was to provide a clearer understanding, in operational terms, of the construct of attention. This led to refinement in definition of the subject matter of the current study. That done, we are now prepared to consider the remediation of sustained attention subsequent to brain trauma.

Remediation of Sustained Attention

Interest in remediation of sustained attention derives from three sources: prevalence of general attention disorders among the head injured, potential importance of attention remediation to the overall process of rehabilitation, and inconsistency of research findings which has led to controversy regarding the efficacy of sustained-attention remediation. To firmly establish a rationale for the current study, let us briefly examine the first two before turning to a detailed consideration of the third. The latter will lead to identification of factors that have promoted inconsistency of results. This, in turn, will lead to identification of some of the basic characteristics of a research design chosen to minimize such inconsistencies.

Prevalence of Attention Deficits

The pervasive nature of attention deficits among the head injured has been noted by many authors. For support they refer to anecdotal clinical experiences common to all who have worked with such patients, and to several surveys of the enduring complaints of survivors and their families. Such surveys have consisted of symptom checklists including items thought to relate to attention, such as "concentration difficulty," along with a multitude of other frequent complaints such as "headaches" and "fatigue." In light of the fact that an operational definition of attention has only recently begun to emerge, we must bear in mind that these pre-operational estimates of prevalence may differ from those based upon more recent conceptualizations of attention.

In a sample of Korean War veterans, Caveness (1969) found 41% of those who had suffered head injury to report concentration difficulties five years after injury, compared to only 14% of non-head-injured respondents. As McLean et al (1984) have pointed out, it is important to consider the base rate of attention complaints in the general population when attempting to draw conclusions regarding the contribution of head injury in this regard. In a thorough review of the literature, Gronwall (1987) not only referred to the Caveness study but also reviewed the findings of more recent studies. As indicated by studies employing control groups, the base rate of concentration complaints tends to fall in the 4 - 8% range (Gronwall, 1987), the one exception being the 14% rate reported by Caveness (1969). Most studies, including that of Caveness, found incidence of concentration complaints to be much higher among head injury survivors.

Unfortunately the validity of self-report following brain injury is open to question. There are factors which could produce overestimates of disability, as indicated by McKinlay et al (1983) who found only a 9% incidence of attention complaints 12 months post injury among patients without compensation claims pending, but a 38% incidence among patients with pending claims. On the other hand brain injured patients are notoriously poor observers of their own behavior, which could produce underestimates of disability. For example the above-cited McKinlay et al study found relatives of brain injured patients without pending compensation claims to report a 25% incidence, compared to the 9% incidence reported by patients themselves.

Despite shortcomings inherent in the self-report of survivors, and the poorly operationalized nature of "concentration" in the above studies, it is nevertheless generally accepted that failures of concentration are among the most common of posttraumatic symptoms (Lidvall et al, 1974; McLean et al, 1984). This may be taken as circumstantial evidence that attention, as currently defined, is frequently impaired following head injury. This also fits with clinical observations of the present author, who has noted a high incidence of performance inconsistencies in head injured patients.

Attentional difficulties are not only common, but can pose significant limitations on the entire process of rehabilitation. Let us now briefly consider the implications of such impairment.

Importance of Attention to Rehabilitation

As described by Gronwall (1987), statements thought to reflect attention difficulties include: "'...couldn't listen and take notes at the same time.'; 'Fine for the first part, then I just kind of drifted off.'; 'Every time someone near me turned a page or made a noise, I lost the thread.'; and 'I couldn't seem to keep my mind on what I was there for.'" (p. 355 - 356). The variety of behaviors represented by these statements highlights the fact that attention has been a rather vague construct. However, as described above and as reflected in the examples of Gronwall, concern with consistency of task performance across conditions outlined in Table 1.1 unifies the various conceptualizations or subtypes of attention.

The notion that attention refers to skill acquisition and execution sufficiently attests to the importance of this topic to rehabilitation (e.g., see Newcombe, 1982). All aspects of rehabilitation are essentially educational, with patients engaged in the process of acquiring or reacquiring skills specific to each therapy. Clients who exhibit deficits in attention are, under some conditions, limited in the performance of activities that are the focus of therapy. This can frustrate attempts of therapists to instill targeted behaviors and, to the extent that different conditions will be encountered out of therapy, can complicate generalization of these behaviors from the clinic to functional settings. The degree to which conditions contributing to poor performance can be identified, and training to perform under those conditions be provided, the effectiveness of instruction in any therapeutic endeavor should be enhanced. Thus, disorders of attention can limit progress in all areas of therapy (e.g., see Wood, 1984) and efficacious attention retraining would offer promise as a foundation upon which other rehabilitative efforts might build.

The above-mentioned findings with regard to incidence of attentional impairment (albeit loosely defined), even five years after injury, point to the potentially chronic nature of such deficits and underscore the fact that they cannot be expected to spontaneously resolve in their entirety. Furthermore if methods for improving general attentional functions exist, there is little question of the importance of their application among head injury survivors involved in the arduous process of rehabilitation. Having established a satisfactory rationale for investigating attention remediation in general, let us now return to consideration of sustained-attention remediation in particular.

Evaluating Efficacy of Attention Remediation

Review of Attention Retraining Literature

To date, there have been few reported studies of attention retraining after head injury. In fact the primary body of this research is comprised of only six studies (see Table 1.2). Sohlberg and Mateer (1987), Ben-Yishay, Piasetsky, and Rattok (1987), and Gray and Robertson (1989) obtained encouraging results, while Ponsford and Kinsella (1988) and Malec, Jones, Rao, and Stubbs (1984) drew negative conclusions, and Wood (1986) reported mixed results. As reflected by the variety of conclusions drawn by these authors, consensus regarding the efficacy of such therapy has been elusive.

Table 1.2
Summary of Attention Retraining Literature

Authors / Training	Dependent Measures	±
Malec et al (1984) video game practice	Stroop Cancellation Visual RT	- - -
Wood (1986) auditory vigilance task of fixed difficulty, visual vigilance task of variable difficulty	WAIS Digit Span WMS Logical Memory Rey AVLT %-time on task	- - - +
Ben-Yishay et al (1987) multi-component training program	WAIS Digit Span WAIS Picture Completion Visual RT	+ + +
Sohlberg & Mateer (1987) multi-component training program	PASAT	+
Ponsford & Kinsella (1988) computerized tasks of fixed difficulty level	Symbol Digit Modalities Cancellation Choice RT Rating Scale %-time on task	- - - - -
Gray & Robertson (1989) computerized tasks of fixed difficulty level	WAIS Digit Span-arithmetic Composite	+

"±" indicates whether authors drew positive (+) or negative (-) conclusions regarding efficacy of training.

The paucity of research is surprising, in light of evidence suggesting a growing tendency for clinicians to employ "attention remediation" methods with their head-injured clients. This evidence is largely circumstantial; the current author knows of no formal surveys indicating the frequency with which such methods are applied. However, given the high demand for improvements in consistency of performance across the various "attention conditions" outlined in Table 1.1, and the recent appearance on the market of considerable materials ostensibly for such remediation, one can assume the practice of "attention remediation" to be fairly widespread. Also indicative of a high prevalence of attention retraining is the widespread use of personal computers in clinical practice. (A 1984 informal survey of the readership of Cognitive Rehabilitation, a trade magazine, found that 73% of the 63 respondents used computers in therapy.) The fact that much of the computer software marketed for "cognitive rehabilitation" is billed as having been designed for "attention remediation" suggests that these computers are often used in such activities.

The paucity of research despite the clear need for studies of efficacy is probably attributable, in part, to ambiguity in the definition of attention. In the words of Ponsford and Kinsella (1988), "One possible reason for the relative lack of studies of attention following brain injury is the continuing controversy as to the meaning of the concept of attention and its mechanisms." While it is hoped that discussions such as that presented above will eventually lead to consensus regarding the definition of attention, this has not yet been achieved.

In the absence of a universally accepted operational definition of attention, it is not surprising that these few studies have not shared common approaches to training nor have used comparable dependent measures. It is also not surprising, therefore, that considerable variation in results has obtained. We will now turn to specific consideration of the major studies in the area, which will permit greater understanding of the factors possibly contributing to disparity in conclusions. This will, in turn, permit development of a protocol for evaluating attention retraining which might lessen the influence of these factors. For the sake of brevity, only three of the six existing studies will be reviewed in depth. This should convey an understanding of this body of literature sufficient for present purposes.

Most recently, Ponsford and Kinsella (1988) evaluated the effectiveness of such training in a group of ten survivors of severe traumatic brain injury (TBI). Subjects were 17 - 38 years of age, had experienced 10 days - 12 weeks of posttraumatic amnesia, and had begun treatment 6 - 34 weeks following onset.

All showed deficient performance on at least three of five measures derived from a total of three different tasks. Symbol Digit Modalities (Smith, 1968, 1973) involved written and oral translation of a series of printed symbols to numbers, producing a measure for each response modality consisting of the number of correctly translated symbols in a 90-second period. Two-Letter Cancellation (Ben-Yishay et al, 1978) involved a sheet of apparently random letters printed in rows, requiring the subject to mark off every instance of two target letters; two measures, Time Taken and Percent Correct, were produced from this task. Choice Reaction Time (Van Zomeran, 1981) required the subject to monitor a display consisting of four lit buttons all equidistant from a single "home" button, to keep the home button depressed until one of the others lit, then to press the lit button as quickly as possible. The time between lighting of the target button and release of the home button (Decision Time) was the single measure contributed by this task.

In addition, a Rating Scale of Attentional Behaviours (developed by Ponsford and Kinsella) was completed by each patient's Occupational Therapist, and 30-minute videotapes of each patient performing clerical duties were evaluated with regard to percent-time spent on-task. This set of dependent measures, including the above mentioned five and the two developed by the authors, is described fully in Appendix A and was used as the standard by which effectiveness of training was judged.

Impairment on these measures was defined as performance at least one standard deviation below the mean obtained from a group of 16 orthopedically injured control subjects. All head trauma subjects were judged by their Occupational Therapist as displaying "impaired attentional behavior" and were judged by neuropsychological evaluation to exhibit "slow information processing" and "poor selective attention." Most were also said to exhibit "memory problems, as well as impulsivity, poor self-regulation of behavior, and difficulties with planning and problem solving." The authors did not elaborate on the objective indices which led to their subjects being so characterized.

Training entailed performance of five different computerized tasks which provided "repeated practice in responding rapidly, but selectively, to information presented visually on a computer screen," and which "allowed for measurement of changes in accuracy as well as speed of response over time." These five: React, Search, Red Square/Green Square, Spot the Letter, and Evens and Fives are fully described in Appendix B.

The training tasks were administered in two contexts. First, for three weeks patients were required to practice them without explicit feedback. Then for another three weeks the tasks were performed with a therapist providing feedback and graphing results.

Half of the subjects underwent three, and the other half six, weeks of pretreatment baseline. Following baseline, all were treated identically: three weeks of practice, followed by three weeks of practice with feedback, and three weeks of return to baseline. The nature of feedback was not fully described, but appeared to consist primarily of post-session feedback and graphing of performance. During all phases, the psychometric measures of speed of information processing were collected three times per week. The rating scale and videotaping were administered at the beginning and end of each phase.

Performance on training tasks was not reported. As a group, however, the subjects showed gradual improvement on the five psychometric measures. This improvement occurred primarily during baseline and was therefore not attributable to treatment, leading the authors to conclude that their remedial approach had been ineffective. When analyzing their data on a case-by-case basis, however, they noted that five of the ten subjects showed significant improvement on some measures when therapist feedback was provided. One subject improved on four of the five measures, two on three of the five, and two subjects on one measure.

Although the group data indicated no significant effect of training upon the chosen dependent measures, it nevertheless appeared that, for some patients, improvement in psychometric performance was related to receiving therapist feedback regarding training-task performance. The authors were unable to find any characteristics which distinguished those who profited from those who did not. Nevertheless, given encouraging results among some of the subjects, it seems reasonable to wonder whether refinements of treatment methodology might have produced more consistency across subjects, despite the authors' negative conclusions.

Also worthy of note is that the videotaped behavioral measure was particularly insensitive to the attentional deficits that had been ascribed to these patients; a pronounced ceiling effect in percent-time attending to task was noted. This highlights a particularly frustrating feature of studying attentional dysfunction. It is an elusive phenomenon in that patients are commonly able to rise to the occasion when asked to perform a familiar task over a prescribed period of time, when they know they are being critically observed. This is often the case even when reports of both family and staff, as well as results of psychometric evaluation, suggest that attentional performance is indeed impaired.

The other study producing negative findings was that of Malec et al (1984). In a study involving ten subjects who were within six months of injury, these authors reported no effect of performance of an arcade-like video game on three dependent measures derived from the Stroop Tests (Golden, 1978), two paper-and-pencil cancellation tasks, and a visual reaction time task. However, the authors were careful to refer to the application of the video game as "practice" rather than "training," since treatment consisted of simply allowing subjects to play the standard video game without explicitly incorporating aspects of training (other than the immediate feedback inherent in such games). Since data regarding performance of the video game itself were not reported, and the authors themselves pointed out that "one obvious shortcoming of manufactured video games is that they offer little opportunity to individualize performance demands to suit a particular patient's training needs," one might question whether there were indeed any effects of treatment to generalize to the dependent measures.

Sohlberg and Mateer (1987) employed a single-case experimental design with multiple baselines across two dependent measures, and four subjects ranging in age from 25 to 30 years. Time since onset ranged from 12 - 72 months, and coma duration from 24 hours to 7 weeks. At least three experimental phases were provided each subject: baseline (pre-treatment assessment), attention training, and visual process training (training designed to remediate deficits in complex visual stimulus discrimination). The latter training was provided as a means of demonstrating the specific benefits derived from attention training, by showing absence of effect of visual process training on measures of attention.

Two dependent measures were used. Paced Auditory Serial Addition Task (PASAT) involving taped presentation of single-digit numbers at a steady rate of 2.4 seconds per stimulus, requiring the subject to add the current number to the preceding number and to speak the sum aloud, was intended to represent attentional function. The Spatial Relations subtest of the Woodcock-Johnson Psychoeducational Battery, involving presentation of a series of geometric figures accompanied by sets of irregular shapes and requiring the subject to identify those shapes that could be combined to form the given geometric figure, was intended to represent "visual processing abilities." (Description of both of these tasks is included in Appendix A.) Each was administered only occasionally, as a probe.

Attention retraining consisted of a multicomponent treatment package, described briefly in Appendix B. Tasks intended to remediate those aspects of performance described by Sohlberg and Mateer (1987) as sustained, focused, selective, alternating, and divided attention were all administered. Several different tasks thought to address each of these deficit areas were provided.

As was the case for Ponsford and Kinsella, no mention was made of training task performance. However, Sohlberg and Mateer reported evidence for both the generalization and the specificity of their attention retraining procedures, interpreting results as showing that improvements in PASAT scores occurred only during attention-retraining phases and improvements in Spatial Relations scores occurred only during visual process training. Unfortunately, as the authors themselves pointed out, visual inspection of their graphed results was compromised by the infrequency with which probe measures were taken. Adding to difficulties with interpretation was the fact that probes were not administered at equal intervals, and stable baselines were not always obtained on both measures. Additional concern stems from the fact that the PASAT, unlike standard vigilance tasks, involves varied behavioral demands, not all of which would be interpreted as "attentional" in character (e.g., requiring rapid numerical addition rather than simple target detection). Assuming that improved speed of addition is not synonymous with improved attention, it is conceivable that training tasks involving arithmetic could have conferred a benefit to PASAT performance unrelated to generalized improvement in sustained attention. We will return to this issue below.

Ben-Yishay et al (1987) described a program of attention remediation that, like that of Sohlberg and Mateer, consisted of several components providing a variety of training tasks. These have been extensively described elsewhere (Ben-Yishay et al, 1987; Ben-Yishay, Rattok, and Diller, 1979) and will not be repeated here. These authors found significant improvement not only on the training tasks themselves but also, using a pretest-posttest design, on all four psychometric measures employed (visual reaction time, WAIS Digit Span, WAIS Picture Completion, and an in-house Picture Description test). Although the authors reported stability of these dependent measures over the months preceding participation in their study, ostensibly mollifying methodological concerns regarding the use of a pretest-posttest design, this failed to control for nonspecific effects (e.g., a Hawthorne effect) which could have enhanced or emulated effects specific to treatment.

Also reporting positive findings were Gray and Robertson (1989) who, using a single-case design across three subjects, demonstrated improvement across treatment sessions in a composite score reflecting Digit Span (forward and backward) performance and performance of a "simple" arithmetic task involving addition with carrying. Unfortunately the interpretability of these results is compromised by the brief baselines obtained, in conjunction with the visual impression conveyed by their graphed data indicating that the trend of improvement during treatment may have reflected a continuation of a trend established during baseline. Bolstering the authors' claims of demonstrated treatment efficacy was the use of a control measure for each subject. In all three cases the control measure remained level throughout the study, which the authors took as indication that spontaneous recovery (and, presumably, practice effects) was not underlying improvements noted on the attention measure. However the control measure appeared level by comparison to the attention measure even during the brief baseline periods, and the possibility of differential effects of practice on the two measures was not considered.

Wood (1986) investigated the effects of attention retraining among four subjects, ages 19, 25, 30, and 46 years. All had suffered lengthy posttraumatic amnesia, were many years post onset, exhibited "problems of attentional focusing, distractibility, and were unable to sustain attention."

Two classes of outcome measures were employed (see Appendix A). One consisted of "attention-to-task behavior" observed via a time-sampling procedure implemented during standard rehabilitation therapy sessions. The other measures were standard tests of "auditory-verbal memory" (Digit Span, Logical Memory, and Rey Auditory-Verbal

Learning). Presumably, depending upon the extent to which performance on the latter tests was limited by attentional impairment, one might expect to see improvements resulting from attention training.

Training involved performance of two tasks (see Appendix B), one auditory the other visual. Each was administered daily in separate 30-minute sessions over the same 28-day period. The auditory task consisted of rapid presentation of digits, with subjects required to signal at every occurrence of three consecutive odd numbers. The visual task required subjects to monitor a screen and to position a moving light onto various target positions by making a series of well-timed button presses. For the latter task correct responses produced a tone, and incorrect responses were followed by a buzz and repetition of the trial. There was no mention of such feedback on the auditory task.

For the visual task, difficulty was modified by manipulating the speed of the scanning light. There was no mention of such manipulation for the auditory task. A token-reinforcement system was in place for both tasks, that for the auditory task involving both reinforcement and punishment (awarding a token for correct responses and deducting a token for incorrect responses) and that for the visual task involving only punishment (deducting tokens from a bank of tokens provided at the outset). Tokens were exchanged for a presumed reinforcer at termination of each session.

As a group, performance averages on the training tasks clearly improved over the course of 28 therapy sessions. Comparing pretreatment and posttreatment averages of attention-to-task behavior, significant improvement was noted. However methodological problems with the use of a pretest-posttest design limit the conclusions that can be confidently drawn from this result. With regard to the auditory-verbal memory measures, no significant treatment effect obtained, though this could be interpreted as suggesting that improved attention is not sufficient to produce improved memory performance rather than indicating lack of improvement in attentional performance. Individual subject data were presented only for the memory measures, but not for the pretest-posttest measures on which significant group results had obtained.

Even after excluding the studies of Malec et al (1984), Ben-Yishay et al (1987), and Gray and Robertson (1989) on the basis of obvious methodological shortcomings, one is still left with inconsistency across those remaining. Nevertheless, several related factors and conclusions emerge as important from a consideration of these studies, with notable implications for future work in the area. Had such factors been identified and controlled, greater consistency of results might have obtained and more uniform application of effective treatment methods might have emerged. This information will be of obvious importance in designing a method for investigating sustained-attention retraining. Let us therefore consider some of these factors before attempting to devise a protocol with which to study treatment efficacy.

Factors Underlying Inconclusive Results

When considering the results of any study employing brain-injured subjects, there are several factors to weigh in terms of their potential for obscuring results. These include time since injury (Bond, 1976; Dikmen, Reitan, & Temkin, 1983; Lezak, 1979; Mandleberg, 1976; Mandleberg & Brooks, 1975; O'Brien & Lezak, 1981; Parsons & Prigatano, 1978; Tabaddor, Mattis, Zazula, & Phil, 1984), severity of injury, (Brooks, Aughton, Bond, Jones, & Rizvi, 1980; Mandleberg, 1976), type of injury

(Parsons & Prigatano, 1978), and age of subjects (Carlsson, von Essen, & Lofgren, 1968; Parsons & Prigatano, 1978; Wassertheil-Smoller, Tabaddor, Feiner, & Shulman, 1984). Each of these variables is thought to be related to rate and extent of recovery from brain injury. By extension, each could have some modulating effect upon treatments administered to this population. Before proceeding, then, we should consider the possibility that previous studies have differed with respect to some of these variables. Failure to control such differences could have contributed to the obtained lack of consistency in results.

Time since onset, the period of time elapsed since injury, is thought to be inversely related to rate of recovery. That is, performance measures showing a decrement immediately after injury tend to show spontaneous improvement which follows a decelerating course over time. As a group, patients tend to show the preponderance of recovery within the first year after injury, but with some gradual improvement noted even after three to five years (Bond, 1976, 1979; Dikmen, Reitan, & Temkin, 1983; Lezak, 1979, 1987; Mandleberg, 1976; Mandleberg & Brooks, 1975; O'Brien & Lezak, 1981; Tabaddor, Mattis, Zazula, & Phil, 1984;).

Depending upon the research design used, spontaneous recovery could produce false-negative or false-positive findings. Pre-testing followed by treatment and a post-test might show significant improvement, but such improvement might be a product of spontaneous recovery rather than treatment, especially in the first year or two after injury. Conversely, repeated-measure studies employing subjects in the early stages of recovery are faced with greater "noise," in the form of ascending baselines associated with spontaneous recovery, which could obscure treatment effects.

When evaluating the efficacy of a given treatment, the potentially contaminating influence of spontaneous recovery would seem to dictate the study of subjects several years post injury. However there are some considerations weighing against adoption of such a research strategy. Some have suggested the possibility of a critical period for therapy following brain injury, treatment administered earlier having greater effect than that administered later (e.g., see Sohlberg & Mateer, 1989, p.10-11). Were this so, study of procedures administered two years post injury could yield results not indicative of their potential had they been administered sooner; results obtained from subjects many years post-onset therefore might not be applicable to patients treated soon after injury. Furthermore, current rehabilitation practice is typically time-limited, with therapy usually concluded within the first year of injury. In the interest of elucidating the effects of such services, it is therefore necessary to employ recently injured subjects.

These latter concerns dictate the study of more recently injured subjects, despite problems associated with spontaneous recovery. As already mentioned, one experimental method involves taking repeated measures of the dependent measure of interest, obtaining a stable baseline, then administering treatment. If the baseline is sufficiently stable, even if gradually ascending with spontaneous recovery or practice effects, one might assume that clinically significant treatment effects would be apparent. In fact one might argue that a chief consideration in determining the significance of clinical intervention is the degree to which the treatment's effects stand out against a backdrop of spontaneous recovery. This single-subject approach has lent itself well to brain-injury rehabilitation research (see Gianutsos & Gianutsos, 1979). Five of the six sustained-attention treatment studies mentioned above employed such a design (the lone exception being Ben-Yishay et al), but varied with respect to time since onset.

Ponsford and Kinsella (1988) employed subjects only 6-34 weeks after injury, while Mateer and Sohlberg (1987) and Wood (1986) used subjects much farther post-onset. This fact, coupled with the likelihood that more rapid spontaneous recovery occurs early on, and that the effects of intervention might therefore be minor compared with the effect of spontaneous recovery at this stage, provides a possible explanation for Ponsford and Kinsella's failure to obtain treatment effects. On the other hand, the majority of improvement occurred during early baseline and quickly leveled off. This observation is most parsimoniously interpreted as a product of practice rather than spontaneous recovery, since the odds seem slight that spontaneous recovery would have been restricted only to the brief period of early baseline. At any rate, the fact that the dependent measures were stable from late baseline and into the training phase indicates that potential treatment effects were not obscured. Finally, the possible influence of time-since-onset on treatment effectiveness (i.e., the "critical period" theory mentioned above) could not have produced the discrepancy between studies, since this position holds that treatments administered sooner after injury are more effective, and the study obtaining negative results was also the study employing subjects soonest after injury.

Apart from time since onset, other potentially influential subject attributes were quite homogeneous across studies. Judging from generally accepted indices of severity of brain injury (duration of posttraumatic injury and duration of coma), the studies cited above were approximately equivalent in this respect. All were also equivalent with regard to type of injury, employing victims of motor-vehicle accidents.

Only Sohlberg and Mateer's study examined other etiologies, including one gunshot victim and one aneurysm victim. No lesion-site information was presented for any subjects in any of the three studies.

Finally, with respect to age, there was also good homogeneity across studies. For the most part all were young adults falling in the age range at greatest risk for head trauma. The relative uniformity in subject attributes precludes appealing to the possibility that such factors produced the variety of results obtained. It seems appropriate, therefore, that factors other than subject attributes should be considered in terms of their possible contribution to inconsistency of results. We will now consider some of these in detail.

First, *a universally accepted operational definition of sustained attention is lacking*. Sustained attention is conceived as a ubiquitous faculty involved in virtually any task. Thus a single task which is generally acknowledged as a benchmark measure of sustained attention has yet to emerge. The current literature is difficult to interpret because each study has used a different set of criteria by which to gauge the effectiveness of sustained-attention retraining. Adoption of a standard measure might improve consistency and interpretability of results.

Second, *it is not clear that the dependent measures used have been entirely appropriate*. Many could be considered insensitive to improvement in behavioral skills which have been characteristically conceptualized as attentional in nature, or sensitive to improvement in behaviors and behavioral contingencies not falling within the domain of attention outlined above. For example, Wood's (1986) failure to find effects of retraining on measures of auditory-verbal memory does not necessarily reflect the training's ineffectiveness with regard to sustained attention; performance of memory tasks could have been limited by problems in delayed versus immediate stimulus control. Conversely, the finding of significant effects upon PASAT performance (Sohlberg & Mateer, 1987) does not indisputably demonstrate the remediation of *sustained attention.* Given that the PASAT employs at least one skill (e.g., rapid mathematical calculation) that would not be considered a defining characteristic of sustained attention, it is conceivable that therapy improved calculation performance without improving sustained attention. The current literature is therefore difficult to interpret because construct validity of the dependent measures used is open to question. Buchtel (1987) alluded to this when arguing that "many of the popular measures of attention have been too complex to permit an understanding of which functions have been compromised" (p. 372).

A third factor contributing to inconsistent results is that *there has been little uniformity in training methods employed.* This is at least partly attributable to lack of definition regarding what it is that is being trained. Were there more general agreement or clarification as to what aspects of performance are in need of training, it seems likely that greater consistency in training methodology would emerge and, in turn, that greater consistency of results might obtain.

Some of the negative results could have resulted from inadequacies in treatment methodology. In fact, much of the training provided is open to criticism from an instructional standpoint. The various treatments employed have, for the most part, lacked immediate reinforcement and failed to selectively reinforce correct and extinguish incorrect responses. Patients typically have been asked to perform each training task for several minutes before seeing a summary score reflecting performance proficiency. Research largely conducted in the 1960's (Renner,1964, 1968; Sgro, Dyal, & Anastasio, 1967; Wist, 1962) has emphasized the importance of immediacy of reinforcement as a factor influencing acquisition and maintenance of behavior. The delayed nature of feedback could therefore have constituted a limitation in training methodology.

Note that Ponsford and Kinsella (1988) found no benefit offered by training in which all feedback was delayed in nature. However when therapist feedback was introduced, some patients showed improvement on the dependent measures employed. Since therapist feedback, as it is most commonly employed with such training devices, involves some immediate feedback but still tends to emphasize feedback at termination of the session, it is conceivable that the weak effect found by these authors is attributable to greater immediacy of reinforcement and that this factor could be further enhanced to improve the effectiveness of training. Note also that Wood (1986) did use a system of immediate feedback in some training, and that positive effects were noted on some measures. Like Ponsford and Kinsella (1988), Sohlberg and Mateer (1987) tended to use tasks which explicitly incorporated delayed feedback, though these could have also included immediate feedback depending upon the involvement of those administering the tasks.

With the exception of Wood's (1986) visual training task, such training has often neglected to incorporate shaping as an explicit training method, failing to rigorously maintain task demands within subjects' range of proficiency while gradually and systematically increasing difficulty as proficiency increases. Training methods have instead emphasized presentation of tasks consisting of fixed parameters. However it seems justifiable to assume that, were there merit to sustained-attention retraining as a therapeutic endeavor, this would be enhanced by employing proven methods of behavior change such as shaping, which would systematically progress from less-demanding to more-demanding task requirements in a very gradual manner based upon proficiency demonstrated by each subject.

In sum, we have identified two central concerns that must be addressed when devising a study of sustained-attention retraining: adherence to an operational definition of sustained attention when selecting dependent measures, and use of well-established principles of operant conditioning in treatment. These two issues must be adequately dealt with before we can develop a study of sustained attention retraining. With respect to the latter, we have already argued for the application of immediate feedback and shaping as useful elements of treatment; this line of reasoning seems sufficiently elaborated for purposes of moving on to discussion of a research design.

However the former concern, regarding identification of dependent measures, is in need of further exploration. Through our earlier efforts we produced an operational definition of attention and its "subtypes." Consequently, we have already identified a set of characteristics which must be incorporated in tasks acceptable as dependent measures of sustained attention. Specifically, our dependent measures must be chosen on the basis of their reflecting performance of a continuous task that is perceptually and motorically simple, consisting of minimal irrelevant stimuli and a single set of invariant contingencies. The task must be continuous, of course, so as to meet the criteria of inclusion for sustained attention tasks. The other requirements are for the purpose of meeting exclusionary criteria. Perceptual and motoric simplicity help ensure that deficits (and improvements) are associated with the stimulus-response contingency rather than ineffective stimuli or motor limitations. Minimal irrelevant stimuli lessen the selective-attention aspects of the task, while maintaining a single set of invariant contingencies limits divided- and alternating-attention aspects. These task attributes will be discussed below, as we sharpen our focus on specific dependent measures.

Operationally Defining Sustained Attention

Described clinically, sustained attention is loosely operationalized as consistent, accurate and timely responding to stimuli critical to effective execution of an ongoing task. It is the ongoing nature of the task, with unpredictable occurrences of discriminative stimuli, that qualifies it as one requiring sustained attention. In such tasks, patients exhibiting deficiencies in sustained attention may be slow in emitting responses, may fail to achieve a level of accuracy reflecting mastery despite having demonstrated knowledge of the contingencies involved, or may show inconsistent performance. Inconsistencies in performance can occur within or between sessions, with the subject sometimes appearing to have mastered a particular task, while at other times committing glaringly simple errors or slow responses. These difficulties are manifested despite overt maintenance of orientation to the task at hand, a phenomenon familiar to clinicians and supported in the research literature by reports of non-correlation between measures such as eye-fixation and signal-detection rate (Baker, 1960; Jerison & Wing, 1961).

To be portrayed as a measure of sustained attention, a task should be as free as possible of attributes characteristic of other attentional tasks. That is, it should involve presentation of no irrelevant or distracting stimuli, and there should be a single stimulus class which demands a response throughout. Furthermore stimulus discrimination and response demands should be as simple as possible, to limit contributions of factors other than the demand for sustained performance. This would lessen the chance that difficulties with stimulus discrimination or response emission were limiting performance, as well as lessening the chance that improvements in task performance were attributable to improvements in these areas rather than in sustained performance.

Sustained attention has been a topic of laboratory study since World War Two, at which time investigations arose from practical concerns with the watchkeeping behavior of radar operators and the like. The basic paradigm, still employed by these studies, requires subjects to perform vigilance tasks consisting of continuous monitoring of a single sensory channel and to report events presented via that channel. Simple signal-detection tasks are, conceptually, as pure an indicator of sustained attention as possible, and provide an operational definition of sustained attention in terms of the efficiency with which they are performed (Jerison, 1977).

There has been similar work reported in the literature of instrumental conditioning. This literature conceptualizes observing behavior as an operant which produces a discriminative stimulus controlling subsequent behavior. Such research, employing animals, dates at least from the early 1950's (e.g., Wyckoff, 1952). Later, Holland (1958) pointed out the relevance of this literature to human vigilance studies. In so doing, he demonstrated that signal detections operate as positive reinforcers in that the pattern of detections can influence observing behavior just as schedules of reinforcement can influence other operants.

Efficiency of vigilance task performance, or the magnitude of control exerted by discriminative stimuli, can be viewed in two ways: in terms of both accuracy and speed (Davies & Parasuraman, 1981) of response. Apparently these represent two levels of sensitivity to inattention, with errors indicating severe deficiency and slow reaction times indicating subtler deficiency in stimulus control. Thus, when using tasks which engender very low error rates, reaction time has been the measure of choice (e.g., McCormack, 1962). As with error rate, reaction time has been found to increase (e.g., Buck, 1966) and to show heightened variability (e.g., Faulkner, 1962) with time on task.

The few published findings of vigilance performance following head injury indicate higher error rates and response latencies for such subjects than for normal controls, but no difference between groups with respect to vigilance decrement (i.e., within-session diminution in performance) has been reported (Buchtel, 1987). Pilot research using a vigilance task recently developed by Seidel and Joschko (1990, 1991) has shown brain-injured patients to show reasonably good accuracy but apparently slowed average response speed and increased variability in reaction time, compared to non-brain-injured subjects (Seidel, Van Doren, & Joschko, unpublished data). Thus the current study, employing this same simple vigilance task, will use mean and variability of reaction time as primary dependent measures.

A number of studies have demonstrated the effectiveness of such tasks in detecting brain injury (Brouwer & van Wolffelar, 1985; Dencker & Lofving, 1958; Ewing, McCarthy, Gronwall, & Wrightson, 1980; Kaspar, Millichap, Backus, Child, & Schulman, 1971; Mirsky & Orren, 1977; Mirsky, Primac, Ajmone-Marsan, Rosvold, & Stevens, 1960; Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956), and evidence has been presented supporting the construct validity of Seidel and Joschko's vigilance task as a measure of sustained attention in school-aged children (Seidel & Joschko, 1990, 1991).

This task, the Seidel Computerized Attention Test (SCAT; see Appendix A), was validated on both normative and clinical (Attention Deficit Disorder) children in the 6 - 11 year age range. Both parents and teachers were administered Revised Conners questionnaires (Goyette, Conners, & Ulrich, 1978) and, within groups, SCAT measures correlated significantly with attentional but not with nonattentional dimensions of those questionnaires. Significant correlations with psychometric measures were obtained only with those thought to load on a "freedom from distractibility" factor. Other authors have noted that attention improves with age among children, and SCAT performance indeed did improve with age. Significant differences between the normative and clinical groups were obtained on all SCAT measures except Response Time and Response Bias, the former exception conceivably due to a floor effect at this age. Three A.D.D. children were administered the SCAT both on and off methylphenidate medication, with two of the three showing better performance on than off. The third showed good SCAT performance whether on or off medication, and this child's psychiatrist, teacher, and family, all blind to SCAT results, agreed that A.D.D. symptoms had abated and that discontinuation of medication was in order. These findings all support the validity of SCAT as a measure of sustained attention, a measure with sensitivity to experimental manipulation that did not obtain with other attentional measures. The findings of Seidel and Joschko (1990, 1991) therefore support the use of SCAT in the current study.

Recent pilot study using SCAT in brain-injured and normal adult populations indicates that Reaction Time continues to improve beyond the life span studied by Seidel and Joschko (1990, 1991), with normal female adults (N = 9) showing an average reaction time of 567 msec.

(s.d., 56) and average variability of 63 (s.d., 17) on the SCAT X-task. Seidel and Joschko had found 10 - 11 year-old girls to show average reaction time of 683 (s.d., 76) and variability of 100 (s.d., 27). Brain-injured males (N = 9) showed an average Reaction Time of 667 (s.d., 85) and variability of 95 (s.d., 44), performance comparable to that shown by the 9-year-old boys studied by Seidel and Joschko. Although normative SCAT data on adult males is not yet available, Seidel and Joschko's findings showed boys to exhibit faster reaction times than girls. Assuming this pattern is not reversed in adults, comparison of male brain-injured reaction time with female norms will yield conservative estimates of impairment. Thus, unlike Seidel and Joschko's finding with children, current pilot data suggests that Response Time may be sensitive to differences between clinical (brain injured) and normal adult populations.

Pilot study has also shown that SCAT is sensitive to effects of attention training in adult brain-injured patients. Appendix C presents graphic data from two patients, both showing stable baseline followed by dramatic improvement in reaction time with training of the type to be used in the current study.

In summary, SCAT reaction time average and variability measures seem valid as indices of attentional function and are likely to be sensitive to treatment effects. Although this task has been previously validated only in children, current pilot data suggest it is also sensitive to deficits of brain injured adult males. This, in conjunction with the fact that the SCAT was readily available, is computerized (enhancing its ease of use and its accuracy of stimulus delivery and response timing), has a wealth of data (albeit children's) collected on it, and meets the criteria established above in our definition of attention, led to its selection as the primary task to serve as the operationalization of sustained attention.

Having derived a plausible operational definition of sustained attention, we now have a dependent measure to study the effects of retraining. The facts which led us to selection of SCAT are also pertinent to those aspects of performance that will be the target of training. Since pilot study indicated that error rate itself does not appear to be very great among the head injured, treatment will focus on speed as well as accuracy of response. Immediate auditory feedback will be provided, coupled with a shaping procedure that will vary the speed of stimulus presentation according to performance accuracy. Training will not be provided directly on SCAT, but will be administered on tasks very similar to it, reserving SCAT itself as one indicator of generalization of training.

Since attention retraining is ostensibly directed at improving a basic skill with implications for performance of a wide variety of tasks not directly involved in training, the current study will not only evaluate performance of the trained task but will also consider generalization to untrained tasks. The issue of generalization is central to the study of treatment efficacy, especially when evaluating the plausibility of remediating an aspect of behavior that is so basic to the performance of tasks of all kinds. The concern, of course, is that "sustained attention" is not a broadly trainable behavior, that it is instead an aspect of task performance which may respond to training only within the narrow confines of the task at hand. Were this the case, it would be important to define the boundaries of generalization for such training, thereby indicating the extent and variety of training necessary to achieve a given patient's rehabilitative goals.

Thus, to claim that training has produced improvement in sustained attention, one must demonstrate improvement in non-trained tasks thought to represent this construct. Otherwise improvements in performance of the trained task might more accurately be described in terms specific to the task itself rather than by more global claims of attentional improvement. As mentioned above, training will consist of a shaping procedure in concert with immediate reinforcement. Tasks used for training and testing will conform to the simple vigilance paradigm described above, with effects upon sustained attention evaluated by SCAT and similar task performance. By noting the extent to which improvements on untrained but similar tasks occur, we may reach conclusions with respect to the limits on generalization of such training. Let us now turn to a more detailed description of the research protocol.

SECTION 2: METHOD

Objective

The current study addresses two questions: Does immediate feedback, in combination with shaping, produce vigilance task performance superior to that obtained with delayed feedback and, if so, to what degree does improved performance generalize to untrained tasks? The former relates to concerns regarding what may have been suboptimal methods of training used in some prior studies of attention retraining, a factor which could have led to the weak results reported by some authors and thereby contributed to inconsistent findings across studies. The latter relates to concerns regarding the practical significance of attention retraining, and may provide information useful in identifying patterns of generalization arising from such training. This could be a step toward developing systematic methods of promoting desired generalization effects.

Design

A total of four different tasks were administered: one training task, and three generalization tasks to assess the effects of training. Performance of the three generalization tasks was monitored across two conditions: Baseline, during which performance of the three tasks was permitted to stabilize across sessions, and Treatment, during which the training task was administered and generalization tasks were monitored for treatment effects.

Single-subject methodology was used, employing the above mentioned generalization tasks as control measures. To the extent possible, treatment was withheld until all control measures had achieved stability over sessions. It was recognized that the longer the period of stability, the more confidently one might ascribe changes in performance to treatment. On the other hand practical considerations mitigated against adoption of lengthy baseline since subjects were drawn from a population of patients whose length of stay was often unpredictably brief.

The use of multiple control measures is a procedure well suited to exploration of generalization effects. It is more commonly used as a method for supporting a causal relationship between treatment and effect by showing the effect to be specific to the trained task, an application which is actually the converse of demonstrating generalization. However, by selecting control measures varying in presumed similarity to training, such an approach may allow one to draw conclusions regarding the degree of generalization arising from training. That is, it would be expected that maximum treatment effects would be observed on the task being trained, and that the magnitude of this effect would diminish across progressively more dissimilar tasks.

Thus the present study was designed to permit rudimentary mapping of generalization effects obtained from training on a single sustained attention task. Tasks were chosen to allow conclusions regarding intra- vs. cross-modal generalization and to thereby provide information pertinent to systematic planning of the range of rehabilitation tasks necessary to achieve desired generalization goals. For example, a finding of little or no generalization across modality but considerable generalization within modality would suggest that training should be conducted in both auditory and visual formats but that the specific content area of training (e.g., letters geometric shapes) is of less consequence when considering simple vigilance performance.

Of course, stimulus modality marks only one of many possible dimensions to be considered when designing a battery of training tasks for generalization of rehabilitation. Considerably more research would be necessary to produce a body of information sufficient to allow methodical construction of treatment batteries to promote maximum stimulus and response generalization with a minimum of individual treatment tasks.

Subjects

Nine subjects, 14 - 44 years of age, participated. All were clients of the Phoenix, Arizona Good Samaritan Medical Center Head Injury Day Hospital, a full-day interdisciplinary outpatient rehabilitation program involving occupational, physical, speech, and recreational therapy as well as neuropsychological services. All had suffered closed head trauma, with initial Glasgow Coma Scores ranging from 3 - 14 and time since onset ranging from 40 - 224 days. Neither subjects with previous neurologic or psychiatric history, nor those whose length of stay in the Day Hospital would be insufficient for completion of the study, were employed. Please refer to table 2.1 for a case-by-case description of some relevant subject traits.

Tasks (See also Appendices A & B)

As mentioned above, one training and three generalization tasks were administered. All four were computerized vigilance tasks requiring subjects to press the computer's spacebar each time a specified target was presented. Generalization tasks presented stimuli at a fixed onset-to-onset interstimulus interval of 1.5 seconds and stimulus duration of 200 msec, and recorded subjects' accuracy and reaction time to targets. These tasks differed from one another only in terms of the type of stimuli employed (described below). For each session, four measures were derived from each generalization task: mean reaction time of correct responses, standard deviation of reaction time, number of incorrect rejections, and number of false positive responses.

The training task for any given subject was very similar to the generalization tasks. In fact, stimuli employed in the training task were identical to stimuli in one of the three generalization tasks. The training task differed only in that immediate feedback was provided on a trial-by-trial basis, and rate of stimulus presentation varied with performance accuracy. That is, each error (either a false alarm or a missed target) was followed by a buzz and addition of one iteration to the BASIC delay loop between trials, producing a slightly longer interstimulus interval. (See Appendix B.) Each correct target detection produced a beep, and subtracted one iteration from the delay loop. Thus, in the training task, the period allowed for response varied across trials, depending upon accuracy. The initial interstimulus interval of each training session was long enough to allow ample time for subject response, but short enough to insure that reductions in interstimulus interval over the course of the session would encroach on subject reaction time and thereby demand faster responses. In this way, correct responding at faster rates of stimulus presentation was shaped.

Training for most subjects was conducted with the Figural Computerized Attention Task (FCAT), a vigilance task presenting a series of geometric shapes and requiring press of the spacebar upon detection of a specified figure. (For the last two subjects a process of systematic replication was followed, using FCAT only as a generalization measure and instead using the Auditory Computerized Attention Task (ACAT) in training. The latter task involved presentation of speech-synthesized letters of the alphabet, requiring press of the spacebar on hearing a specified letter.)

Table 2.1
Subject Profiles

S#	Age	Ed	Sex	GCS	TSO
1	37	12	M	7	50
2	20	12	M	NA	224
3	30	7	M	8	53
4	19	11	F	4	145
5	21	10	M	14	78
6	20	12	M	NA	66
7	21	15	M	7	126
8	14	8	F	3	40
9	44	18	M	NA	137

S#=Subject Number. Ed=Education (years). GCS=initial Glasgow Coma Score. TSO=Time Since Onset (days). NA=Not Available (medical records did not always report initial Glasgow Coma Score).

As mentioned above, the generalization task employing the same stimuli as those used in the training task was different from training only in that figures were presented at a fixed rate without feedback. This most-similar "generalization" task was so similar to training that one might question its being considered a generalization task rather than a direct indicator of training. Indeed, given that the training task itself could not be administered during baseline, the most-similar generalization task was in fact used to gauge the effectiveness of training. Had training produced no effect upon the most-similar task, there would have been questions regarding whether there had been any effect of treatment at all, and questions regarding generalization would have been rendered moot. (As we shall see, this proved not to be the case.)

Generalization was more fully evaluated via the two vigilance tasks employing stimuli different from training. The Seidel Computerized Attention Task (SCAT) involved sequential presentation of letters on the computer monitor, with the subject required to press the spacebar upon detection of each "X". The Auditory Computerized Attention Task (ACAT), referred to above, was an auditory task analogous to SCAT and used as a measure of generalization for the first four subjects. FCAT, also described above, served as a measure of generalization for the last two subjects since these subjects were trained on ACAT rather than FCAT. Please refer to Appendices A and B for more thorough task descriptions.

Finally, regarding the above tasks, a note concerning their classification as "sustained attention" tasks is in order. They can, of course, be considered sustained since they consist of multiple trials over time. However, to some degree the other types of attention are also involved. For example there are stimulus features irrelevant to the task and therefore incorporating some demand for selective attention. Although some vigilance tasks might have been conceptually "purer" representations of sustained attention (i.e., involving simpler stimuli), concerns regarding subjects' willingness to perform the many hundreds of trials necessary prompted the use of more complex stimuli. Thus, the choice of stimuli used in the above vigilance tasks was dictated by the need to balance demand for simplicity (to minimize non-sustained types of attention) with demand for complexity (to minimize subject boredom).

Dependent Measures

As mentioned above, generalization tasks each produced four measures per session: mean and standard deviation of reaction time, number of misses, and number of false alarms. Each hit contributed a single observation to the sum used to derive mean and standard deviation of reaction time. The number of observations determining the latter two measures was therefore not necessarily constant across sessions, but varied with accuracy since misses and false alarms were excluded.

Misses were defined as failure to respond to a target within that trial's onset-to-onset interstimulus interval (excluding a very brief period between trials during which responses would not register, when the computer was removing the previous stimulus and creating the next. False alarms were, of course, defined by response to a non-target within that trial's interstimulus interval.

A number of considerations weighed against use of the same dependent measures to reflect training task performance. Since interstimulus interval continually shortened with correct responding, the commission of errors was unavoidable during training. Given training's emphasis on speed, the absolute number of these errors was not deemed a useful indicator of performance since it would be the pattern of errors within a session determining speeds attained (e.g., many consecutive correct followed by many consecutive incorrect would produce speedier trials than would alternation of correct with incorrect trials, although both situations could produce the same number of errors). Similarly, mean reaction time of correct responses could be insensitive to changes produced by training (e.g., earlier trials in a session would afford the opportunity for slower reaction times which, when averaged with faster reaction times from later trials, could produce mean reaction time identical to that of a session in which reaction time was unchanging and which failed to achieve interstimulus intervals as brief as those in the former session). Thus, because rate of presentation was not fixed, measures of reaction time and accuracy were judged to be poor indicators of training task performance. It was therefore decided to use minimum interstimulus interval attained per block of 200 trials as a gauge of training-task performance, since this measure would be a direct indicator of maximum speed of accurate responding. Training sessions consisted of one to three blocks, depending upon time permitted.

An unfortunate consequence of using different dependent measures for training tasks and generalization tasks was that performance on one could not be directly compared to the other. However, as may be evident from the above discussion, even if mean reaction time had been used to represent training task performance, it could not have been meaningfully compared to that from the fixed-interstimulus-interval generalization tasks. This is a consequence of the fact that the variable-interstimulus-interval training task would produce a biased sample of reaction times; as the interval between stimuli was reduced in the training task, slower responses would not be detected thereby biasing average speed and spuriously making performance of the training task appear faster than that of the generalization tasks.

Apparatus

As mentioned above, all vigilance tasks were computerized. FCAT was presented on an Apple IIgs running at 2.8 MHz, with standard Apple analog RGB color monitor, and Timemaster clock card by Applied Engineering for measuring reaction time in milliseconds. ACAT was presented on the same system, with addition of an Echo IIb speech synthesis card for production of auditory letters over Sennheiser HD 40 headphones equipped with monaural-to-stereo converter. SCAT was presented on an Apple IIe equipped with standard Apple composite color monitor adjusted to monochrome display; reaction timing was performed by machine language code incorporated into the program.

Procedure

Throughout both conditions generalization tasks were administered in half-hour sessions, a total of two to three times per week for each task. These 30-minute sessions allowed either administration of SCAT (15 minutes) or of FCAT followed by ACAT (5 minutes each). This allowed ample time for delayed feedback, describing performance in terms of average reaction time and overall accuracy, and graphically comparing that day's session to previous sessions. Variation in time of day of task administration was held to a minimum within subjects, varying by no more than one hour.

Additional brief sessions approximately 15 minutes in duration were administered two or four times per week (depending upon constraints imposed by each subject's schedule of therapies). During baseline these sessions consisted of presentation of commercially available computerized tasks that were unpaced in nature, thereby keeping time spent working on the computer constant across conditions. During treatment, these sessions were devoted to administration of the above mentioned training task. Early subjects were administered training in which targets became progressively more complex over sessions, as a plateau in training task performance was reached. This practice was discontinued for later subjects, as it appeared to have no effect upon generalization task performance.

For each session, the subject was seated directly in front of the computer, with the center of the monitor at eye level and at arm's length. SCAT was administered in a different room and on a different computer than ACAT and FCAT (see Apparatus). Both rooms were relatively free of overt distraction, though muffled conversation could occasionally be heard. After being seated, standard instructions were read at each session (see Appendix D). These instructions differed slightly for generalization and treatment sessions. In both cases, both accuracy and speed of performance were emphasized. After completion of each task the subject was engaged in conversation in which speed and accuracy measures were reviewed, compared to those of previous sessions, and graphically recorded.

SECTION 3: RESULTS and DISCUSSION

Results will be presented individually for each subject. Minimum training-task interstimulus interval achieved in each 200-trial block will be graphically portrayed over time, as will mean reaction time, standard deviation of reaction time, and number of errors for each session of each generalization task. Error graphs will be presented as histograms, with total height of each histogram representing total errors per session and ratio of false alarms to misses proportionately represented on each.

Table 3.1 presents the SCAT performance data of each subject, with comparison to group data obtained from non-brain-injured control subjects. As inspection of this table reveals, only Subject 3 was within one standard deviation of control mean on all SCAT measures. Since this is the commonly used cutoff for establishing deficit on a particular measure, we would be justified in concluding that all subjects, save number 3, exhibited deficit in sustained attention as defined in the current study. No distinguishing features which could account for Subject 3's discrepancy from the others were discerned. Further inspection of Table 3.1 shows mean reaction time to be the measure most sensitive to attention deficit, with 8/9 subjects performing below the cutoff.

In the spirit of single-case experimental approaches, each subject will be presented as an individual experiment. As will be described, procedure was essentially held constant over the first four subjects and results were very consistent, producing confidence in the reliability of results obtained with this procedure. At that point, major procedural changes were introduced to clarify some ambiguities in interpretation of results. After describing results obtained from each of three subjects exposed to this second procedure, a return to the original procedure was undertaken with minor variations. These minor

Table 3.1
SCAT Performance of Current Subjects vs. Normative Sample

Subject	Mean RT	SD of RT	Hits	False Alarms
Control Grp	571.6 (56.6)	63.5 (19.5)	90 (0.0)	0.3 (0.6)
1	**685 (2.0)**	**116 (2.7)**	90	0 (-0.5)
2	**819 (4.4)**	81 (0.9)	90	**2 (2.8)**
3	556 (-0.3)	76 (0.6)	90	0 (-0.5)
4	**785 (3.8)**	80 (0.8)	90	0 (-0.5)
5	**663 (1.6)**	**150 (4.4)**	**88**	**1 (1.2)**
6	**656 (1.5)**	**85 (1.1)**	**89**	**1 (1.2)**
7	**740 (3.0)**	**140 (4.0)**	**89**	**1 (1.2)**
8	**889 (5.6)**	**97 (1.7)**	90	0 (-0.5)
9	**954 (6.8)**	**199 (6.9)**	**86**	**2 (2.8)**

Control Group (N=28) information is reported as means (standard deviations in parentheses), based upon unpublished data (Seidel, Van Doren, and Joschko).

Information for Subjects 1-9 is that obtained from each's first session of SCAT, (z-scores, referenced to control group, in parentheses).

All data greater than one standard deviation below control mean is printed in boldface.

variations produced successful systematic replication of the original results, further increasing confidence in reliability of the present findings.

Thus subjects may be thought to be divided into three groups corresponding to the three major procedures applied. After presentation of the subjects of each group, a brief summary and set of conclusions will be presented. This will not only serve to integrate findings obtained within each procedural group, but will establish the rationale for each shift in procedure.

Subject 1

Procedure

Standard procedure was followed, as described in the preceding Method section, except that an unforeseen (and still unexplained) equipment failure aborted a single ACAT session. Following baseline, treatment was administered four times per week.

Results

Subject 1's mean reaction time in the first SCAT session was 2.0 standard deviations below the mean of an uninjured control group (Seidel, Van Doren, & Joschko, unpublished data). This attests to Subject 1's impairment with respect to sustained attention, as currently defined.

Training Task

This task was very similar to FCAT, differing only in that immediate feedback was provided and speed of stimulus presentation varied. Improvement was noted within and across the first two training sessions (Figure 3.1.0). This improvement was reflected in diminution in minimum interstimulus interval achieved in each block of 200 trials Recall that interstimulus interval was dependent on training-task performance accuracy. The rationale for gauging performance of the training task in terms of interstimulus interval has been presented above, in the preceding Method section.

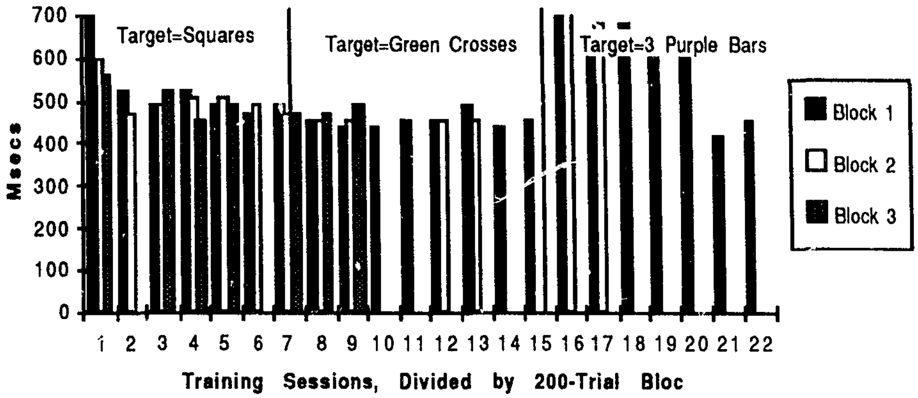


Figure 3.1.0: Subject 1's Minimum ISI Achieved in Training, Msecs.

FCAT

Mean reaction time (Figure 3.1.1) was stable during baseline, varying from 605 to 685 milliseconds across sessions. With treatment, mean abruptly dropped to 555 milliseconds and progressively diminished to 507 milliseconds by the final session.

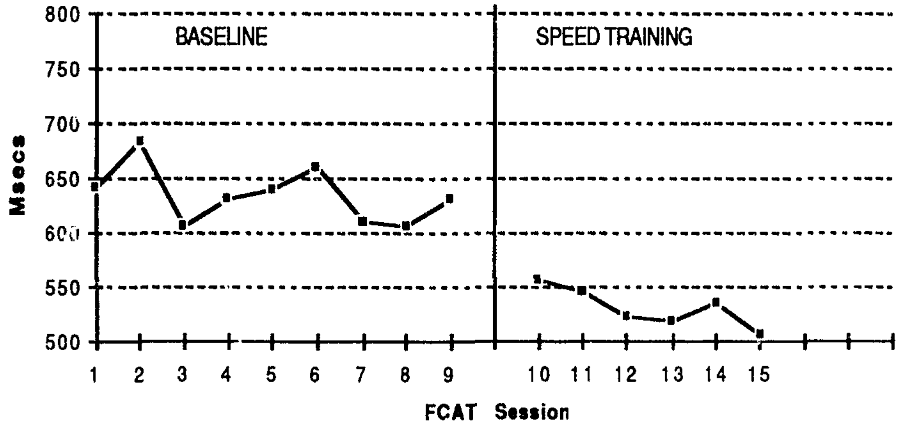
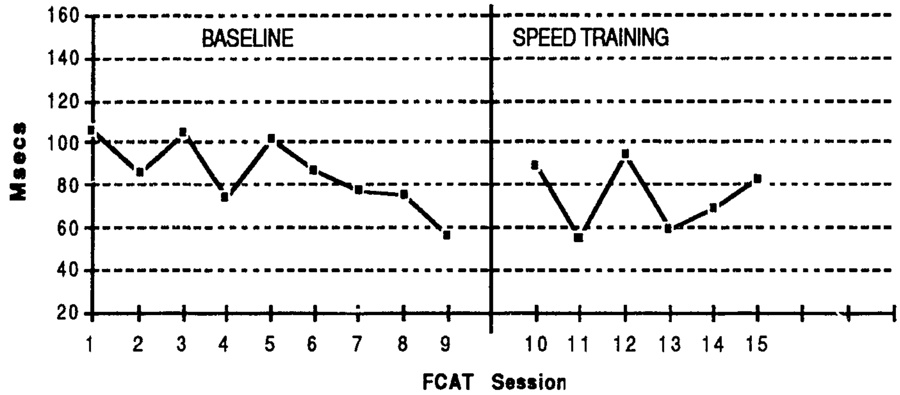


Figure 3.1.1: Subject 1's Mean FCAT Reaction Time, in Millisecon

Standard deviation of reaction time (Figure 3.1.2) diminished somewhat during baseline, presumably reflecting an effect of practice.

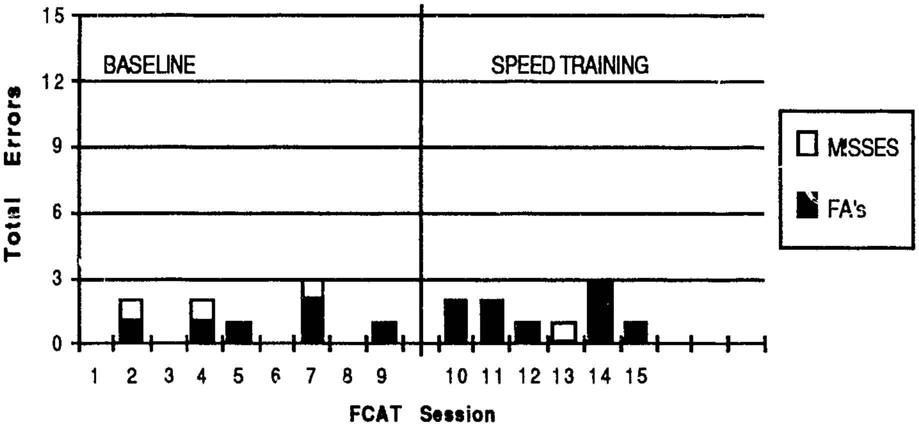
Treatment produced no apparent effect, varying within the range established during baseline.

Figure 3.1.2: Subject 1's Standard Deviation of FCAT Reaction Ti in Milliseconds.



As for error rate (Figure 3.1.3), sessionwise frequency of false alarms appears to have increased very slightly with treatment, while that of misses remained stable (and very low) across conditions.

Figure 3.1.3: Subject 1's Total FCAT Errors, by Ty

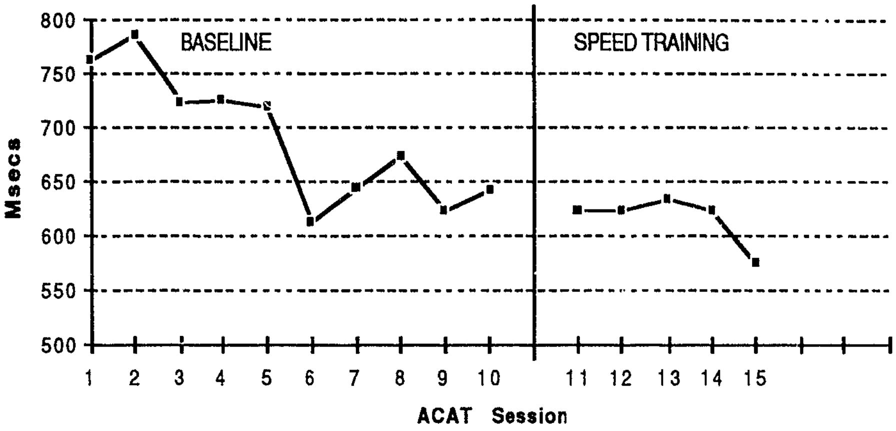


ACAT

Mean reaction time (Figure 3.1.4) diminished during the first half of baseline, remaining stable through the second half and into Treatment.

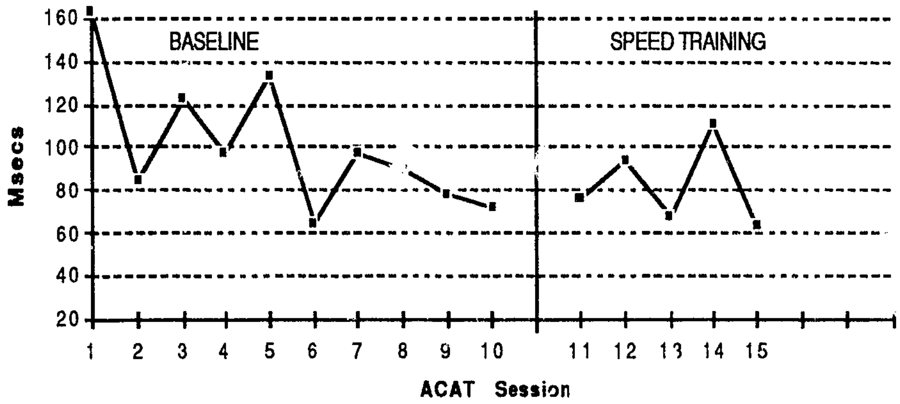
With treatment, reaction time remained extraordinarily stable across sessions, dropping only in the final session. This final drop was the only suggestion of a possible treatment effect.

Figure 3.1.4: Subject 1's Mean ACAT Reaction Time, in Millisecon



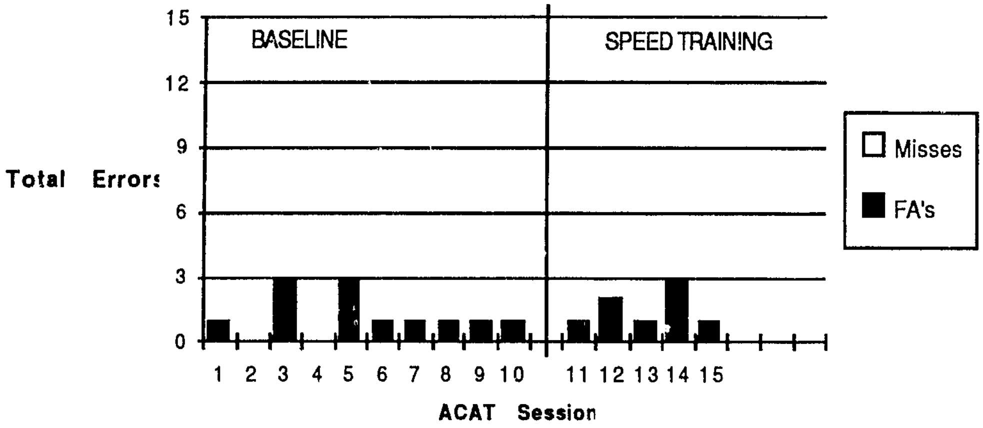
Standard Deviation of reaction time (Figure 3.1.5), like mean, became relatively stable during the last half of baseline and remained so thereafter.

Figure 3.1.5: Subject 1's Standard Deviation of ACAT Reaction Ti in Milliseconds.



Inspection of error rate (Figure 3.1.6) shows that no misses occurred. False alarm rate was quite consistent across the last half of baseline, rising negligibly with treatment.

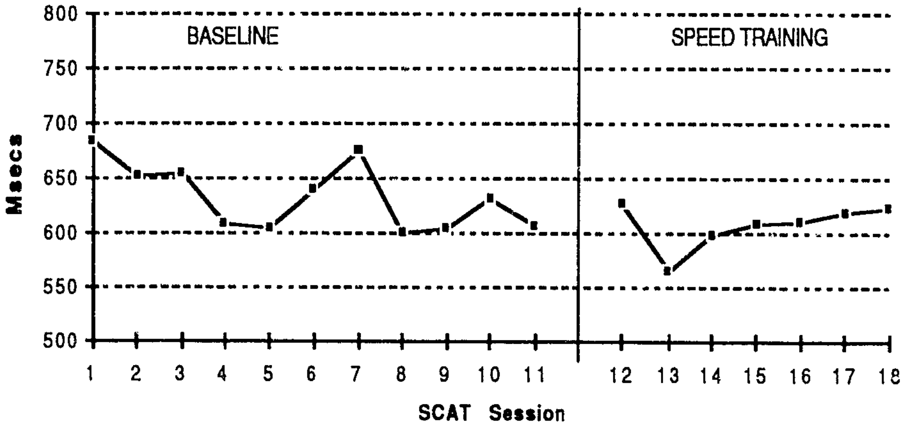
Figure 3.1.6: Subject 1's Total ACAT Errors, by Ty



SCAT

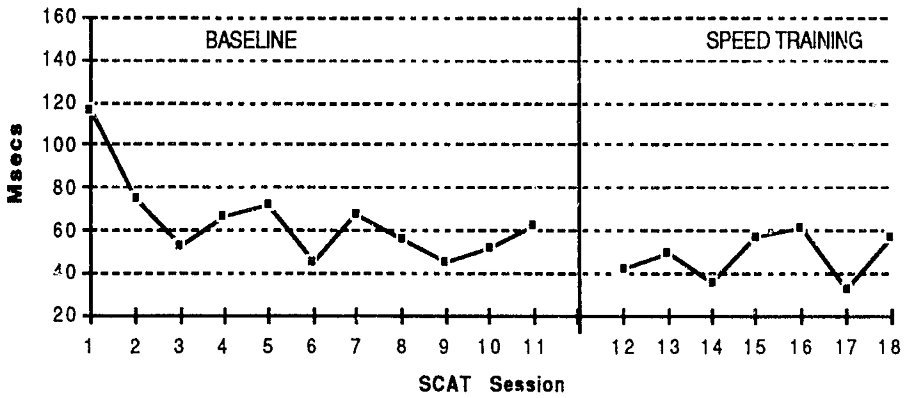
Mean reaction time (Figure 3.1.7) produced a shallow negative slope across baseline sessions. Treatment produced no clear effect on this measure.

Figure 3.1.7: Subject 1's Mean SCAT Reaction Time, in Millisecon



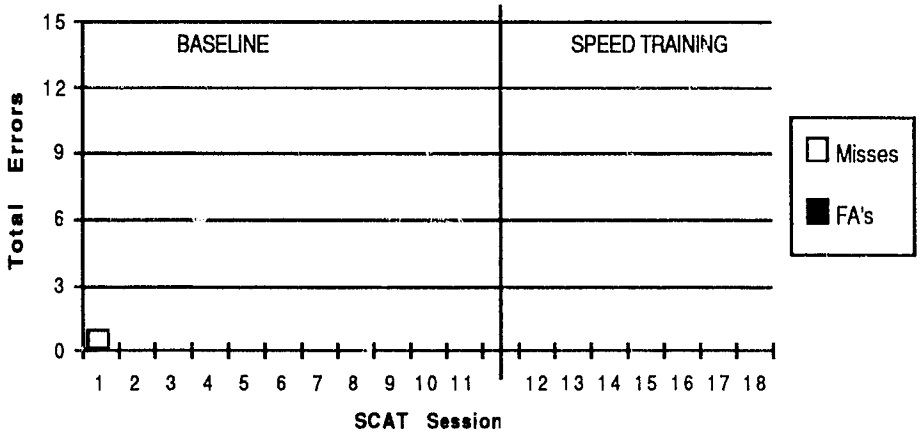
Standard deviation of reaction time (Figure 3.1.8) was stable throughout most of baseline, and remained so across conditions.

Figure 3.1.8: Subject 1's Standard Deviation of SCAT Reaction Ti in Milliseconds.



The only error (Figure 3.1.9) was a single false alarm, occurring in the first session.

Figure 3.1.9: Subject 1's Total SCAT Errors, by Ty



Conclusions

Skill Acquisition:

Combination of immediate feedback with speed-shaping was effective in producing improved performance of the training task, in that maximum speed of response improved within and between sessions. Furthermore, although baseline conditions produced some improvement in vigilance-task performance over sessions, application of treatment employing immediate feedback and speed-shaping produced additional positive effect. Unfortunately this was limited to average reaction time on the task which was nearly identical to training.

Generalization Across Tasks:

As mentioned above, FCAT showed faster mean reaction time following FCAT-like speed training. Generalization was limited to FCAT, however, and was not evident with respect to ACAT and SCAT measures.

Specific vs. Nonspecific Treatment Effects:

Treatment effects were attributable to the subject's having learned to perform the training task more quickly or accurately, and having generalized performance improvements to similar tasks. The effects of treatment therefore were not attributable to nonspecific (e.g., "motivational") influences. Evidence for this interpretation includes:

Learning of the trained task. A learning curve was evident on the training task, performance improving gradually both within and between sessions. This observation is most consistent with the inference that the subject had learned to perform the FCAT-like training task more quickly or precisely. Nonspecific factors would not be expected to produce results indicative of gradual skill acquisition.

Pattern of improvement on untrained tasks. Performance improvements with treatment on tasks other than those on the training task were largely limited to the generalization task most similar to training. This observation is compatible with an interpretation of generalization of specific skill-learning across tasks. To reconcile this result with a "nonspecific" interpretation, one would have to propose that such nonspecific influences would affect only the task most similar to training.

Subject 2

Procedure

The procedure used with this subject differed from that employed with Subject 1, in that a single discrimination training session employing immediate feedback was administered prior to baseline. The change in procedure was necessary because the subject had demonstrated inability to perform FCAT without such training. The dependent measures, graphed in figures 3.2.1 through 3.2.9, represent performance of the standard computerized tasks lacking immediate feedback.

Discrimination training involved one 600-trial session in which stimulus rate was constant and equal to that used by the computerized dependent measures, differing from the FCAT dependent measure only in that beeps followed hits and buzzes followed errors. Discrimination training differed from speed training only in that rate of stimulus presentation was constant rather than tied to performance accuracy.

Prior to any FCAT discrimination training, Subject 2 had successfully performed one session of SCAT. Therefore Session 1 of SCAT preceded training, but all sessions of FCAT and ACAT followed the single session of discrimination training.

Phase 2 consisted of speed training, employing both accuracy feedback and rate of stimulus presentation linked to performance accuracy. This condition was identical to that administered Subject 1, except that training was administered only twice per week. Only three sessions of treatment were administered due to this subject's abrupt discharge from therapy.

Finally, as had been the case for Subject 1, one ACAT session was not able to be administered due to equipment failure.

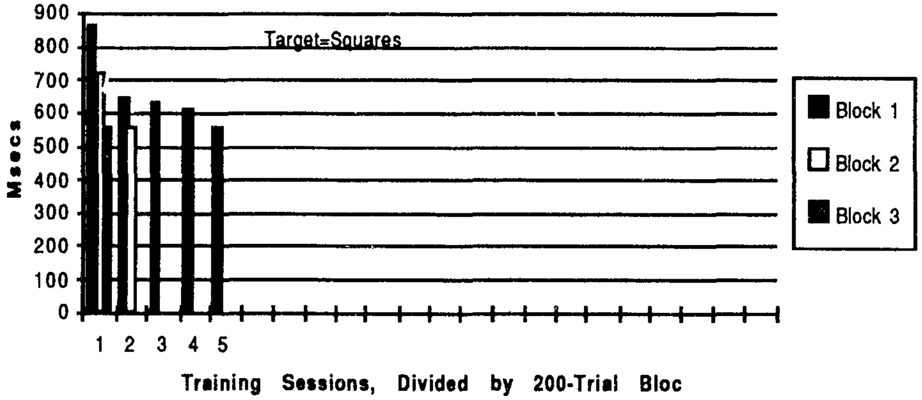
Results

With respect to mean reaction time in the first SCAT session, Subject 2 was 4.4 standard deviations below the normative group mean mentioned previously (Seidel, Van Doren, and Joschko, unpublished data). As was the case for Subject 1, this indicates significant impairment in sustained attention as it has been operationalized in the present study.

Training Task

As had been the case for Subject 1, consistent improvement in performance of the training task was noted within and between sessions but quickly reached asymptote (Figure 3.2.0).

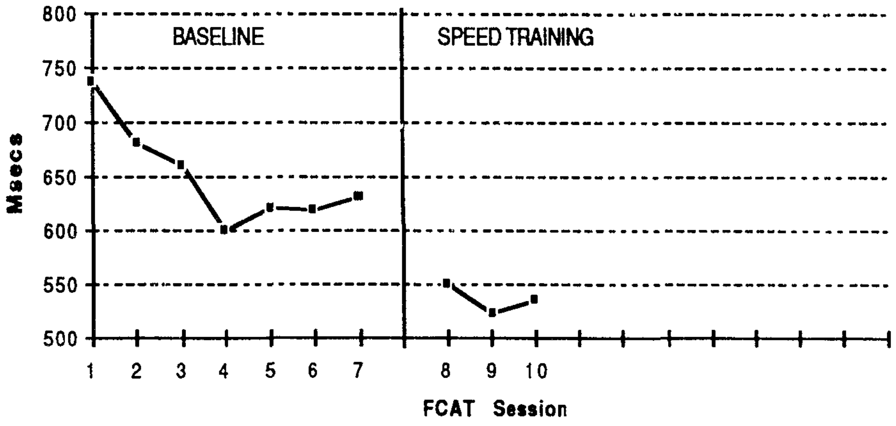
Figure 3.2.0: Subject 2's Minimum ISI Achieved in Training, Msecs.



ECAT

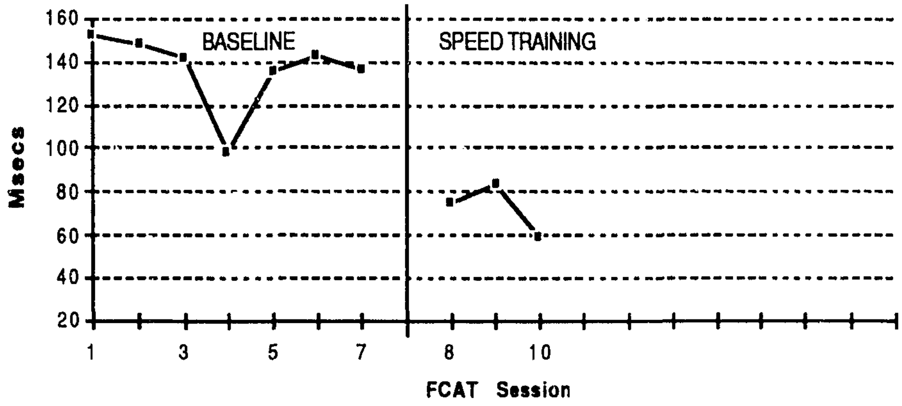
Mean reaction time (Figure 3.2.1) exhibited dramatic drop following treatment.

Figure 3.2.1: Subject 2's Mean FCAT Reaction Time, in Millisecon



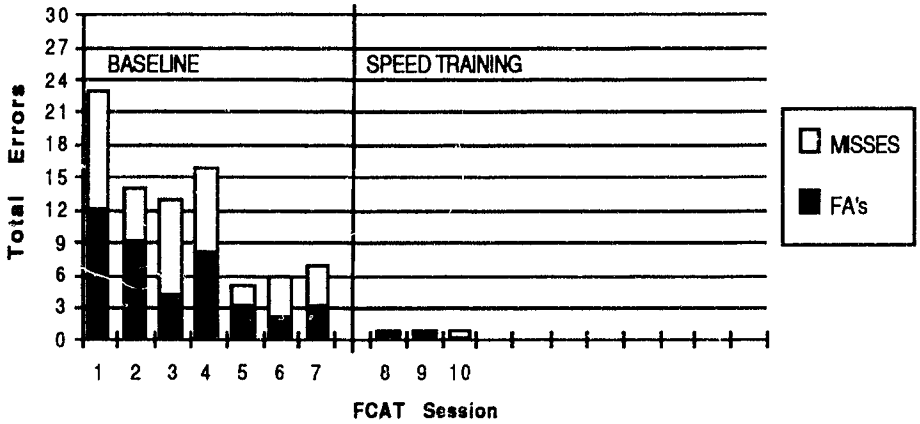
Standard deviation of reaction time (Figure 3.2.2), like the mean, clearly diminished following treatment.

Figure 3.2.2: Subject 2's Standard Deviation of FCAT Reaction Ti
in Milliseconds.



Error rate (Figure 3.2.3) showed steady diminution throughout baseline. Although errors were consistently less frequent with treatment, the downward baseline trend obscured possible treatment effect.

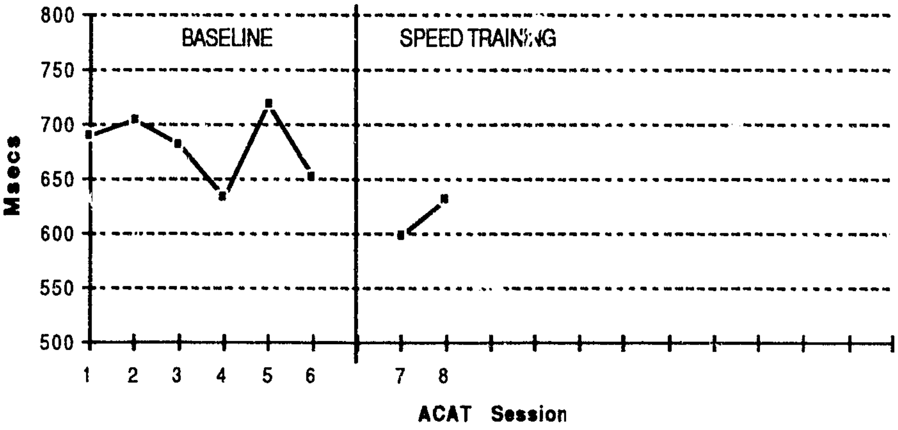
Figure 3.2.3: Subject 2's Total FCAT Errors, by Ty



ACAT

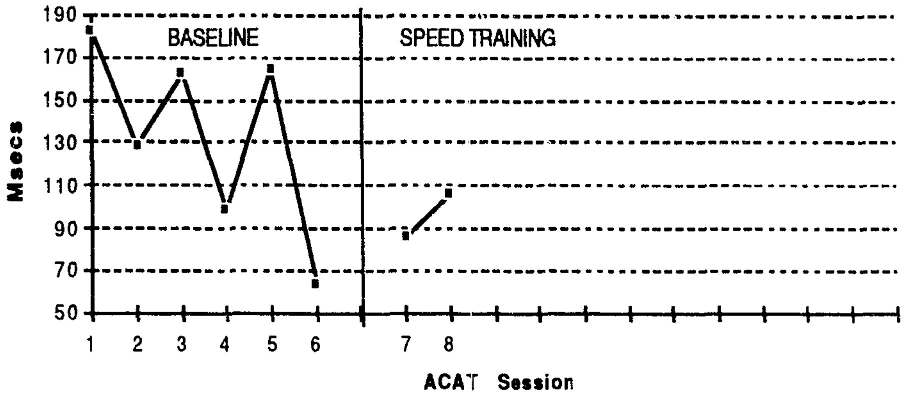
Mean reaction time (Figure 3.2.4) showed no effect of practice during baseline. The first session after treatment dropped, but the second session rose back to within baseline levels. It is unfortunate that additional sessions were not possible.

Figure 3.2.4: Subject 2's Mean ACAT Reaction Time, in Millisecon



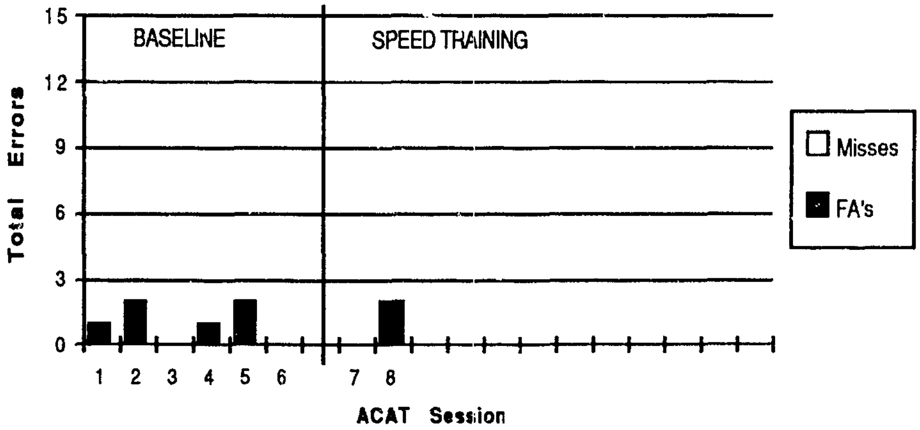
Standard deviation of reaction time (Figure 3.2.5) showed considerable intersession variability, though an overall trend of diminution during baseline. There was a suggestion of reduced cross-session variability in standard deviation following treatment.

Figure 3.2.5: Subject 2's Standard Deviation of ACAT Reaction Ti in Milliseconds.



Inspection of error rate (Figure 3.2.6) shows that no misses occurred and false alarms were sporadic throughout. No effect of treatment was seen.

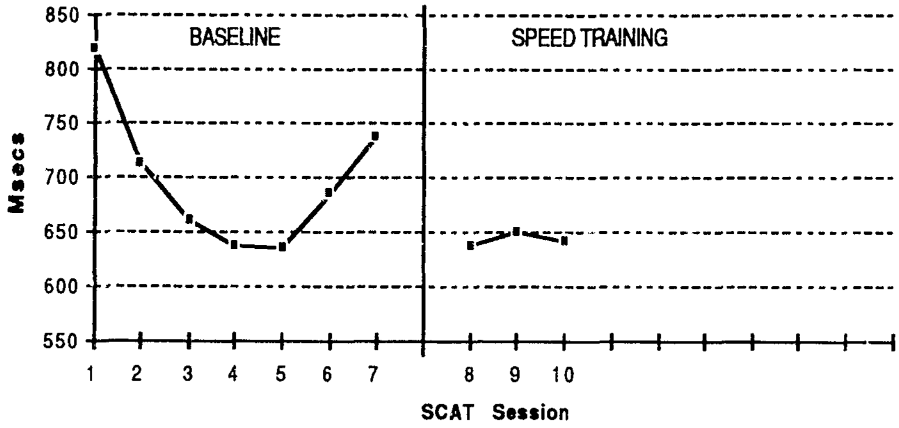
Figure 3.2.6: Subject 2's Total ACAT Errors, by Ty



SCAT

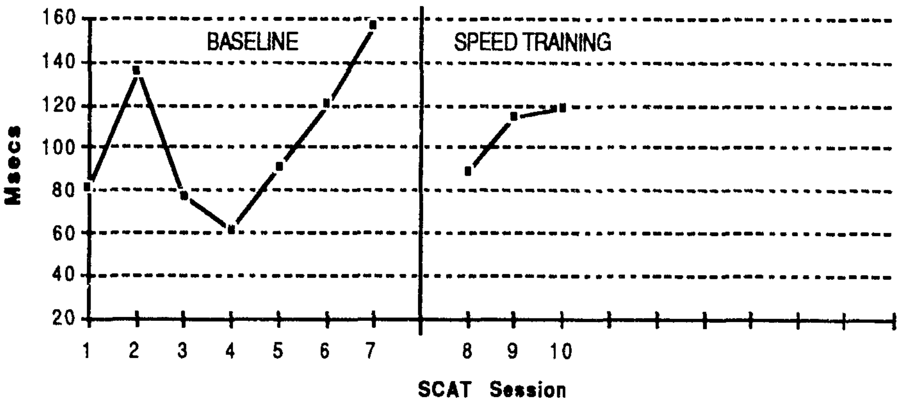
With respect to mean reaction time (Figure 3.2.7), treatment reversed the upward trend that had characterized baseline, but failed to achieve faster means than had been achieved during baseline.

Figure 3.2.7: Subject 2's Mean SCAT Reaction Time, in Millisecon



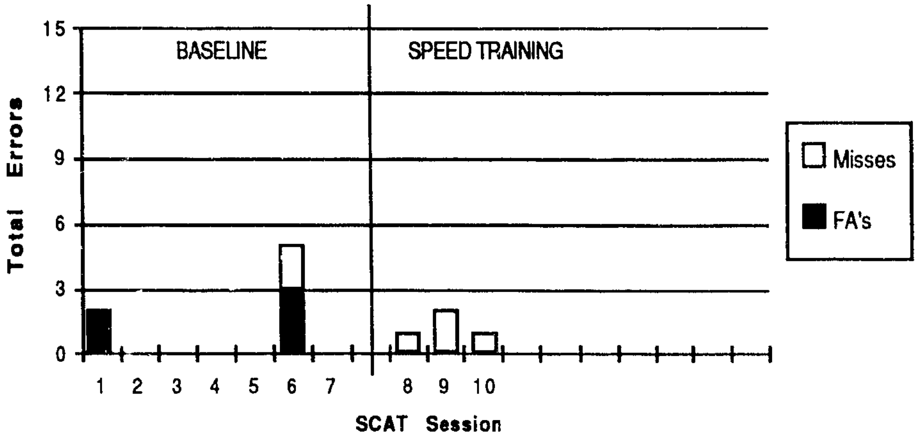
There was no effect of treatment on standard deviation of reaction time (Figure 3.2.8).

Figure 3.2.8: Subject 2's Standard Deviation of SCAT Reaction Ti in Milliseconds.



Error rate (Figure 3.2.9) showed a slight rise in misses following treatment.

Figure 3.2.9: Subject 2's Total SCAT Errors, by Ty



Conclusions

Skill Acquisition

Immediate feedback. Since there was only one session of discrimination training at the outset, it is not possible to numerically support the observation that provision of immediate feedback was essential in acquisition of skills necessary for performance of the training task. However, as will be described below, generalization in the form of improved performance of the vigilance task most similar to training was noted. This is indicative of learning associated with discrimination training.

Speed shaping. Speed shaping, which included immediate feedback, produced improved performance of the training task, both within and between training sessions.

Generalization Across Tasks

Discrimination training. As indicated above, the effect of FCAT-like discrimination training was not reflected in numerical measures of task performance since a pre-treatment session was not completed. However, following a single session of discrimination training the subject was able to complete vigilance sessions without catastrophic reaction. This in itself represented significant improvement and was indicative of generalized performance improvement.

Speed training. Following FCAT-like speed shaping, FCAT mean reaction time, standard deviation, and error rate dropped. No such evidence of generalization was found for ACAT and SCAT.

Specific vs. Nonspecific Treatment Effects

Evidence supporting the contention that treatment effects were specifically related to training was the same as for Subject 1.

Learning of the trained task. Performance of the speed-shaping training task produced a learning curve within and across sessions, a result that could not readily be attributed to nonspecific factors.

Pattern of improvement on untrained tasks. Again, as for Subject 1, improvements on untrained tasks were limited to the one most similar to training. This result is consistent with generalization of skills specific to rapid performance of FCAT-like tasks.

Subject 3

Procedure

There were two ways in which the procedure used with this subject differed from that outlined above: an additional experimental phase was employed, and there was a three-week hiatus between first administration of SCAT and subsequent data collection.

As had been the case with Subject 2, Subject 3 showed consistently high FCAT error rate during baseline, chiefly consisting of false alarms. However Subject 3 displayed lower error rate than Subject 2 and did not exhibit the highly emotional reaction exhibited by the prior subject. This permitted administration of several sessions prior to initiating discrimination training, thereby establishing a true pretreatment baseline and allowing some exploration of the separate effects of immediate feedback and speed shaping.

Due to practical considerations regarding the total number of sessions possible, baseline was limited to only four sessions. The relatively stable FCAT reaction time and error rates obtained over those four sessions mollified concerns regarding brevity of baseline.

Discrimination training involved training sessions, conducted four times per week, in which stimulus rate was constant and equal to that used by the generalization tasks, differing from FCAT only in that beeps followed hits and buzzes followed errors. Thus this training differed from speed training only in that stimulus presentation rate was constant.

The final phase consisted of FCAT speed training, also administered four times per week, employing both accuracy feedback and adjustment of rate of stimulus presentation as a function of performance accuracy. This phase was identical to the speed-shaping administered prior subjects. The brevity of the final phase was a result of discharge from therapy.

The three-week hiatus following initial SCAT administration was due to the fact that, at time of admission to the Day Hospital, the subject was severely disoriented and in considerable pain due to orthopedic injuries. Participation was therefore suspended pending the subject's emergence from posttraumatic amnesia and the easing of his pain.

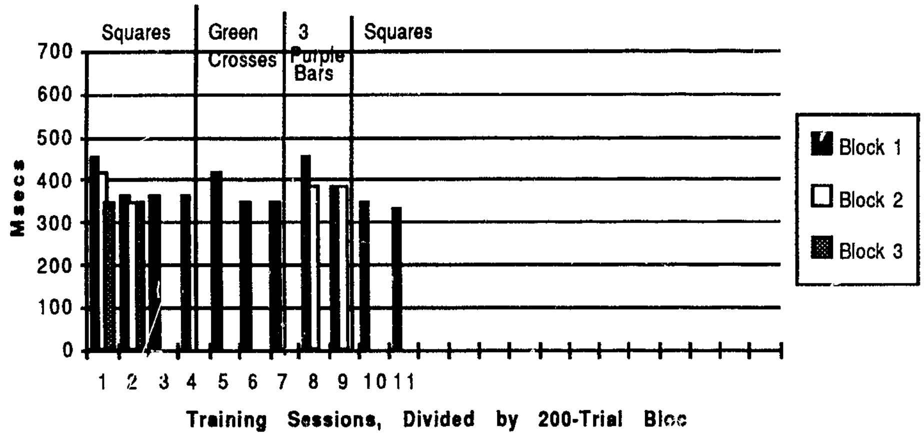
Finally, a note regarding a result that was frustratingly consistent across the first three subjects is in order. As had been the case for the previous two subjects, one session of ACAT administration was unsuccessful due to equipment failure.

Results

Training Task

As had been the case for both previous subjects, maximum rate of performance of the FCAT-like training task improved within and between sessions (Figure 3.3.0).

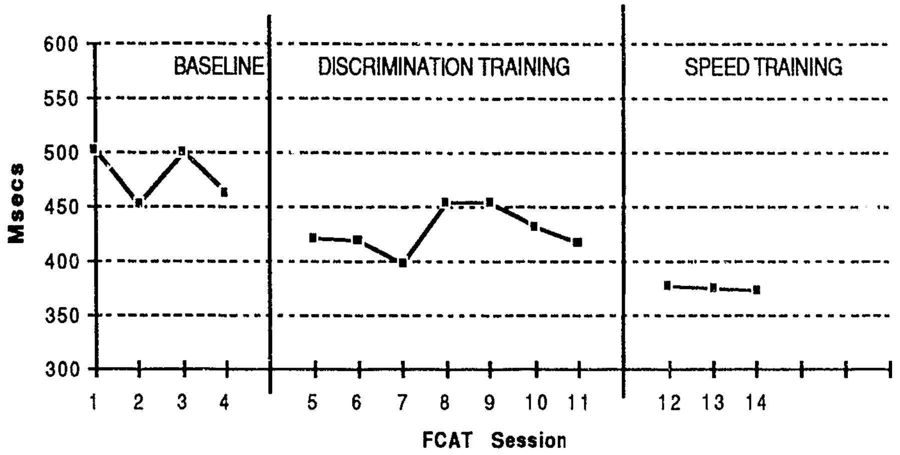
Figure 3.3.0: Subject 3's Minimum ISI Achieved in Training, Msecs.



FCAT

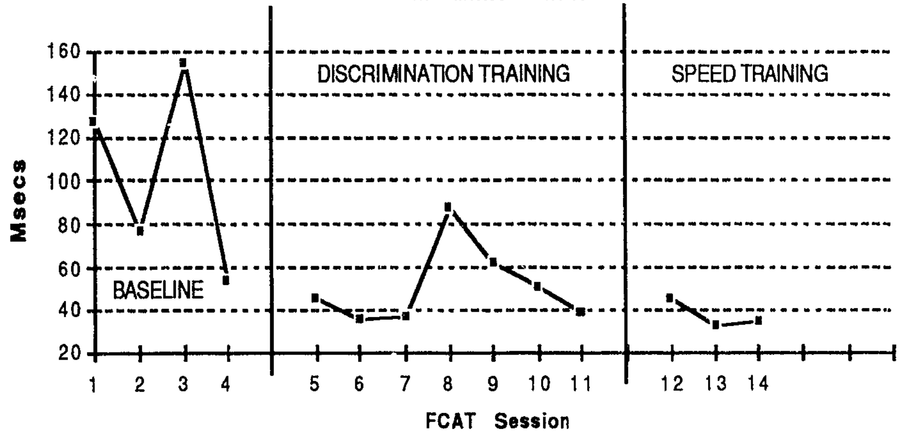
Mean reaction time (Figure 3.3.1) initially dropped with immediate feedback, but was inconsistent across sessions (though five of seven scores were in fact lower). However it fell even more and became very consistent across sessions when speed-shaping was introduced.

Figure 3.3.1: Subject 3's Mean FCAT Reaction Time, in Millisecon



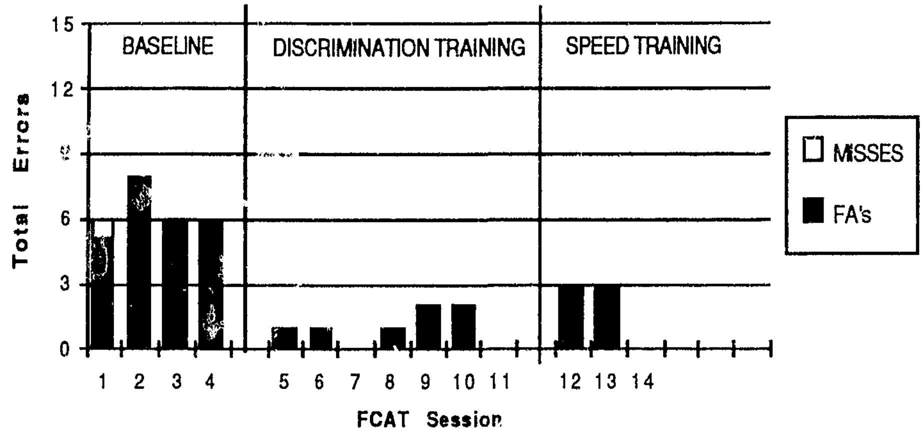
Standard deviation (Figure 3.3.2) diminished and became less variable over conditions.

Figure 3.3.2: Subject 3's Standard Deviation of FCAT Reaction Ti in Milliseconds.



Inspection of error rate (Figure 3.3.3) shows that false alarms dropped with immediate feedback, and were unaffected or rose slightly with speed-shaping.

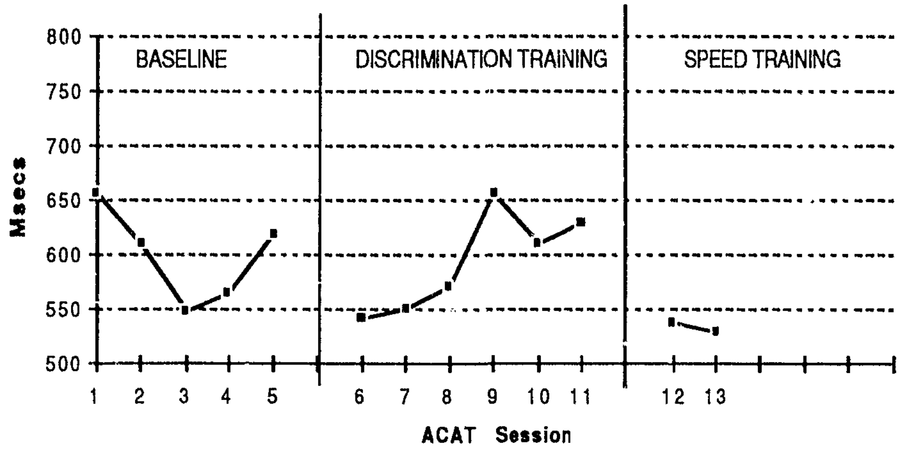
Figure 3.3.3: Subject 3's Total FCAT Errors, by Ty



ACAT

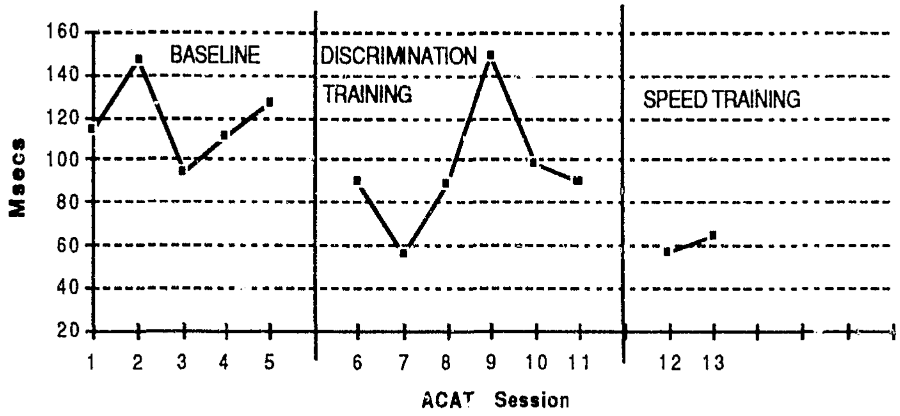
Mean reaction time (Figure 3.3.4) was largely unaffected by either treatment, although a treatment effect of speed training might have been apparent if the initial trend established across the two sessions of this condition would have persisted across additional sessions.

Figure 3.3.4 Subject 3's Mean ACAT Reaction Time, in Millisecon



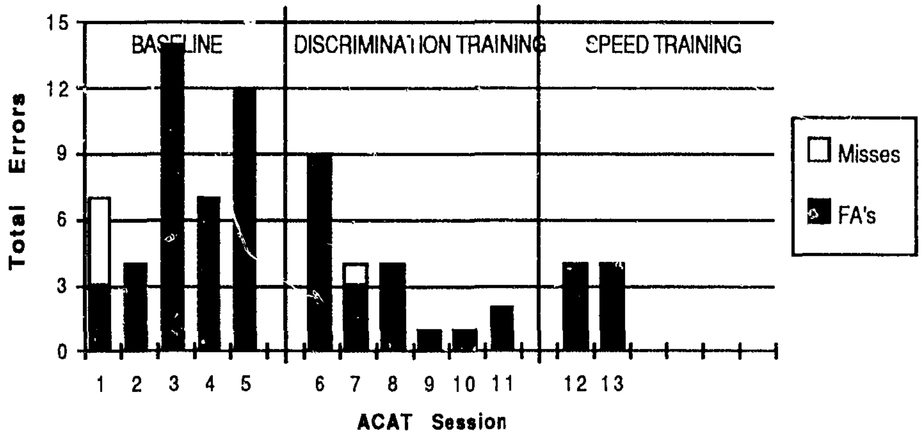
Standard deviation of reaction time (Figure 3.3.5) showed results very similar to those for mean reaction time, apparently unaffected by either treatment condition although the two sessions of speed training might have led to decrement had additional sessions been administered.

Figure 3.3.5: Subject 3's Standard Deviation of ACAT Reaction Ti in Milliseconds.



Errors (Figure 3.3.6), consisting chiefly of false alarms, showed some reduction following discrimination training.

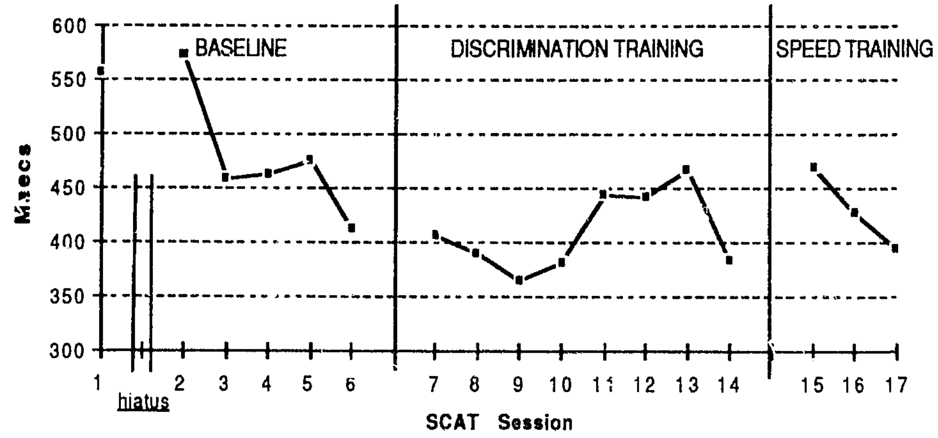
Figure 3.3.6: Subject 3's Total ACAT Errors, by Ty



SCAT

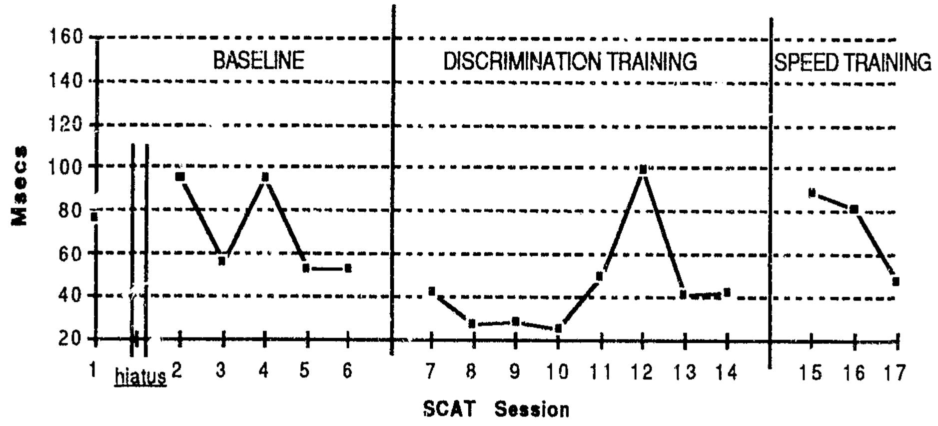
Mean (Figure 3.3.7) was not clearly affected by either treatment. The diminishing trend established during baseline continued into the Immediate Feedback (discrimination training) condition, but soon reversed itself. All sessions during Speed Training fell within the range established in previous conditions.

Figure 3.3.7: Subject 3's Mean SCAT Reaction Time, ir ilisecon



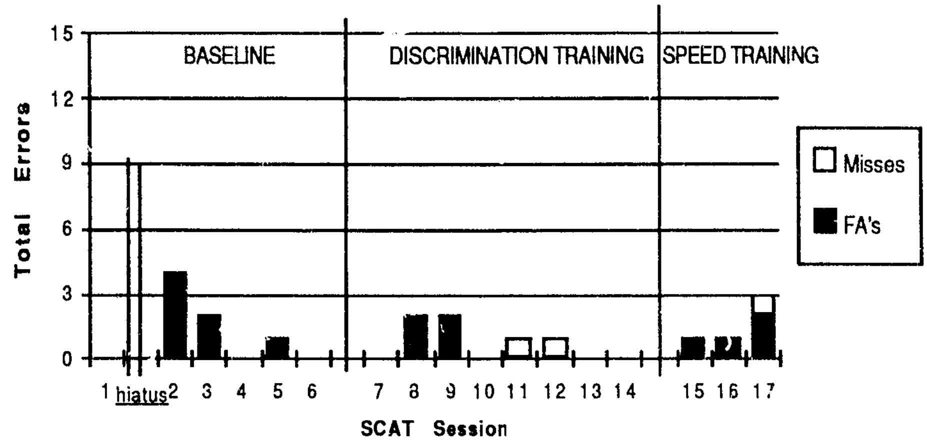
Again, standard deviation (Figure 3.3.8) mirrored mean, showing no conclusive evidence for treatment effect.

Figure 3.3.8: Subject 3's Standard Deviation of SCAT Reaction Ti in Milliseconds.



With regard to error rate (Figure 3.3.9), false alarms disappeared during the last half of discrimination training, with a slight resurgence during speed-shaping. With the disappearance of false alarms during discrimination training, the first misses occurred.

Figure 3.3.9: Subject 3's Total SCAT Errors, by Ty



Conclusions

Skill Acquisition

Discrimination training. With immediate feedback, performance of the FCAT-like training task improved in terms of decreased rate of false alarms. Reaction time findings were consistent but less pronounced, changing little over training sessions.

Speed Training. Adding the speed-shaping component to training, there was indication that the subject gradually learned to emit faster responses. As had been the case with prior subjects, this was evidenced by systematic improvement in minimum response time on the training task, both within and between sessions.

Generalization Across Tasks

Discrimination training. FCAT-like discrimination training produced improvement in FCAT mean and standard deviation of reaction time, though not consistently across sessions. FCAT false alarm frequency was consistently curtailed across sessions. ACAT and SCAT provided no evidence of generalization.

Speed Training. FCAT-like speed training produced further improvement in FCAT mean reaction time, and enhanced intersession consistency of this measure. No effect was noted on ACAT and SCAT.

Specific vs. Nonspecific Treatment Effects

Evidence suggested that effects obtained were specifically attributable to treatment.

Learning of the trained task. As had been noted with previous subjects, performance of the training task improved both within and across sessions.

Pattern of improvement on untrained tasks. Also noted in previous subjects, improvement on untrained tasks was limited to the task most similar to training.

Subject 4

Procedure

This subject received fewer SCAT sessions during baseline than called for by the standard protocol, due to equipment failure. This may have been the reason for failure to achieve a stabilized SCAT baseline prior to treatment. The decision to proceed with treatment rather than wait for SCAT stability was based upon concerns that delaying treatment would not permit sufficient post-treatment observation. This concern proved justified. In fact, following treatment only two sessions of each of the three vigilance tasks were administered prior to this subject's discharge from therapy.

Inability to administer SCAT during a large portion of baseline, producing an alteration in the standard sequence and frequency of task administration, was not seen as entirely negative. Failure to reliably replicate pilot results with respect to generalization to SCAT was suspected to be a result of overtesting, administering too many vigilance tasks too frequently (Seidel, personal communication, 1988). That is, exposure to several monotonous vigilance tasks on a frequent basis may have produced a deleterious effect on performance and on cross-task generalization. Pilot subjects, exposed only to a single vigilance task on a less-frequent basis, presumably would have been less affected by such a factor.

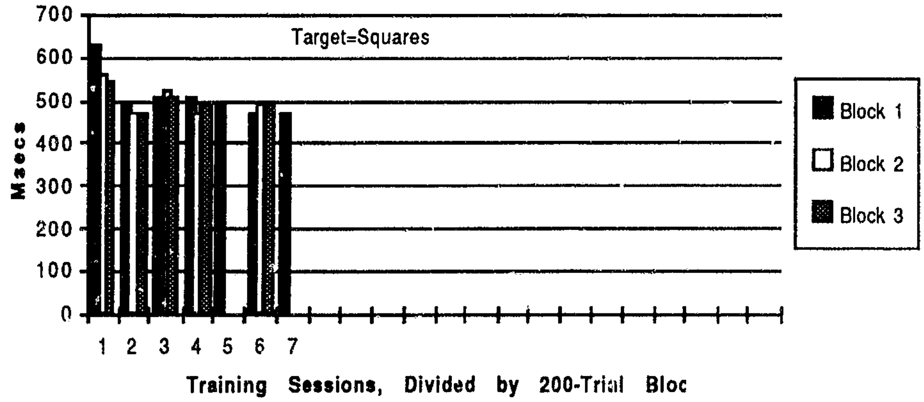
The unplanned variation in frequency of testing afforded an opportunity to observe the effect of such variation on task performance. Initially during baseline, two tasks were administered every day, followed by a period during which an average of one task was administered per day (0, 1, or 2). The latter frequency of generalization-task administration was maintained into the treatment phase, during which training was conducted four times per week.

Results

Training Task

Once again, obvious gains were evident in the FCAT-like training task (Figure 3.4.0). This parallels findings obtained from the preceding three subjects.

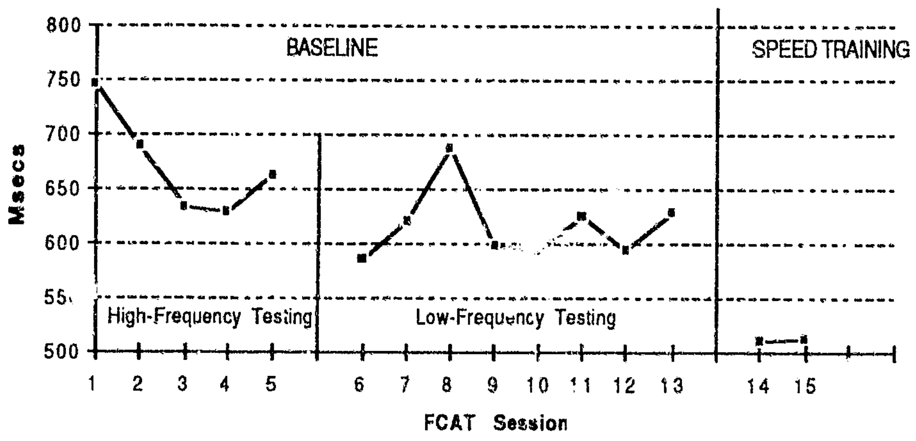
Figure 3.4.0: Subject 4's Minimum ISI Achieved in Training, Msecs.



FCAT

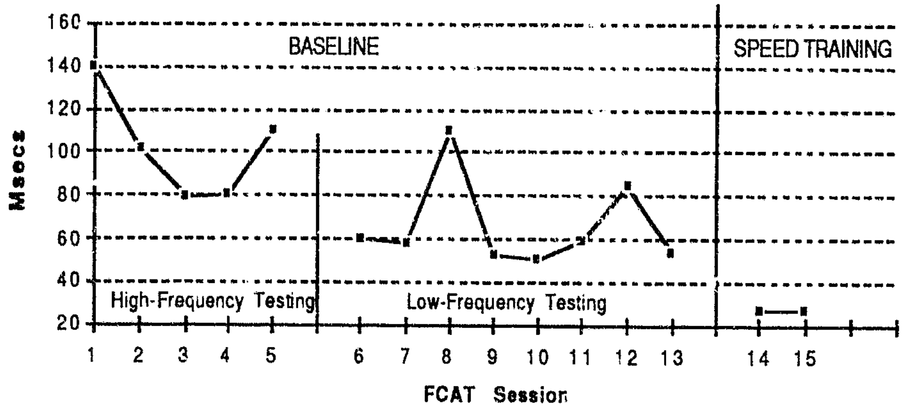
Inspection of mean reaction time (Figure 3.4.1) shows that varying frequency of testing had no effect on this measure. Speed-shaping had strong effect, mean reaction time dropping dramatically with training.

Figure 3.4.1: Subject 4's Mean FCAT Reaction Time, in Millisecon



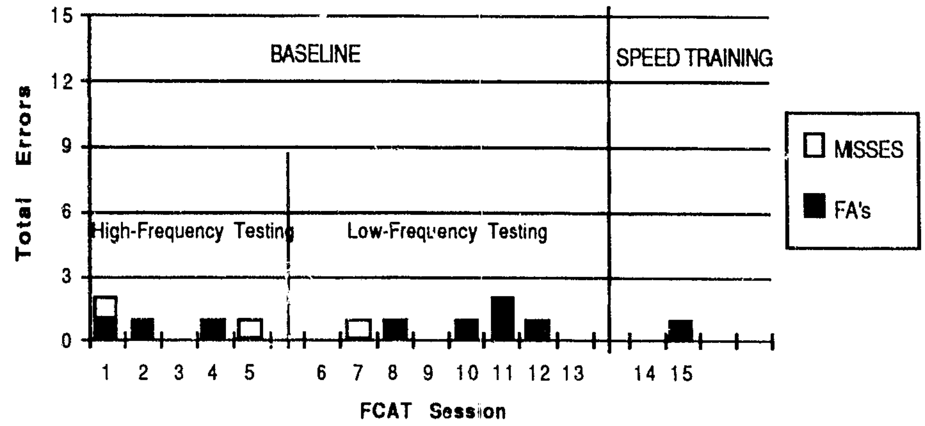
Standard deviation (Figure 3.4.2) showed an overall trend of diminution before treatment, with considerable intersession variability. Although there was a drop following treatment. this may have been a continuation of pretreatment trend.

Figure 3.4.2: Subject 4's Standard Deviation of FCAT Reaction Ti in Milliseconds.



Error Rate (Figure 3.4.3) was unaffected by all conditions.

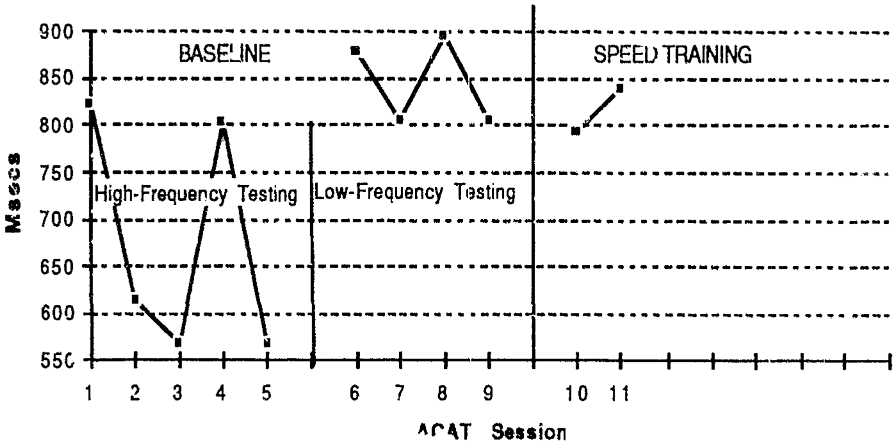
Figure 3.4.3: Subject 4's Total FCAT Errors, by Ty



ACAT

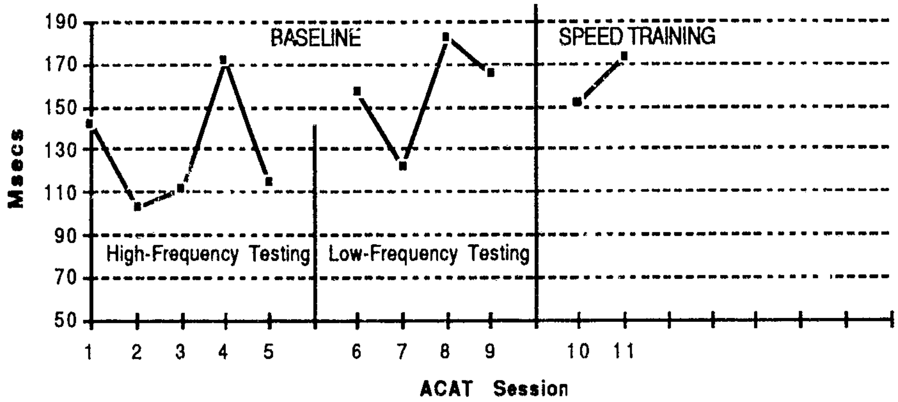
Mean ACAT reaction time (Figure 3.4.4) increased with decreased frequency of vigilance testing. No effect of Speed-shaping was noted.

Figure 3.4.4: Subject 4's Mean ACAT Reaction Time, in Millisecon



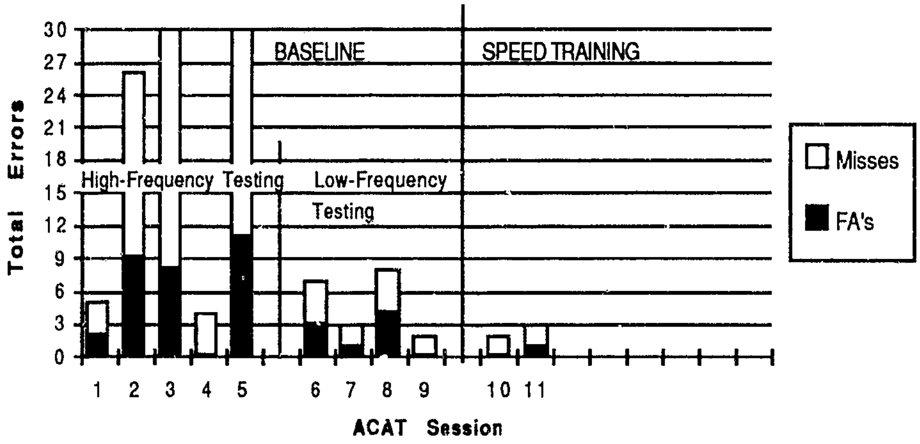
Standard deviation (Figure 3.4.5) was unaffected by all conditions.

Figure 3.4.5: Subject 4's Standard Deviation of ACAT Reaction Ti in Milliseconds.



Errors (Figure 3.4.6) diminished with decreased frequency of testing. Speed shaping produced no further effect. Gross changes in reaction time were inversely related to gross changes in error rate. This gross speed-accuracy tradeoff held throughout for ACAT. Overall performance efficiency, therefore, could be characterized as having remained relatively constant for this task.

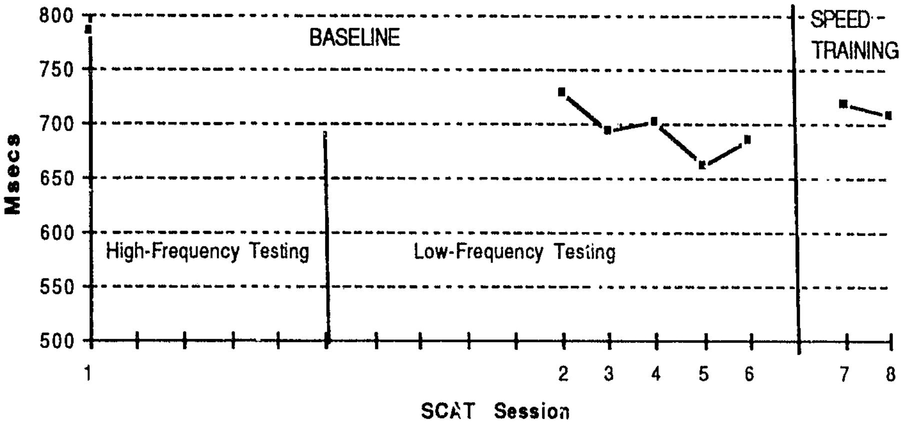
Figure 3.4.6: Subject 4's Total ACAT Errors, by Ty



SCAT

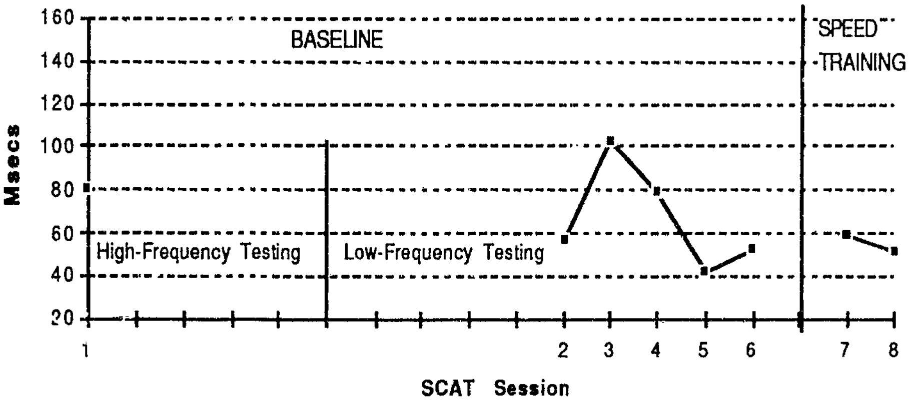
Mean (Figure 3.4.7) was unaffected by speed shaping. SCAT was not administered during the high-frequency testing condition, so the effect of test frequency on this measure could not be determined.

Figure 3.4.7: Subject 4's Mean SCAT Reaction Time, in Millisecon



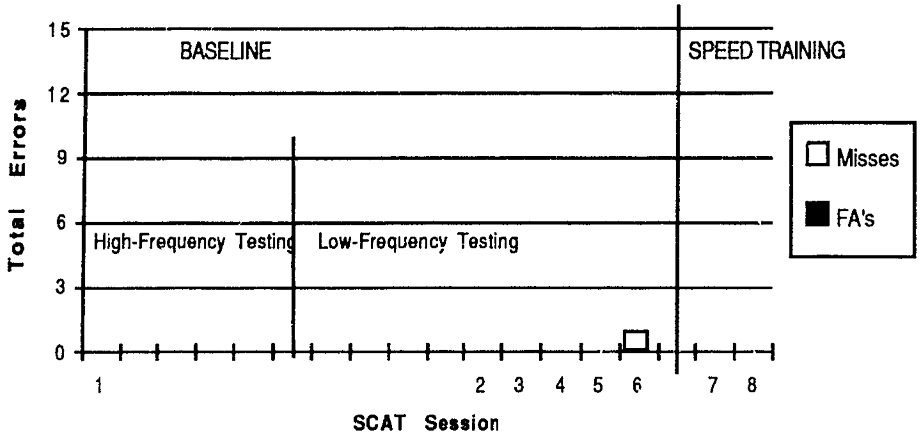
Standard deviation (Figure 3.4.8) was unaffected by treatment.

Figure 3.4.8: Subject 4's Standard Deviation of SCAT Reaction Tir in Milliseconds.



Error Rate (Figure 3.4.9) was consistently low, and unaffected by treatment.

Figure 3.4.9: Subject 4's Total SCAT Errors, by Ty



Conclusions

Skill Acquisition

As had been the case for all prior subjects, performance of the FCAT-like training task improved consistently across initial sessions. The presence of a learning curve indicated that training was indeed useful in improving rate of performance of that task.

Generalization Across Tasks

FCAT-like speed training markedly reduced mean and standard deviation of FCAT reaction time. This was not the case for ACAT and SCAT.

Specific vs. Nonspecific Treatment Effects

Evidence supporting the conclusion that effects were specifically attributable to treatment remained consistent with that of previous subjects: A clear learning curve was evident on the training task, and only the generalization task most similar to training showed improvement.

Discussion of Results, Subjects 1 - 4

After only four subjects, a pattern of results had emerged. Despite consistency of findings across subjects, however, there were some ambiguities in interpretation. The procedure for subsequent subjects was therefore altered to provide clarification.

Before describing results obtained from additional subjects, it will be instructive to summarize results obtained from the first four. This will allow tentative conclusions regarding the major questions stated at the outset, will identify alternative explanations of results thus far obtained, and will provide preliminary rationale for procedural variants intended to remove these ambiguities.

First we will examine basic skill acquisition, the extent to which training improved performance of the treatment task itself. Then we will consider the issue of generalization, the effect of treatment on untrained tasks. Within our description of both skill acquisition and generalization, we will separately consider the effects of discrimination training and speed training. Bear in mind that the current study was primarily intended to document the effects of speed training, a treatment consisting of both immediate feedback and gradual alteration of stimulus presentation rate. Consequently, data permitting consideration of the effects of immediate feedback isolated from speed training are limited. In the interest of identifying critical components of treatment, however, separate analysis will be attempted.

Skill Acquisition

Examination of performance on the training task, within and between sessions, indicated whether treatment was at all successful. Failure of subjects to improve on this task would have shown treatment to be wholly ineffectual, and questions regarding generalization would have been moot. As will be seen, however, this proved not to be the case.

We will first outline limited conclusions regarding the effect of immediate feedback (discrimination training) on treatment-task performance, then will summarize results obtained with combined immediate feedback and speed-shaping (speed training).

Discrimination Training

Of four subjects, only the second and third were provided discrimination training separate from speed training. Furthermore, because Subject 2 received only one block of discrimination training, it was not possible to demonstrate a learning curve for this participant. Thus only Subject 3 provided data relevant to determining the effect of discrimination training apart from speed training. For this subject, immediate feedback on a task with fixed rate of stimulus presentation appeared to reduce error rate, but not response latency.

Speed Training

A consistent finding was one of improvement on the trained task, reflected in subjects' attaining faster interstimulus intervals both within and between sessions. Interstimulus interval, it should be recalled, gradually speeded and slowed depending upon performance accuracy.

Effects of speed training on error rate could not be determined independently of effects on response latency, since error rate and speed were interrelated in the speed-training task. That is, although maximum rate of stimulus presentation did increase with training, and this increase did represent improvement in performance, it was not possible to directly determine whether improvement was attributable to improved accuracy or speed of response since either could have produced faster rates of stimulus presentation. Had the task involved a forced-choice paradigm in which subjects were required to respond "yes" or "no" on each trial, depending upon presence of a target, the factor(s) underlying improved performance would have been clearer. This would have differentiated misses in the current training task into two types: failure to respond at all within the allowed time span, and incorrect responding "no" when a target was indeed present. The former could be considered "speed of response misses" and the latter "accuracy misses." Such a task would have permitted direct analysis, in terms of accuracy vs. speed, of the underlying behavior producing improvement in performance of the training tasks.

Indirect evidence was obtained from performance on FCAT, the task most similar to training. Although FCAT results will appropriately be considered in the following discussion of generalization, they also have some bearing on the present question. This task maintained a constant rate of stimulus presentation and thereby allowed independent analysis of speed and accuracy measures. In general it was noted that FCAT mean speed of response improved following speed training, but that error rate was minimally affected. Therefore increased speed of response was the chief factor underlying improvement on FCAT. Since this change was presumably a function of the changed contingencies in the training task, it seems reasonable to assume that improvements in training-task performance, like FCAT, consisted chiefly of enhanced speed of response.

Generalization Across Tasks

Given that training indeed produced improvement on the trained task, the next question was whether this improvement occurred on untrained tasks as well.

Discrimination Training

Immediate feedback produced limited generalization across tasks, as reflected in data from both subjects 2 and 3, the only two receiving discrimination training. Admittedly, limited inferences could be drawn from Subject 2 in the absence of a true pretreatment baseline. Nevertheless the fact that this subject had failed to perform FCAT prior to FCAT-like discrimination training, yet successfully performed FCAT subsequent to training, demonstrated generalization to the task most similar to training.

Data from Subject 3 corroborated that from Subject 2, more specifically showing improvement in both error rate and speed of response on the task most similar to training (FCAT). Tasks other than FCAT showed only transient improvement of response speed, or improvement of error rate at the expense of speed. (See Figures 3.3.4 and 3.3.6).

Speed Training

Shaping to briefer interstimulus intervals on the FCAT-like task produced shorter mean response latencies on FCAT, an observation not obtained on measures less similar to FCAT-like training.

Specific vs. Nonspecific Treatment Effects

The dual finding of a learning curve on the training task and generalization only to the most-similar task was consistent with an interpretation that training had indeed been effective in promoting task performance, but that the skill acquired showed limited generalization across tasks. It would have been difficult to contend that some factors other than treatment were responsible for the learning curve consistently obtained on the training task. With respect to the pattern of generalization across tasks, however, one could argue that factors unrelated to skill acquisition had been at work. These other factors would be considered nonspecific in nature, reflecting changes in tonic state rather than being related to learning the trained skill. Such factors could have produced results resembling generalization effects, or they could have served to suppress generalization that might otherwise have occurred. Let us now consider each of these possibilities, weighing them against the conclusion that the pattern of generalization was specifically a product of training.

Nonspecific Effects Producing Apparent Generalization

One could argue that effects of training were limited to the very task on which training was conducted, and that apparent generalization across tasks was merely the result of change in conditions rather than training *per se*. A basic assumption of this viewpoint would be that effects nonspecific with respect to the precise nature of change in condition were limited to tasks similar to the one associated with the change. Were this the case, nonspecific effects emerging from introduction of the FCAT-like training task could have produced change in FCAT without producing like change in ACAT and SCAT.

For example, it could be suggested that change in tonic state occurred with introduction of a new task that was very demanding, involving an accelerated rate of stimulus presentation. One might then posit that this change became a conditioned response to the stimuli involved in the new condition. Performance on a like task might therefore show improvement unrelated to skill acquisition. One means of determining whether this was the case would be to provide intervention consisting of delayed consequences with a high degree of significance to the subject. Presumedly this would constitute a generally arousing condition without introducing specific training (i.e., immediate reinforcement and shaping). Were nonspecific factors of the type mentioned above contributing to the current findings, one might expect introduction of such treatment to produce results similar to those thus far obtained.

Weighing against the likelihood that nonspecific factors were responsible for apparent training effects, there was some evidence to suggest that such factors in fact did influence tasks beyond those most similar to training. SCAT mean reaction time for Subject 2 and ACAT mean reaction time for Subject 3 both showed deterioration after reaching a minimum part-way through baseline. This was viewed as a likely effect of overtesting, a potential nonspecific factor. With introduction of treatment, both of these measures dropped back to minimum levels. Apparently change in condition offset the factor which had been producing decrement during latter baseline. The fact that this finding was isolated to two cases (separate tasks and separate subjects) which had shown deterioration during baseline suggests that treatment effects manifested as "improvement" on these dissimilar-to-training tasks was dependent on the occurrence of pretreatment deterioration, a finding difficult to interpret as a training effect. Therefore, changing condition by introducing FCAT-like training did, in two cases, influence even nonsimilar ACAT and SCAT tasks, but this was attributable to nonspecific effects not limited to the task most similar to training. Thus nonspecific effects may have occurred but were readily distinguishable from effects attributed here to learning, and were not limited to the generalization task most similar to training. This makes less plausible an interpretation suggesting that nonspecific factors could have produced results mimicking the generalization of learning to the task most similar to training.

Nonspecific Effects Suppressing Generalization

Another manner in which nonspecific factors could have produced the current results would be through suppression of generalization, perhaps through mechanisms similar to those producing the performance suppressing effect of overtesting mentioned above. This could account for failure to replicate generalization effects obtained in pilot research (see Appendix C), since these early subjects were followed only on SCAT and were exposed to less frequent testing.

Frequency of testing was varied for Subject 4, with no apparent effect on baseline of all three tasks. However, this variation of test frequency occurred over a narrow range which did not include that applied to pilot subjects (see Appendix C). Also this provided no information regarding effects of test frequency on generalization across tasks, since variation occurred only during baseline. Furthermore, the suggestion that FCAT training influenced SCAT performance, obtained from pilot subjects who had received less frequent testing, was consistent with the interpretation that overtesting suppresses generalization.

It therefore seemed that an attempt at direct replication of procedures applied to pilot subjects was in order, further exploring the possibility of obtaining generalization across tasks. Success in obtaining such results would lend credence to the notion that nonspecific effects (specifically, overtesting) had suppressed generalization.

General Conclusions, Subjects 1 - 4

Skill Acquisition

Speed training consistently produced improvement in performance of the trained task. Evidence obtained from the FCAT generalization task pointed to reduced response latency as the factor most likely underlying this improvement. Conversely, results obtained from the one subject receiving several sessions of discrimination training were consistent with the notion that such training improves accuracy rather than response latency on the trained task.

Generalization Across Tasks

Although the treatments employed, discrimination and speed training, were effective in improving FCAT-like task performance, limited generalization to other tasks obtained. Only FCAT, the task most similar to training, showed evidence of generalization.

Possible Contribution of Nonspecific Factors

Results thus far were consistent with the notion that speed training was effective in improving speed of task performance. However, influence of nonspecific effects on generalization could not be entirely discounted. It was therefore decided to modify the experimental protocol for subsequent subjects, allowing prospective exploration of potential nonspecific influences.

Subject 5

Procedure

The procedure for Subject 5 was modified in that SCAT was the only vigilance task administered repeatedly as a test of generalization. This was done in an attempt to directly replicate results of pilot research (see Appendix C).

Furthermore, a new condition was introduced in an attempt to explore the possible contribution of nonspecific factors to previously observed treatment effects. This entailed introduction of a change from baseline conditions, designed to increase the subject's general level of arousal or interest in the task at hand without introducing immediate feedback. After baseline the subject was asked to select, from a menu of positive reinforcers, one which would be made contingent on SCAT performance but delivered *en masse* after the end-of-session performance summary. As during baseline, feedback during this new condition would consist only of a performance summary presented at each session's end.

Monetary payoff was the reinforcer chosen, with the understanding that every correct detection during a five-minute period averaging less than 540 milliseconds would earn five cents, and that false positive responses would be debited five cents. Perfect performance (no misses and no false positives) would produce a bonus of fifty cents. These values were chosen so as to allow total potential earnings per session of $5.00, an amount that seemed significant to the subject judging from his reaction to the suggestion. As mentioned above, reinforcement was delayed in nature, provided only at each session's conclusion.

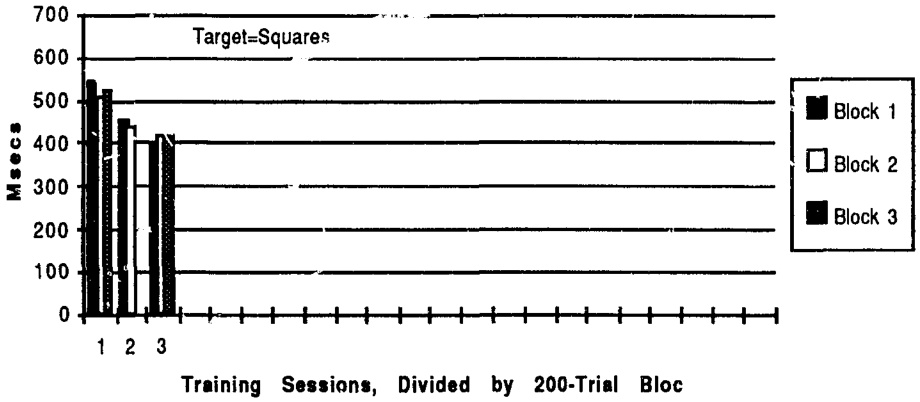
The delayed-reinforcement phase was followed by FCAT-like speed training. As was the case for baseline, the final condition was identical to that of previous subjects, except that SCAT was the only generalization task administered.

Results

Training Task

As has thus far been consistently demonstrated across subjects, performance of the FCAT-like training task improved within sessions and over sessions (Figure 3.5.0).

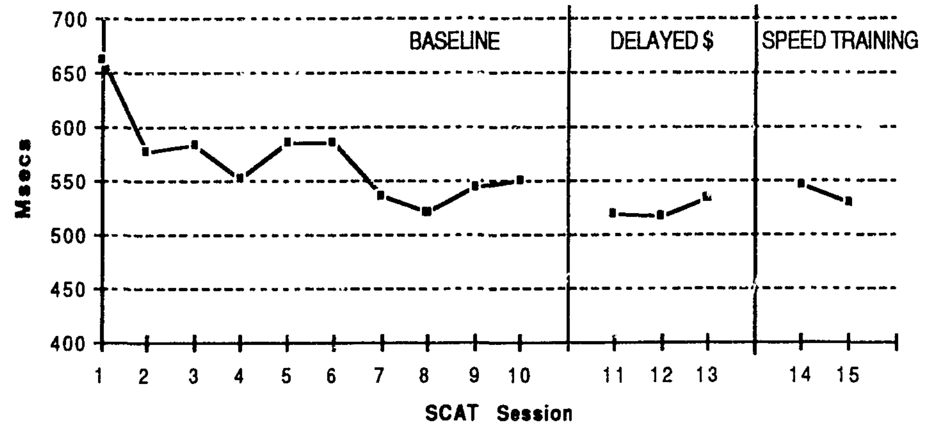
Figure 3.5.0: Subject 5's Minimum ISI Achieved in Training, Msecs.



SCAT

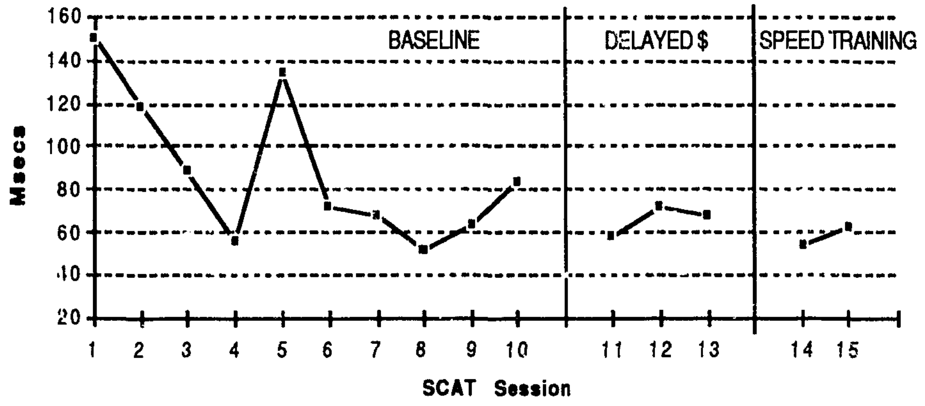
Mean reaction time (Figure 3.5.1) dropped during baseline, but had plateaued by the end of this condition. Neither treatment condition produced observable effects upon mean reaction time.

Figure 3.5.1: Subject 5's Mean SCAT Reaction Time, in Millisecon



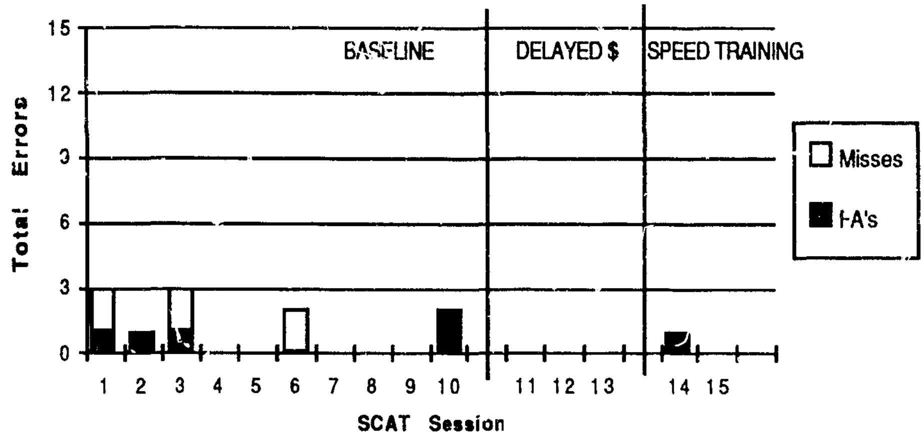
Standard deviation of reaction time (Figure 3.5.2) showed considerable intersession variability during baseline, but this seemed to diminish over sessions. No effect of either treatment was noted on this measure.

Figure 3.5.2: Subject 5's Standard Deviation of SCAT Reaction Ti
in Milliseconds.



Error rate (Figure 3.5.3) fell from initial baseline levels, improving some with practice. No errors were committed during the delayed monetary reinforcement condition, perhaps an effect of this treatment. Only one error, a false alarm, occurred during Speed-shaping.

Figure 3.5.3: Subject 5's Total SCAT Errors, by Ty



During the delayed reinforcement phase, Subject 5 showed perfect performance during all three sessions, earning $5.00 per session.

Conclusions

Skill Acquisition

Performance of FCAT-like training task improved within and between sessions.

Generalization Across Tasks

No effect of FCAT-like speed training was noted on SCAT.

Specific vs. Nonspecific Treatment Effects

Monetary reinforcement, a condition intended to maximize the influence of nonspecific factors without providing specific training, produced no effect on SCAT. Assuming that Subject 5 was not markedly different from subjects 1 - 4, and that the effects of FCAT-like training were therefore the same for all five subjects, this finding is consistent with the conclusion that previously noted treatment effects were specifically a product of training and not the result of more general motivational factors.

Unlike results obtained with pilot subjects (see Appendix C), no effect of FCAT-like training was noted on SCAT. Based upon the results of this single subject, it would be concluded that the simple influence of frequency of task administration was not producing the failure of subjects 1 - 4 to replicate pilot data. It was decided to gauge the reliability of results obtained with Subject 5 by applying the same procedure to subsequent subjects.

Subject 6

Procedure

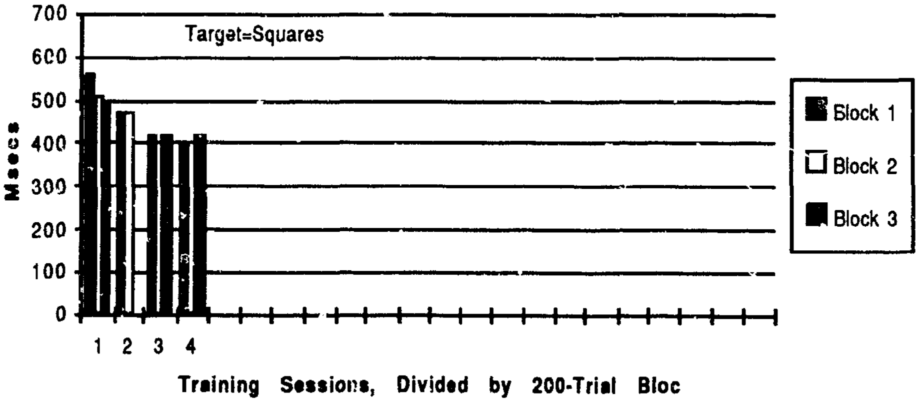
The procedure for Subject 6 was identical to that for Subject 5. As did the preceding subject, Subject 6 chose monetary reward.

Results

Training Task

Once again, performance of the FCAT-like speed shaping task showed improvement over sessions (Figure 3.6.0).

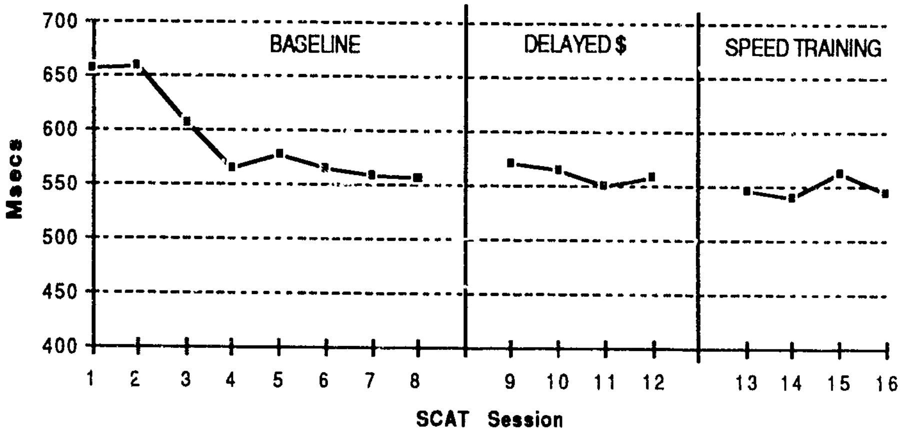
Figure 3.6.0: Subject 6's Minimum ISI Achieved in Training, Msecs.



SCAT

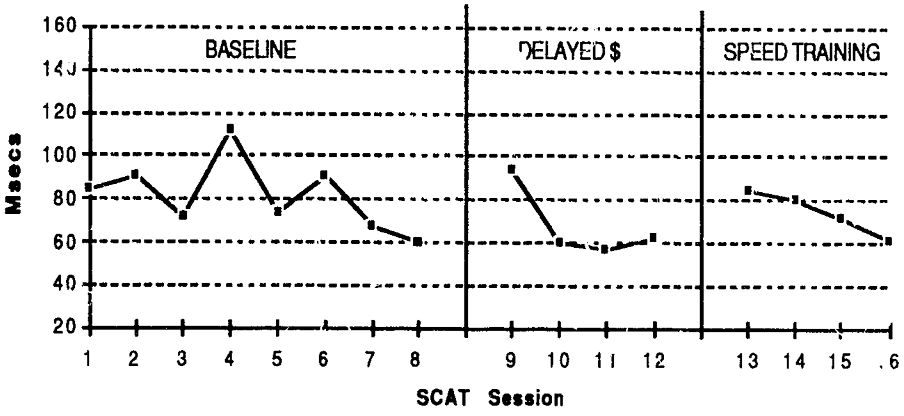
Mean reaction time (Figure 3.6.1) had levelled off by mid-baseline, and remained stable throughout the remainder of the study.

Figure 3.6.1: Subject 6's Mean SCAT Reaction Time, in Millisecon



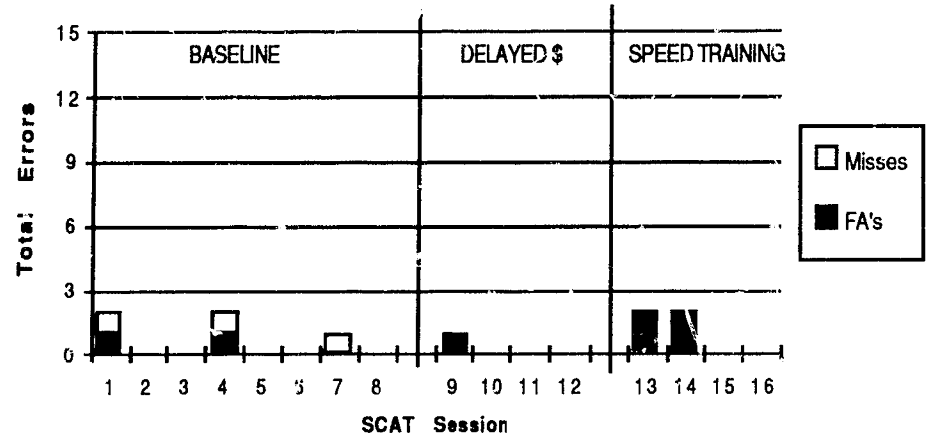
Standard deviation of reaction time (Figure 3.6.2) was variable across sessions, and presented no trends within or between conditions.

Figure 3.6.2: Subject 6's Standard Deviation of SCAT Reaction Ti in Milliseconds.



Errors (Figure 3.6.3) were sporadically evident throughout the period of study. After introduction of contingent feedback, there were no misses. With introduction of speed-shaping there was a slight rise in false alarms over the first two sessions of this condition.

Figure 3.6.3: Subject 6's Total SCAT Errors, by Ty



Conclusions

Results matched those obtained with Subject 5 very precisely. Improvement with practice during early baseline was noted, and no effect of contingent reinforcement nor of FCAT training was noted on SCAT performance. These results add credibility to those obtained with Subject 5; using the same procedure, the same results obtained for both subjects. The reason for inability to replicate pilot data (see Appendix C) remained obscure. Replication with one more subject would be attempted.

Subject 7

Procedure

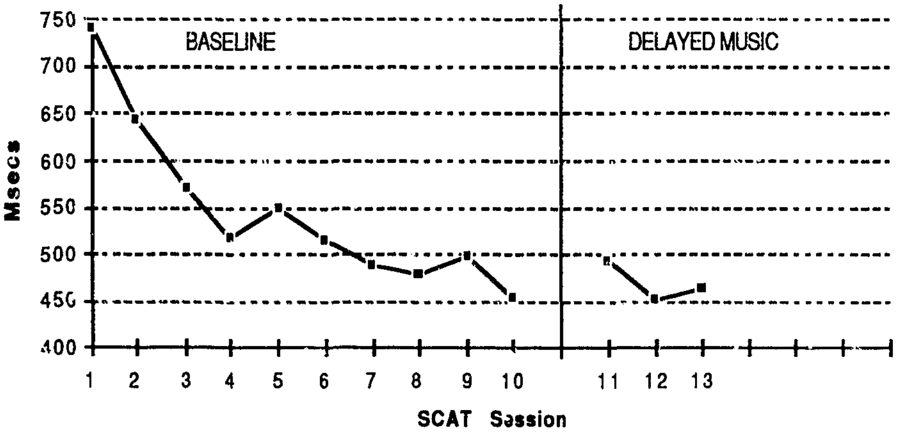
The procedure for Subject 7 matched that for subjects 5 and 6. However, this subject terminated participation prior to FCAT training. Therefore data was obtained only for baseline and the delayed contingent-reinforcement phase. Another difference was that Subject 7 chose access to music as a reinforcer. He was given permission to sit quietly after completing the SCAT session, headphones on, listening to the radio station of his choice for a period of time determined by SCAT performance. Using a formula similar to that used for monetary reinforcement, ten seconds of music was awarded for each correct detection during any five-minute block averaging less than 540 milliseconds per response. Ten seconds of music was deducted for each false positive. This provided a maximum of 15 minutes of music, to which a five-minute bonus was added for perfect performance.

Results

SCAT

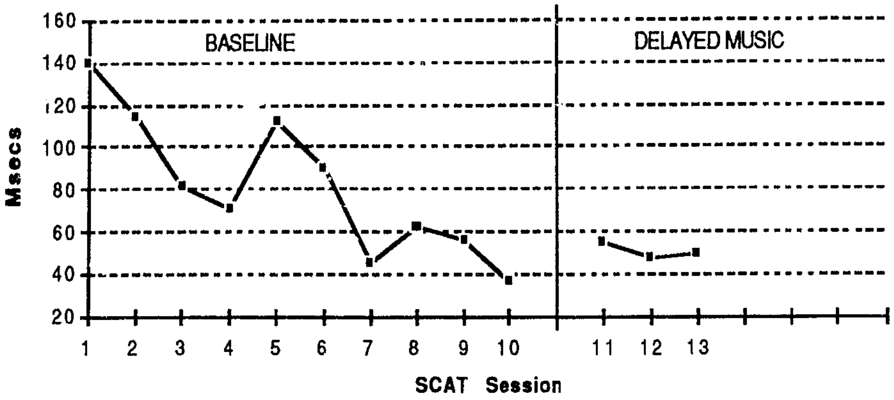
Mean reaction time (Figure 3.7.1) dropped precipitously over the course of baseline. No evidence of treatment effects obtained for this measure.

Figure 3.7.1: Subject 7's Mean SCAT Reaction Time, in Millisecon



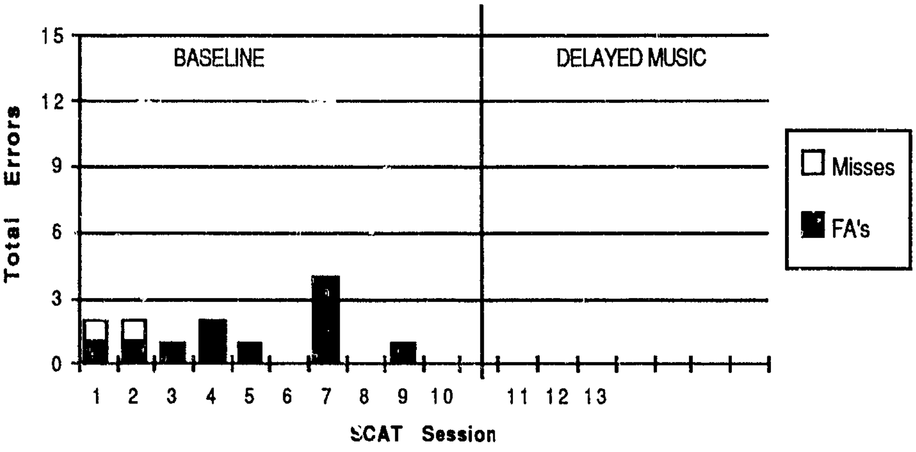
Standard deviation (Figure 3.7.2), like mean, progressively diminished throughout baseline. Again, no treatment effect was evident.

Figure 3.7.2: Subject 7's Standard Deviation of SCAT Reaction Ti in Milliseconds.



Frequency of errors (Figure 3.7.3) also fell during baseline, and because of this trend it was not clear whether absence of errors during Music Reward was attributable to treatment. However prior to treatment this subject had not completed two consecutive sessions without false alarm, suggesting a slight decrease in likelihood of false alarms. In all three treatment sessions, Subject 7 earned maximum music-listening time.

Figure 3.7.3: Subject 7's Total SCAT Errors, by Ty



Conclusions

In the absence of FCAT-like speed training, data obtained from this subject were limited in terms of adding to our understanding of the factors underlying previously observed behavior. However, given the absence of observed effects of delayed reinforcement on SCAT performance, we have seen another replication of the failure to obtain evidence supporting the notion that previous treatment results were a product of nonspecific factors.

Discussion of Results, Subjects 5 - 7

Having administered the revised procedure to three subjects, certain consistencies in results were apparent. As in the summary for subjects 1 - 4, we will first weigh the implications of the present findings with respect to basic skill acquisition, will then turn to consideration of the issue of generalization across tasks, and will finally summarize conclusions pertaining to the question of whether obtained effects were a product of specific or nonspecific treatment effects.

Skill Acquisition

Both subjects 5 and 6 exhibited a learning curve on the FCAT-like speed-training task . This replicated results obtained with the first four subjects. Since Subject 7 had terminated participation prior to FCAT-like training, there was no opportunity to collect from him data relevant to this question.

Generalization Across Tasks

The two subjects receiving FCAT-like speed training showed no effect of such training on SCAT performance. This, too, replicated findings obtained from subjects 1 - 4 but failed to replicate initial results obtained from pilot subjects (see Appendix C). The hypothesis that frequency of task administration was a factor in determining cross-task generalization was not supported by results from subjects 5 and 6. These subjects had experienced a frequency of task administration equivalent to that of the pilot subjects, but produced results equivalent to those of subjects 1 - 4.

Specific vs. Nonspecific Treatment Effects

The delayed-reinforcement condition, applied only to subjects 5 - 7, was intended to explore the possible contribution of nonspecific factors to previously observed treatment effects. All three subjects produced similar results, showing no change in SCAT performance when delayed reinforcement was made contingent on this performance. This fails to support an interpretation of previous results being attributable to nonspecifc factors, and may therefore be taken as added support to the notion that these results were actually a product of training.

The fact that subjects 5 - 7 showed rather dramatic diminution of mean SCAT reaction time over the course of baseline might produce concerns regarding a possible floor effect obviating any potential treatment effects. However, the fact that Subject 3 had obtained speeds substantially faster than those shown by the present subjects mitigates this concern.

General Conclusions, Subjects 5 - 7

At this juncture, evidence seemed to indicate that speed-training was effective in improving performance of the trained task and a nearly identical training task. However tasks differing from training in terms of stimuli employed showed no benefit from training. In order to confirm that this finding was not unique to the task used in training, it was decided to attempt replicating the present results with a different training task. Thus, for subsequent subjects, an ACAT-like task was used in training. The same generalization tasks, FCAT, ACAT, and SCAT, were used.

Subject 8

Procedure

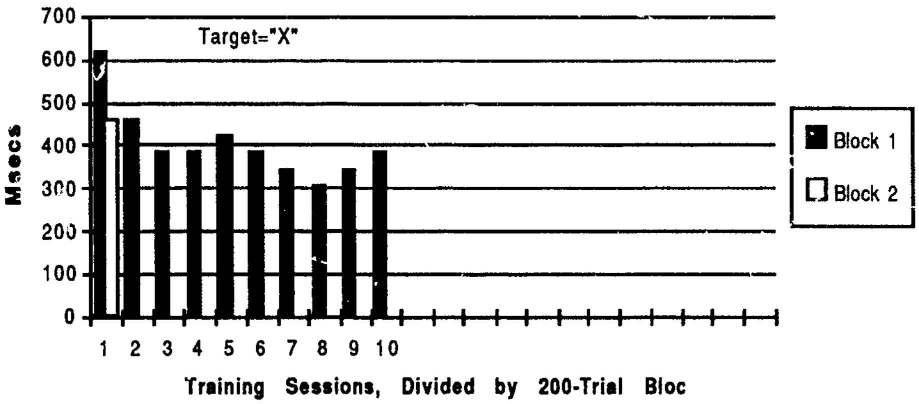
The procedure for Subject 8 was very similar to that for Subject 1, closely following the procedure originally described. The single exception was that Subject 8, instead of being trained with an FCAT-like speed-shaping task, was trained with an ACAT-like task which involved auditory stimuli but which was otherwise identical to the previous training task. Training was conducted twice per week and all sessions, save the first, involved only one block of trials. This was the case due to Subject 8's limited availability during treatment hours.

Results

Training Task

As has been the case with previous subjects receiving FCAT-like training, performance of the ACAT-like training task improved over sessions (Figure 3.8.0).

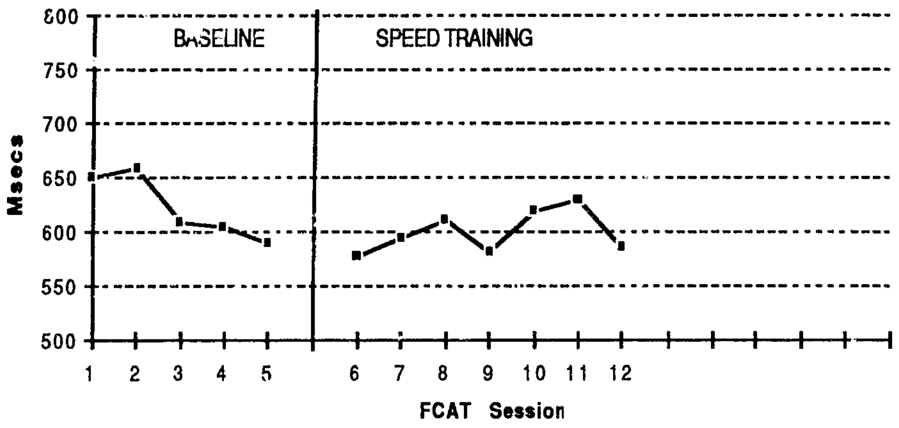
Figure 3.8.0: Subject 8's Minimum ISI Achieved in Training, Milliseconds.



FCAT

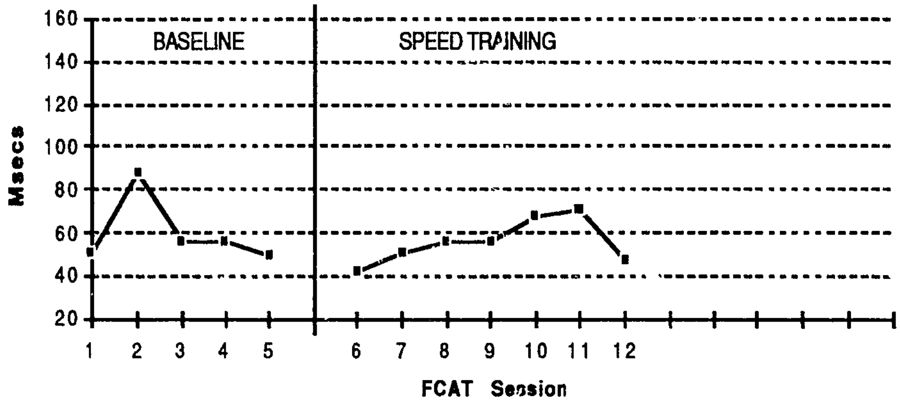
FCAT mean reaction time (Figure 3.8.1) did not appear to be influenced by ACAT-like speed shaping.

Figure 3.8.1: Subject 8's Mean FCAT Reaction Time, in Millisecon



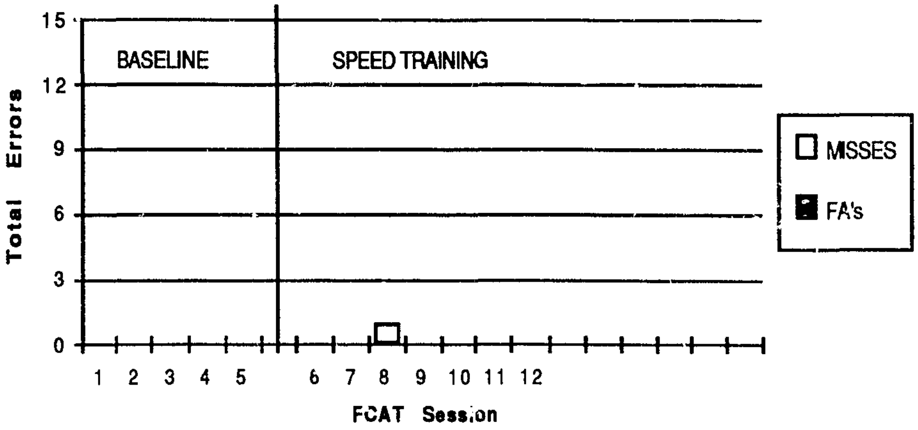
Neither was there observable effect upon standard deviation of reaction time (Figure 3.8.2).

Figure 3.8.2: Subject 8's Standard Deviation of FCAT Reaction Ti
in Milliseconds.



Error rate (Figure 3.8.3) was virtually zero throughout.

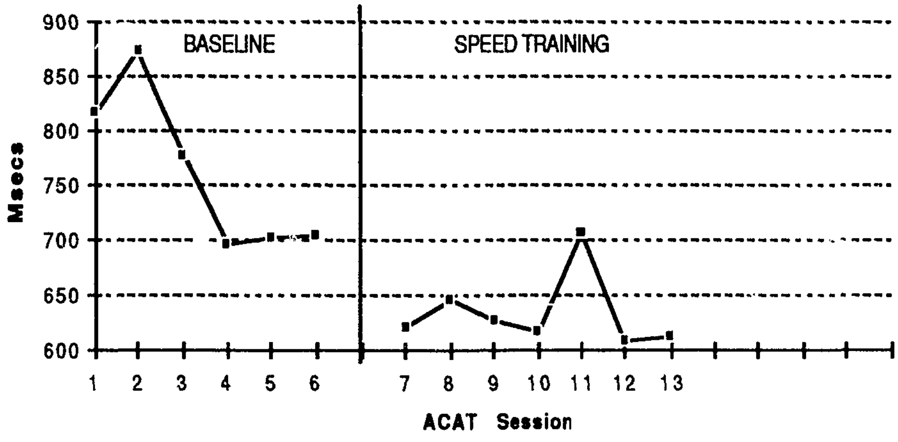
Figure 3.8.3: Subject 8's Total FCAT Errors, by Ty



ACAT

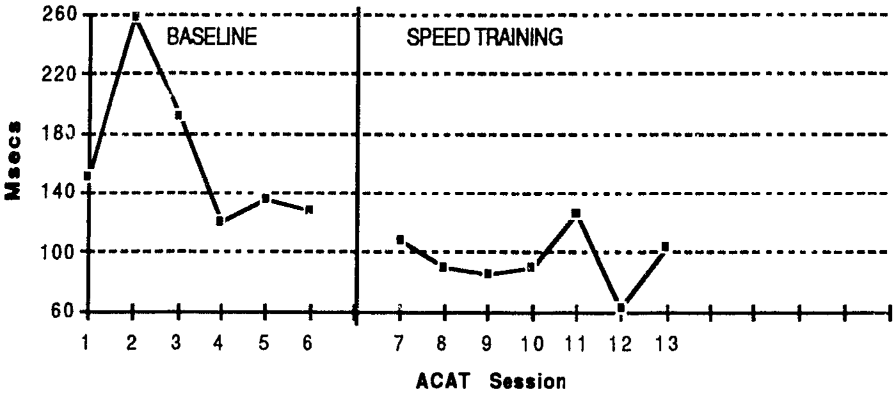
On the other hand, ACAT-like speed shaping coincided with lowering of ACAT mean reaction time (Figure 3.8.4). Of the seven post-treatment sessions, only one was at baseline levels.

Figure 3.8.4: Subject 8's Mean ACAT Reaction Time, in Millisecon



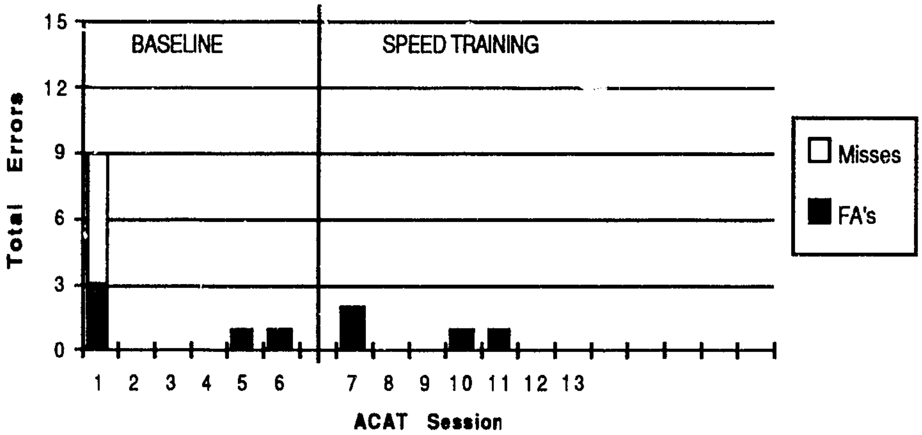
Standard deviation of reaction time (Figure 3.8.5) did not show clear effect of training.

Figure 3.8.5: Subject 8's Standard Deviation of ACAT Reaction Ti in Milliseconds.



Inspection of error rate (Figure 3.8.6) showed that several misses occurred in the first baseline session, but none thereafter. There was no apparent treatment effect.

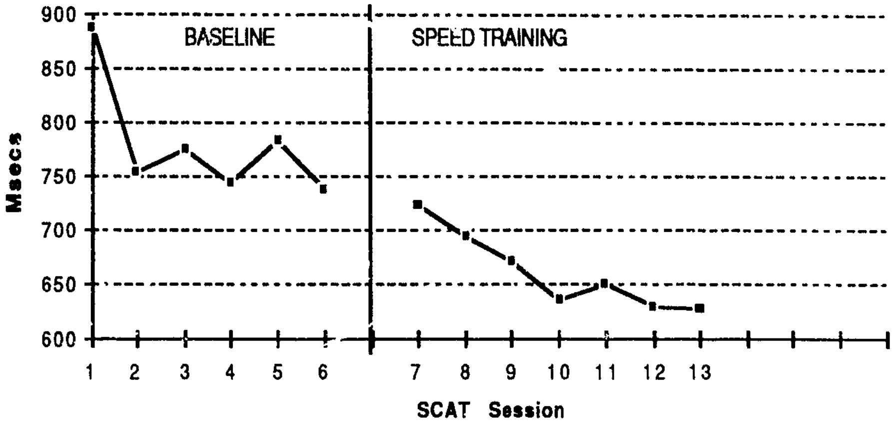
Figure 3.8.6: Subject 8's Total ACAT Errors, by Ty



SCAT

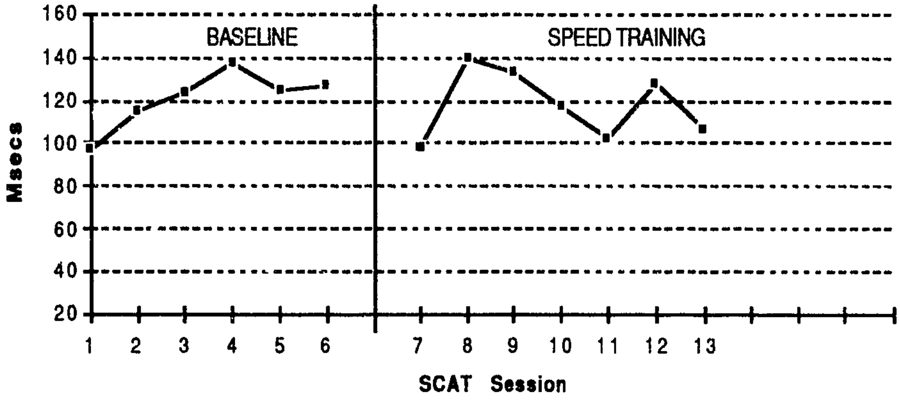
Mean SCAT reaction time (Figure 3.8.7) was stable through most of the latter part of baseline. With treatment, however, there was steady diminution across sessions.

Figure 3.8.7: Subject 8's Mean SCAT Reaction Time, in Millisecon



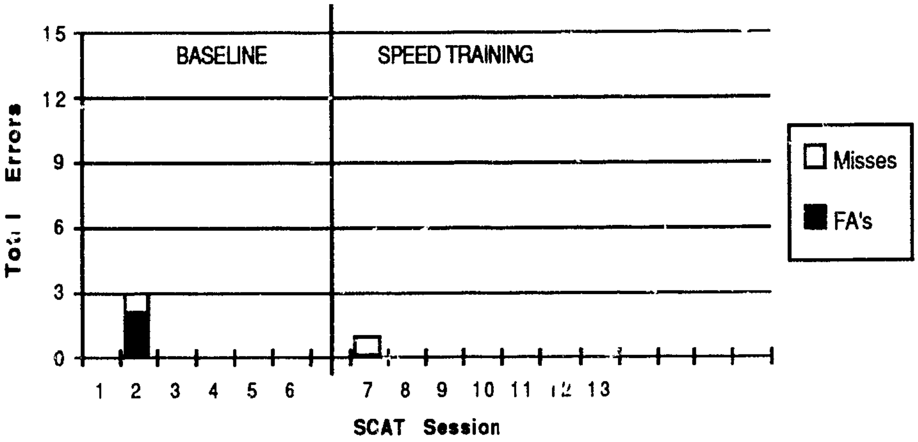
Standard deviation (Figure 3.8.8) was unchanged across phases, with no trend evident either within or between conditions.

Figure 3.8.8: Subject 8's Standard Deviation of SCAT Reaction Ti in Milliseconds.



Error Rate (Figure 3.8.9) did not show treatment effect, remaining very low throughout.

Figure 3.8.9: Subject 8's Total SCAT Errors, by Ty|



Conclusions

Skill Acquisition

As had been the case with FCAT-like training, ACAT-like training produced consistent improvement in performance of the training task, within and between sessions.

Generalization Across Tasks

ACAT-like speed training produced drop in ACAT mean reaction time. Further generalization was noted with respect to SCAT mean reaction time, but was not seen in FCAT performance.

Subject 9

Procedure

Because Subject 9 initially exhibited a very high FCAT error rate, a fact with significant emotional impact on this subject, FCAT-like discrimination training was administered after a brief baseline phase. Further complicating the procedure for this subject, the apparatus for delivery of auditory stimuli used in the ACAT task initially malfunctioned and was not as readily repaired as it had been for subjects 1 - 3. This resulted in loss of several ACAT sessions and failure to obtain an ACAT baseline prior to beginning the above mentioned FCAT-like discrimination training.

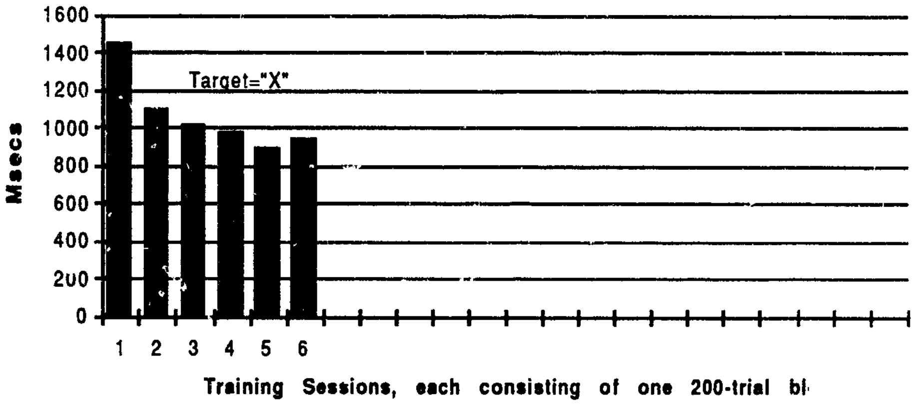
Thus, after a baseline period for FCAT and SCAT tasks, Subject 9 was administered FCAT-like discrimination training, at which time ACAT administration was also begun. The condition employing FCAT-like discrimination training was followed by one consisting of ACAT-like speed shaping. Training sessions were conducted twice per week.

Results

Training Task

As was the case for all prior subjects, improvement in training task performance was observed over sessions. Like Subject 8, this training consisted of ACAT-like speed shaping (Figure 3.9.0).

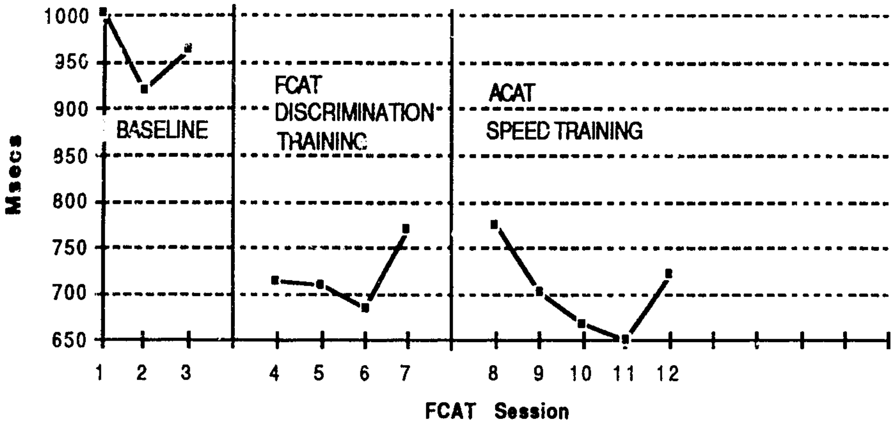
Figure 3.9.0: Subject 9's Minimum ISI Achieved in Training, Milliseconds.



FCAT

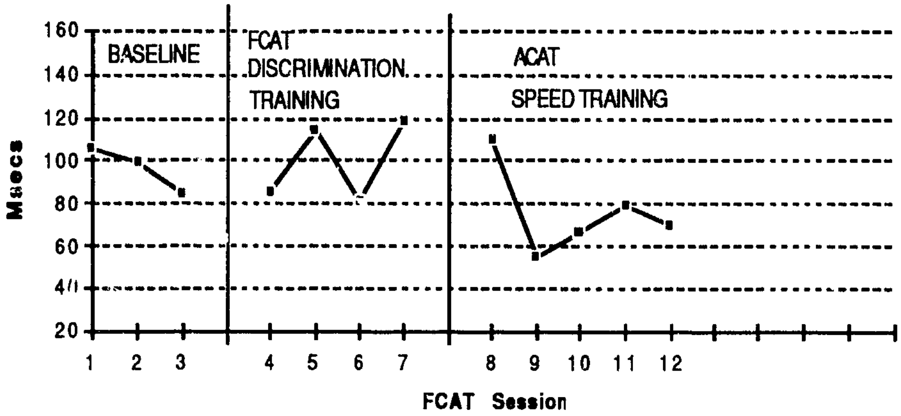
With FCAT-like discrimination training, a dramatic drop in mean reaction time obtained (Figure 3.9.1). Introduction of ACAT-like speed training exerted no further effect on this measure.

Figure 3.9.1: Subject 9's Mean FCAT Reaction Time, in Millisecon



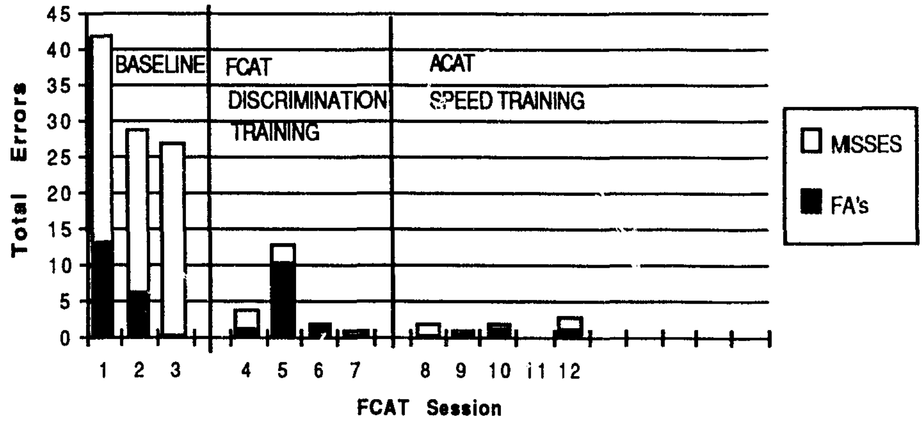
Standard deviation of reaction time (Figure 3.9.2) remained stable across baseline and discrimination training conditions. With ACAT-like speed shaping there was indication of possible slight, though inconsistent, drop in standard deviation

Figure 3.9.2: Subject 9's Stardard Deviation of FCAT Reaction Ti in Milliseconds.



FCAT-like discrimination training produced an abrupt drop in error rate (Figure 3.9.3), especially with respect to missed targets. ACAT-like speed training produced no additional effect on FCAT errors.

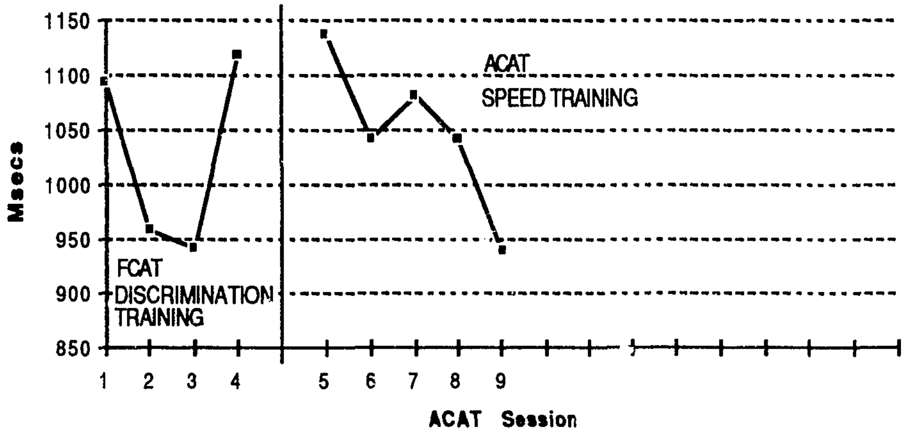
Figure 3.9.3: Subject 9's Total FCAT Errors, by Ty



ACAT

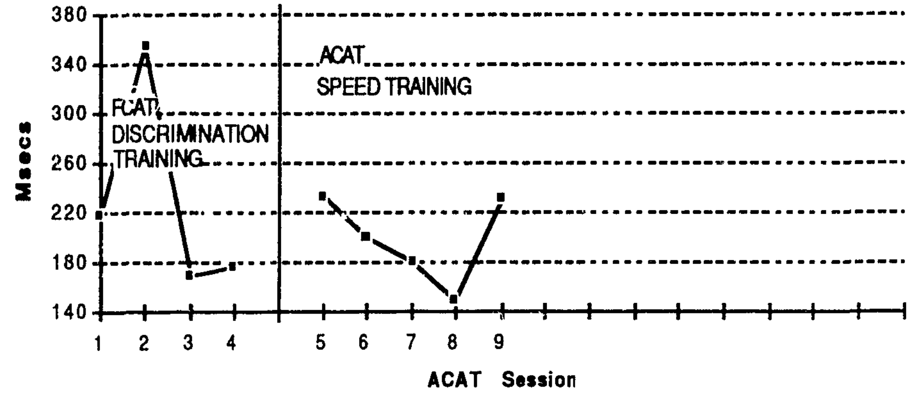
ACAT mean reaction time (Figure 3.9.4) was quite variable during FCAT-like discrimination training, with no trend apparent. Following ACAT-like speed training there was a clear downward trend, but all means were within limits already established in baseline.

Figure 3.9.4: Subject 9's Mean ACAT Reaction Time, in Millisecon



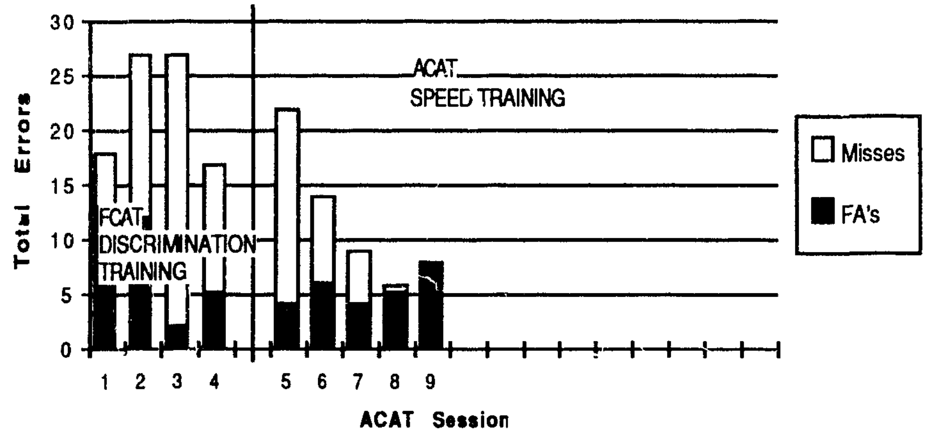
Standard Deviation of reaction time (Figure 3.9.5) produced a pattern very similar to that for mean. There was no clear effect of speed training.

Figure 3.9.5: Subject 9's Standard Deviation of ACAT Reaction Ti in Milliseconds.



Frequency of error (Figure 3.9.6) remained high through baseline, but showed diminishing trend during treatment. This was primarily attributable to reduction in misses following treatment.

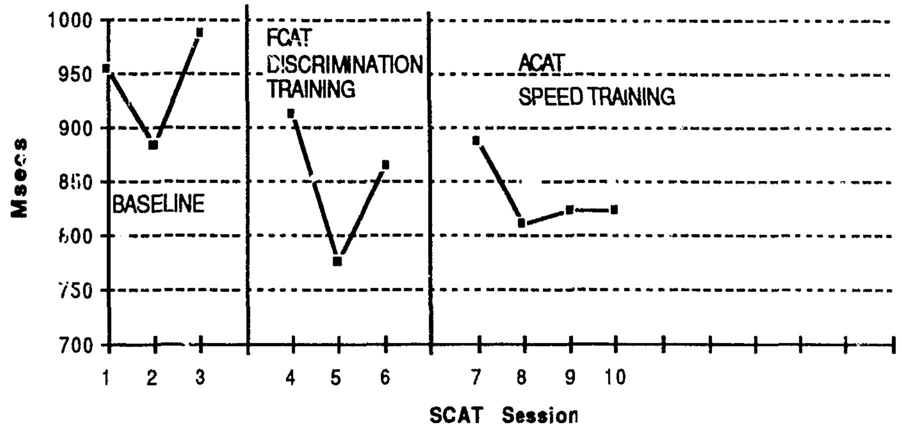
Figure 3.9.6: Subject 9's Total ACAT Errors, by Ty



SCAT

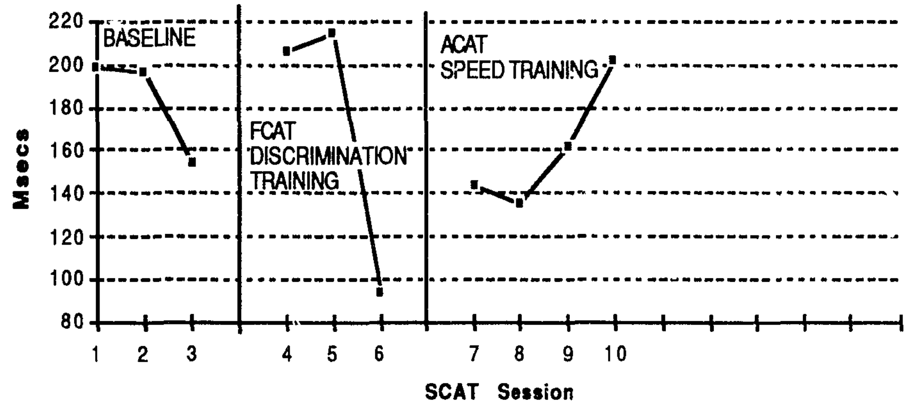
SCAT mean reaction time (Figure 3.9.7) showed slight evidence for diminution with FCAT-like discrimination training, though this impression was primarily attributable to one session. There was no effect of ACAT-like speed training on this measure.

Figure 3.9.7: Subject 9's Mean SCAT Reaction Time, in Millisecon



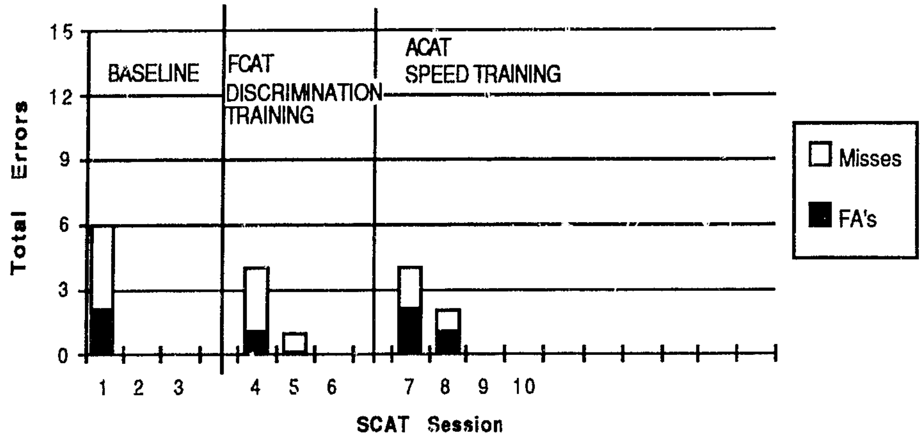
Standard deviation of SCAT reaction time (Figure 3.9.8) was extremely variable across sessions. The number of sessions in each condition was insufficient given the degree of variability witnessed. No effect of treatments was discernible.

Figure 3.9.8: Subject 9's Standard Deviation of SCAT Reaction Ti in Milliseconds.



Error rate (Figure 3.9.9) showed a brief resurgence with each change of condition. However, given the brevity of each of the three phases, reliability of this finding is difficult to gauge. In any event, the fact that error rate reached zero by the end of baseline precludes this as a measure sensitive to potential treatment effects.

Figure 3.9.9: Subject 9's Total SCAT Errors, by Ty



Conclusions

Skill Acquisition

As noted with all previous subjects, training-task performance improved within and across sessions. Not only did this result obtain with all subjects trained on the FCAT-like task, but held for the two subjects trained on the ACAT task as well.

Generalization Across Tasks

At first glance, one might question whether this subject even showed improvement in ACAT, the task most similar to training. A negative finding in this regard would be discrepant from observations of all previous subjects. However careful analysis, simultaneously considering mean reaction time and error rate, reveals improved ACAT performance following ACAT-like training. Although speeds achieved during ACAT-like training did not exceed those attained during the preceding condition, the tradeoff of accuracy for speed during the first phase was not witnessed during ACAT treatment. Thus, ACAT-like training appears to not only have produced lower ACAT error rate, but at speeds that had been associated with the highest error rates of the preceding condition.

SECTION 4: GENERAL DISCUSSION

We may now return to the two questions raised at the outset, first comparing the effectiveness of combined immediate feedback and speed shaping with that of delayed feedback, then turning to the question of whether trained improvements generalize across tasks.

Baseline: Practice Effects vs. Spontaneous Recovery

Baseline on all vigilance tasks generally showed improvement across sessions, with respect to both speed and accuracy. In principle this might have been a product of either practice or spontaneous recovery. In light of the fact that all subjects were in the early months of recovery, serious consideration was given the latter possibility. However, results obtained from subjects 3 and 4 weighed against such an interpretation. Both had received a single session of SCAT, followed by a hiatus of several days. In each case, relatively little change in SCAT performance occurred as a function of time. Instead, it seems that a necessary condition for the majority of improvement during baseline was repeated performance of the task in question.

Therefore improvement noted across baseline sessions seemed most likely to have been a consequence of practice. The substantial nature of this improvement, in the absence of immediate feedback, could be attributed to the fact that subjects were often aware of their own errors. One might consider these tasks to have involved a degree of inherent immediate feedback, which may have led to their being less sensitive to the effects of introducing explicit feedback than more complex or more ambiguous tasks in which subjects would have been less able to monitor their own performance during baseline.

Treatment Effects

Although the simplicity of the tasks could have resulted in a performance ceiling being reached during baseline, this proved not to be the case. While all three vigilance tasks generally plateaued during baseline, some additional improvements attributable to training were noted during the treatment phase. Not only did performance on the training task consistently improve within and between sessions, but the vigilance task employing the same stimuli as those used in training reliably improved across conditions, with benefits of training noted both in mean reaction time and error rate. We may therefore conclude that treatment was effective in improving speed and accuracy on the trained task, and that the baseline condition of presenting performance summaries at sessions' end was less effective than treatment which incorporated immediate feedback and speed-shaping components.

Furthermore, those few cases in which speed and accuracy training were implemented separately showed a differential effect of training-type on performance. Since immediate feedback was considered a *sine qua non* of training, an assumption supported by the current results, speed training included immediate feedback but added demand for speed by gradually altering interstimulus interval. This speed training primarily produced improvement in average speed of response, whereas immediate feedback alone primarily resulted in diminished error rate. Therefore it seems that type of training determined the nature of effects obtained.

Current results thereby support the argument that, in rehabilitative training, immediate feedback should be employed as much as possible. Potential negative effects of immediate feedback, such as emotional reactions or distraction by introduction of additional stimuli to the "natural" task, were not observed. In fact, given the self-evident nature of subject performance on the tasks lacking feedback, poor performance in some cases engendered emotional responses which were allayed by treatment sessions in which immediate feedback was provided. Thus, even with relatively simple tasks that might be considered likely to improve with practice alone, use of immediate feedback appears to produce optimum treatment outcome.

Results also indicated that effects of treatment were, for the most part, specific to the performance parameters emphasized in training. With the exception of results obtained from Subject 9, treatment focussing on accuracy minimized errors, while treatment focussing on speed minimized reaction time. Thus present results support the need to explicitly incorporate the demand for speed in training, when it is desirable to optimize this aspect of performance. Speed shaping, with gradual introduction of faster stimulus presentation based upon performance accuracy, appears to have been an effective approach to this end.

The discrepancy of Subject 9 from other subjects, showing improved mean response latency following discrimination training, led to a search for some unique characteristics of this subject which could account for the observed difference. The most prominent of these was the absolute level of performance attained on the vigilance speed measures. Baseline mean reaction time for subject 9 was in the 900 - 1100 millisecond range across tasks, showing minimum values greater than the maximum values of other subjects. Although data from a single subject should not be overinterpreted, a straightforward line of conjecture could hold that poor stimulus discrimination underlies severely slow reaction times, but does not play such a prominent role in limiting the reaction time of subjects exhibiting more commonly observed baseline levels of mean response latency. Thus, accuracy training in cases of very slow reaction time may produce increased speed without explicit speed training.

Generalization of Treatment

Clearly, treatment influenced performance of the vigilance task employing stimuli identical to those involved in training. For the most part, however, vigilance tasks employing other stimuli failed to show treatment effects. We may therefore infer that speed or accuracy of signal detection was not improved in a general sense, independently of the specific task involved in training. To the contrary, training involving discrimination of squares from other shapes produced improvement only in speed and accuracy of discerning these figures; there were no concomitant effects on detection of X's from among other letters, either visually or aurally. The only possible indication of more extensive generalization occurred with one subject that received training on auditory letter detection. Implications of this finding will be addressed below, in the discussion of Mapping Generalization.

Thus, benefits of treatment were mainly limited to the stimuli actually used in training. This has implications for the functional utility of vigilance training, as well as for application of the construct of sustained attention to rehabilitation.

Functional Utility of Training

In a strict sense, some degree of generalization was evident in the current findings. The training task was not identical to the vigilance task consisting of the same stimuli, since training involved immediate feedback and varying interstimulus interval. Therefore there was generalization from training to a situation in which the same stimuli were used but which was lacking explicit feedback. This corroborates results recently reported by Deacon and Campbell (1991), who found that head-injured subjects receiving immediate feedback and "time windows" (the latter imposing demand for rapid responding) showed faster reaction times which were maintained even when feedback and response windows were removed. Such findings are encouraging with respect to the potential functional utility of training, insofar as tasks in non-therapy settings are likely not to provide immediate feedback.

On the other hand, the limited nature of observed improvements leads one to question the practical utility of such training. Although training on a task employing feedback produced improvement on a no-feedback version of the same task, this improvement was not seen on tasks differing from training with respect to stimulus-type. Other versions, using the same format but different stimuli, showed no such benefit. This points to the extremely limited spontaneous generalization to be expected from such training, perhaps indicating that vigilance training must specifically employ the stimuli that will be encountered in the applied setting.

Current results could therefore be seen as weighing in favor of the contention that training should be provided only on tasks using the selfsame stimuli to be encountered in non-clinical settings. Differences between the natural task and training are unavoidable, by virtue of the fact that training will introduce elements not found elsewhere, but the present study indicates these differences do not prohibit generalization to the actual task. However more substantial (though still relatively minor) differences, as for example existed between FCAT and SCAT, may preclude generalization.

In the absence of specific evidence supporting an expectation of generalized improvement from training on esoteric tasks, it would seem prudent to instead use tasks which the subject will encounter in natural settings. There are some distinct advantages to such an approach. Not only does it reduce (though by no means eliminate) concerns regarding generalization to the applied setting, it also improves the face validity of rehabilitation tasks which in turn increases the likelihood of subject cooperation. Unfortunately there is a potential cost to explicitly training only a select range of functional tasks, in that such an approach could in fact hinder generalization. This will be discussed at greater length below, under Training Generalization.

Implications for the Construct of Sustained Attention

With respect to the question of whether "sustained attention" or "vigilance" are constructs referring to some broad response class which improves uniformly with training, current results suggest not. The fact that improved performance was limited to the task nearly identical to training showed that such improvement failed to spontaneously generalize across tasks, even those sharing many properties with the improved task. Thus, while performance of one task selected as a reasonable operational definition of sustained attention improved with training, other tasks, equally valid as vigilance measures, did not. There was therefore no justification for describing improvement in general terms of sustained attention. To do so would be misleading, ignoring the rather narrow context within which improvements were noted.

Knowledge of the contexts to which benefits of training are limited is of course extremely important to the rehabilitative process. Such constraints dictate the breadth of training necessary to achieve a given goal, and more accurately describe the skills imparted. For example, an occupational therapist working with a patient whose goal is to dress herself independently would need to know whether teaching buttoning and unbuttoning would generalize to zipping and unzipping. If so, training both would be redundant. If not, then describing improved buttoning as improvement in self-dressing would not only be imprecise, but could be misleading. Such improvements would best be described simply as improvements in buttoning, since other behaviors important to dressing, such as zipping, might otherwise be assumed to have improved along with buttoning and might therefore be overlooked in therapy.

This brings us to a general tenet of psychology, that the prediction and control of behavior is limited by our understanding of those circumstances under which the behavior in question occurs. In rehabilitation, psychometric evaluation is applied extensively to the goal of prediction, attempting to identify problematic behaviors and the circumstances in which they will occur. This is useful in rehabilitation planning only to the extent that such predictions are accurate. The constructs of attention are often a major focus of such evaluation, as they are perceived to have some value as shorthand descriptions of behavioral symptoms which tend to occur together and which reliably occur across a wide variety of tasks and situations.

However, the utility of attention as a construct in describing or predicting behavior is a separate matter from its utility in devising means of control. Behaviors which co-occur do not necessarily comprise a response class that uniformly responds to treatment. To the extent that they do not, a given construct loses its utility in describing improvements arising from therapy. Present results could be taken to imply that the construct of sustained attention has limited utility in this regard. That is, if training on one vigilance task does not generalize to other vigilance tasks, then there is no justification for describing imparted skills in broad terms of sustained attention. Rather, they would more accurately be described in terms of the specific task on which improvements were observed. For example, the current results would be best described as reflecting improved detection of squares, not as improved sustained attention. This would avoid misrepresentation of the skill actually acquired, would thereby avoid some potential pitfalls regarding the prediction of vigilance behavior in other settings, and would not be misleading with respect to planning of additional training tasks.

Thus, within constraints imposed by the methodology employed, current results call into question the appropriateness of applying the construct of sustained attention to the description of skills acquired in head injury rehabilitation. The finding that vigilance training produced benefits only on the trained task is consistent with the notion that improvements should be described in terms of the specific aspects of performance showing improvement (e.g., speed of response, error rate), and that the contexts to which improvements are limited (e.g., discrimination of squares from other figures but not discrimination of X's from other letters) should be spelled out as specifically as possible.

However, caution should be exercised in interpreting the current results. Although they might at first glance appear to suggest that vigilance training produces benefits essentially restricted to the trained task, such a conclusion would be premature. The present study did not employ existing methods of encouraging generalization, and should therefore be seen as only a first step in the methodical analysis of generalization patterns occurring with such training. This point will be elaborated below, as we consider training and mapping of generalization.

Training Generalization

As mentioned above, current results should be accepted cautiously, since training methods used did not incorporate strategies for promoting generalization. Stokes and Baer (1977) have outlined several techniques for doing so. Training of sufficient exemplars, an approach described as "the generalization-programming area most prominent and extensive in the present literature" of applied behavior analysis (p. 355), involves teaching several examples of the generalizable lesson, one after the other, until generalization to untrained examples occurs. In the present context, this could involve sequential training or a variety of vigilance tasks while monitoring performance of a set of vigilance tasks never used in training. This might provide useful information regarding the breadth of training necessary to produce more generalized improvement in vigilance.

In fact the reader may have noticed that the bulk of treatment effects occurred very soon after introduction of treatment. This could be taken as indicative that vigilance training should not be conducted extensively on a single task, but that time might be more productively spent briefly training on a wider variety of tasks. By administering several tasks, providing just enough training on each to reach a plateau in performance then moving on, greater generalization might be induced.

To the extent that such generalization could be trained, we would be doing a disservice to recipients of rehabilitation in making our training too specific. Patients who leave the rehabilitative setting trained to perform a very narrow range of tasks in a very narrow range of situations are extremely vulnerable to the vagaries of life which often impose change and demand flexibility. Were it possible to incorporate aspects of training which promote generalization, it might be possible to train broader skills that would more comprehensively prepare patients for varied tasks. That is, training which explicitly promotes generalization might produce improvements which could more reasonably be considered improvements in sustained attention.

However it seems safe to assume that there will always be limits to generalization of such training, and there will always be a need for qualifiers describing these limits. Even should greater generalization than that obtained here be possible, it is likely that the stimulus parameters and situational contexts defining the boundaries of generalization would need to be identified, to avoid false expectations regarding the effects of training.

Mapping Generalization

As indicated above, knowledge of the contextual limits of generalization is extremely important in terms of accurately describing skills taught, which has implications for guiding training toward functional goals. The current results, for example, show improvement in detecting squares after training with squares. Unless the terminal goal of therapy was to improve vigilance to squares, greater variety of training would be warranted.

An understanding of the stimulus parameters along which generalization can be expected would be very useful in terms of identifying the range of tasks needed to achieve a given degree of general improvement in vigilance-task performance. For example, current results indicate that training on a single visual vigilance task employing geometric shapes does not enhance vigilance for visual or auditory letters. Neither intramodal nor intermodal generalization obtained. On the other hand, one subject showed improvement in visual letter vigilance following training on auditory letter vigilance. This hints that "letterness" might be a crucial stimulus characteristic, with training generalizing across letter tasks regardless of modality. If so, then training designed to improve performance on as wide an array of vigilance tasks as possible might only need to employ training of letters in one modality, but would also have to employ training with other stimuli to achieve generalization beyond the domain of letters.

Current Results in Light of Previous Findings

As discussed above, current results do not support the assumption that the construct of sustained attention represents a trainable skill. This is consistent with conclusions drawn by Ponsford and Kinsella (1988), but is at odds with those of Sohlberg and Mateer (1987). This discrepancy can be reconciled by appealing to the potential effects of two factors: similarity of training to generalization tasks, and variety of tasks used in training. The more similar training is to generalization tasks, the more likely is generalization to be detected. By the same token, the more variety used in training, the more likely is generalization to occur.

Ponsford and Kinsella, failing to find substantial evidence for generalization from training to their dependent measures, may have employed tasks too different from training to be sensitive to generalization. They may also have failed to employ sufficient diversity in training to induce generalization, having only employed computerized tasks involving visual stimuli. Sohlberg and Mateer, obtaining positive results, may have chosen a dependent measure more similar to training or may have employed sufficient diversity in training to induce generalization. This latter possibility is especially intriguing, considering that one strength of the training approach used by Sohlberg and Mateer was its diversity. A great variety of tasks (visual and auditory) was presented in a number of contexts (computerized and not) by these authors. Additionally, it may be no coincidence that another study obtaining positive results, methodological shortcomings notwithstanding, was that of Ben-Yishay et al, who also incorporated diversity in training. The two studies entailing the most variety in training were also the two studies presenting the most compelling evidence in favor of the efficacy of training.

The present study represents a starting-point with respect to each of these factors: maximum similarity of training to testing, and minimum variety of training. This maximizes the current design's sensitivity to generalization, while affording an opportunity to observe the effects of training under conditions least favorable to generalization. Results point to the very narrow generalization obtained without incorporating methods for explicitly promoting generalization. Subsequent studies may now systematically introduce progressively greater variety in training to observe the effect on generalization of manipulating this variable.

Summary

In sum, the current study points to limitations in the application of the construct of sustained attention to brain injury rehabilitation. Acquired skills should be defined in specific terms of speed, accuracy, and consistency, and in terms of the situational or contextual constraints which determine whether the behavior in question will occur. Acquisition of the training task itself is optimized through provision of immediate feedback, and response speed is optimized by incorporating demand for speed into training. Results indicate that generalization is largely limited to the trained task when training does not explicitly attempt to train generalization. Future studies should systematically introduce greater variety in training to determine the degree to which generalization can be induced, and should include ranges of dependent measures varying in similarity to training that will permit identification of the limits of generalization.

BIBLIOGRAPHY

Baker, C.H. (1960). Observing behavior in a vigilance task. *Science, 132,* 674-675.

Ben-Yishay, Y., Diller, L., & Rattok, J. (1978). A modular approach to optimizing orientation, psychomotor alertness and purposive behavior in severe head trauma patients. In *Working approaches to remediation of cognitive deficits in brain damaged persons.* Rehabilitation Monograph No. 59 (pp. 63-67). Institute of Rehabilitation Medicine, New York University Medical Center.

Ben-Yishay, Y., Piasetsky, E.B., & Rattok, J. (1987). A systematic method for ameliorating disorders in basic attention. In *Neuropsychological Rehabilitation* . M. Meir, A. Benton, & L. Diller (Eds). New York: Guilford Press.

Ben-Yishay, Y., Rattok, J., & Diller, L. (1979). A clinical strategy for the systematic amelioration of attentional disturbances in severe head trauma patients. In *Working Approaches to Remediation of Cognitive Deficits in Brain Damaged Persons.* Y. Ben-Yishay (Ed.). NYU Medical Center, Rehabilitation Monograph No. 60, pp. 1-27.

Berch, D.B. & Kanter, D.R. (1984). Individual Differences. In *Sustained Attention in Human Performance.* J.S. Warm (Ed.). New York: Wiley & Sons.

Bond, M.R. (1976). Assessment of the psychosocial outcome of severe head injury. *Acta Neurochirurgica, 34,* 57-70.

Bond, M.R. (1979). The stages of recovery from severe head injury with special reference to late outcome. *International Rehabilitation Medicine, 1,* 155-159.

Brooks, D.N., Aughton, M.E., Bond, M.R., Jones, P., & Rizvi, S. (1980). Cognitive sequelae in relationship to early indices of severity of brain damage after severe blunt head injury. *Journal of Neurology, Neurosurgery, and Psychiatry, 43,* 529-534.

Brouwer, W.H. & van Wolffelaar, P.C. (1985). Sustained attention and sustained effort after closed head injury: Detection and 0.10 Hz heart rate variability in a low event rate vigilance task. *Cortex, 21,* 111-119.

Buchtel, H.A. (1987). Attention and vigilance after head trauma. In *Neurobehavioral Recovery From Head Injury.* H.S. Levin, J. Grafman, H.M. Eisenberg (Eds.). N.Y.: Oxford Univ. Press.

Buck, L. (1966). Reaction time as a measure of perceptual vigilance. *Psychological Bulletin, 65,* 291-308.

Carlsson, C.A., von Essen, C., & Lofgren, J. (1968). Factors affecting the clinical course of patients with severe head injuries: 1. Influence of biological factors. 2. Significance of posttraumatic coma. *Journal of Neurosurgery, 29,* 242-251.

Caveness, W.F. (1969). Posttraumatic sequelae. In *The Late Effects of Head Injury*. A.E. Walker, W.F. Caveness, M. Critchley (Eds.). Springfield, Ill.: Charles C. Thomas.

Davies, D.R. & Parasuraman, R. (1981). *The Psychology of Vigilance.* New York: Academic Press.

Deacon, D., & Campbell, K. (1991). Decision-making following closed-head injury: Can response speed be retrained? *Journal of Clinical and Experimental Neuropsychology, 13,* 639-651.

Dencker, S.J. & Lofving, B. (1958). A psychometric study of identical twins discordant for closed head injury. *Acta Psychiatrica Neurologica Scandinavica, 33 (Supplement 122),* 119-126.

Dikmen, S., Reitan, R., & Temkin, N. (1983). Neuropsychological Recovery in Head Injury. *Achives of Neurology, 40,* 333-338.

Ewing, R., McCarthy, D., Gronwall, D., & Wrightson, P. (1980). Persisting effects of minor head injury observable during hypoxic stress. *Journal of Clinical Neuropsychology, 2,* 147-155.

Faulkner, T.W. (1962). Variability of performance in a vigilance task. *Journal of Applied Psychology, 46,* 325-328.

Gianutsos, R., & Gianutsos, J. (1979). Rehabilitating the verbal information processing of brain injured patients: A demonstration using single-case methodology. *Journal of Clinical Neuropsychology, 1,* 117-133.

Gianutsos, R., & Klitzner, C. (1981). *Computer programs for cognitive rehabilitation* (Computer Programs). Bayport, N.Y.: Life Sciences Associates.

Golden, C.J. (1978). *The Stroop Color and Word Test: a manual for clinical and experimental uses.* Chicago: Stoelting Co.

Goyette, C.H., Conners, C.K., & Ulrich, R.F. (1978). Normative data on revised Conners Parent and Teacher Ratings Scales. *Journal of Abnormal Child Psychology, 6,* 221-236.

Gray, J.M. & Robertson, I. (1989). Remediation of attentional difficulties following brain injury: three experimental single case studies. *Brain Injury, 3,* 163-170.

Gronwall, D.M.A. (1977). Paced Auditory Serial-Addition Task: A measure of recovery from concussion. *Perceptual and Motor Skills, 44,* 367-373.

Gronwall, D.M.A. (1987). Advances in the assessment of attention and information processing after head injury. In *Neurobehavioral Recovery From Head Injury.* Harvey Levin, Jordan Grafman, Howard Eisenberg (Eds.). New York: Oxford University Press.

Gronwall, D.M.A. & Sampson, H. (1974). *The psychological effects of concussion.* Auckland: Auckland University Press.

Holland, J.G. (1958). Human Vigilance. *Science, 128,* 61-67.

Jerison, H.J. (1977). Vigilance: Biology, Psychology, Theory, and Practice. In *Vigilance: Theory, Operational Performance, and Physiological Correlates.* Robert D. Mackie (Ed.). New York: Plenum Press.

Jerison, H.J. & Wing, J.F. (1961). Human vigilance and operant behavior. *Science, 133,* 880-881.

Kaspar, J.C., Millichap, J.G., Backus, R., Child, D., & Schulman, J.L. (1971). A study of the relationship between neurological evidence of brain damage in children and activity and distractibility. *Journal of Consulting and Clinical Psychology, 36,* 329-337.

Klauber, M.R., Barrett-Conner, E., Marshall, L.F., & Bowers, S.A. (1981). The epidemiology of head injury: A prospective study of an entire community- San Diego County, California, 1978. *American Journal of Epidemiology, 113,* 500-509.

Lezak, M.D. (1979). Recovery of memory and learning functions following traumatic brain injury. *Cortex, 15,* 63-72.

Lezak, M.D. (1987). Assessment for rehabilitation planning. In *Neuropsychological Rehabilitation.* Manfred Meier, Arthur Benton, Leonard Diller (Eds.). New York: Guilford.

Lidvall, H.F., Linderoth, B., & Norlin, B. (1974). Causes of the post-concussional syndrome. *Acta Neurological Scandinavica, 50,* Supplement 56.

Malec, J., Jones, R., Rao, N., & Stubbs, K. (1984). Video game practice effects on sustained attention in patients with craniocerebral trauma. *Cognitive Rehabilitation, 2,* 18-23.

Mandleberg, I.A. (1976). Cognitive recovery after severe head injury 3. WAIS Verbal and Performance IQ's as a function of post-traumatic amnesia duration and time from injury. *Journal of Neurology, Neurosurgery, and Psychiatry, 39,* 1001-1007.

Mandleberg, I.A., & Brooks, D.N. (1975). Cognitive recovery after severe head injury 1. Serial testing on the Wechsler Adult Intelligence Scale. *Journal of Neurology, Neurosurgery, and Psychiatry, 38,* 1121-1126.

McCormack, P.D. (1962). A two-factor theory of vigilance. *British Journal of Psychology, 53,* 357-363.

McKinlay, W.W., Brooks, D.N., & Bond, M.R. (1983). Post-concussional symptoms, financial compensation, and outcome of severe blunt head injury. *Journal of Neurology, Neurosurgery, and Psychiatry, 46,* 1084-1091.

McLean, A., Dikmen, S., Temkin, N., Wyler, A.R., & Gale, J.R. (1984). Psychosocial functioning at 1 month after head injury. *Neurosurgery, 14,* 393-399.

Meldman, M.J. (1970). *Diseases of Attention and Perception.* New York: Pergammon Press.

Miller, E. (1970). Simple and choice reaction time following severe head injury. *Cortex, 6,* 121-127.

Miller, E. & Cruzat, A. (1981). A note on the effects of irrelevant information on task performance after mild and severe head injury. *British Journal of Social and Clinical Psychology, 20,* 69-70.

Mirsky, A.F., & Orren, M.M. (1977). Attention. In L.H. Miller (Ed.), *Neuropeptide influences on the brain and behavior.* (pp. 233-267). New York: Raven Press.

Mirsky, A.F., Primac, D.W., Ajmone-Marsan, C., Rosvold, H.E., & Stevens, J.R. (1960). A comparison of the psychological test performance of patients with focal and non-focal epilepsy. *Experimental Neurology, 2,* 75-89.

Naugle, R.I. (1990). Epidemiology of Traumatic Brain Injury in Adults. In *Traumatic Brain Injury.* E.D. Bigler (Ed.) (pp.69-103). Austin, Texas: Pro-Ed.

Newcombe, F. (1982). The psychological consequences of closed head injury: assessment and rehabilitation. *Injury: the British Journal of Accident Surgery, 14,* 111 - 136.

O'Brien, K.P., & Lezak, M.D. (1981). *Long-term improvements in intellectual function following brain injury.* Paper presented at the Fourth European Meeting of the International Neuropsychological Society. Bergen, Norway.

Ponsford, J.L. & Kinsella, G. (1988). Evaluation of a remedial programme for attentional deficits following closed-head injury. *Journal of Clinical and Experimental Neuropsychology, 10,* 693-708.

Posner, M.I. & Boies, S.J. (1971). Components of attention. *Psychological Review, 78(5),* 391 - 408.

Posner, M.I. & Rafal, (1987). Cognitive theories of attention and the rehabilitation of attentional deficits. In *Neuropsychological Rehabilitation.* M. Meir, A. Benton, & L. Diller (Eds). New York: Guilford Press.

Renner, K.E. (1964). Delay of reinforcement: A historical review. *Psychological Bulletin, 61,* 341-361.

Renner, K.E. (1968). Temporal integration: Amount of reward and relative utility of immediate and delayed outcomes. *Journal of Comparative and Physiological Psychology, 65,* 182-186.

Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome, E.D., & Beck, L.H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20,* 343-350.

Seidel, W., & Joschko, M. (1990). Evidence of difficulties in sustained attention in children with ADDH. *Journal of Abnormal Child Psychology, 18(2),* 217-229.

Seidel, W., & Joschko, M. (1991). Assessment of attention in children. *The Clinical Neuropsychologist, 5(1),* 53-66.

Sgro, J.A., Dyal, J.A., & Anastasio, E.J. (1967). Effects of constant delay of reinforcement on acquisition asymptote and resistance to extinction. *Journal of Experimental Psychology, 73,* 634-636.

Smith, A. (1968). The Symbol Digit Modalities Test: a neuropsychologic test for economic screening of learning and other cerebral disorders. *Learning Disorders, 3,* 83-91.

Smith, A. (1973). *Symbol Digit Modalities Test Manual.* Los Angeles: Western Psychological Services.

Sohlberg, M.M. & Mateer, C.A. (1987). Effectiveness of an attention-retraining program. *Journal of Clinical and Experimental Neuropsychology, 9(2),* 117 - 130.

Sohlberg, M.M. & Mateer, C.A. (1989). *Introduction to Cognitive Rehabilitation Theory and Practice.* N.Y.: Guilford.

Spreen, O., & Strauss, E. (1991). *A Compendium of Neuropsychological Tests.* New York: Oxford University Press.

Stokes, T.F. & Baer, D.M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10,* 379-367.

Tabaddor, K., Mattis, S., Zazula, T., & Phil, M. (1984). Cognitive sequelae and recovery course after moderate and severe head injury. *Neurosurgery, 14(6),* 701-708.

Van Doren, J. (1984). TIMESENSE: A BASIC time estimation task. *Cognitive Rehabilitation, 2,* 18 - 21.

Van Zomeran, A.H. (1981). *Reaction time and attention after closed head injury.* Lisse: Swets and Zeitlinger.

Van Zomeran, A.H., Brouwer, W.H., & Deelman, D.G. (1984). Attention deficits: the riddles of selectivity, speed, and alertness. In *Closed head injury: psychological, social and family consequences* D.N. Brooks (Ed.) (pp. 74-107). London: Oxford University Press.

Van Zomeran, A.H. & Van Denberg, W. (1985). Residual complaints of patients two years after severe head injury. *Journal of Neurology, Neurosurgery and Psychiatry, 48,* 21 - 28.

Wassertheil-Smoller, S., Tabaddor, K., Feiner, C. & Shulman, K. (1983). Factors affecting short-term outcome of head trauma patients. *Neuroepidemiology, 1,* 154-166.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale - Revised.* New York: Psychological Corporation.

Wechsler, D. (1945). A standarized memory scale for clinical use. *Journal of Psychology, 19,* 87-95.

Wist, E.R. (1962). Amount, delay, and position of delay of reinforcement as parameters of runway performance. *Journal of Experimental Psychology, 63,* 160-166.

Wood, R.L. (1984). The management of attention disorders following brain injury. In B.A. Wilson & N. Moffat (Eds.), *Clinical management of memory problems.* London: Croom Helm.

Wood, R.L. (1986). Rehabilitation of patients with disorders of attention. *Journal of head trauma rehabilitation, 1,* 43-53.

Woodcock, R., & Johnson, B. (1977). *Woodcock-Johnson Psychoeducational Battery.* Boston: Teaching Resources Corporation.

Wyckoff, L.B., Jr. (1952). The role of observing responses in discrimination learning. *Psychological Review, 59,* 431-442.

APPENDIX A

DEPENDENT MEASURES IN ATTENTION RETRAINING STUDIES

Symbol Digit Modalities (Smith, 1968, 1973) is a coding task similar to the Digit Symbol subtest of the Wechsler Adult Intelligence Scale - Revised. The subject is required to numerically encode a printed series of symbols, using a transcription key supplied at the top of the page. The task is presented in two versions: one requiring written, the other oral responses. Performance is gauged according to how many symbols are correctly translated within a 90-second period.

D2 Cancellation (see Spreen and Strauss, 1991) is a cancellation task consisting of rows of lowercase d's and p's on a printed page. Printed either above or below each letter is a number of dots. The subject is required to cross out each d that is associated with two dots, either two above, two below, or one above and one below. Meanwhile the examiner calls out "next line" with the passage of every 20 seconds, at which time the subject moves to the beginning of the next line on the page. This proceeds for a total of 14 lines. Measures derived from this test include Variability (greatest minus fewest number of letters completed on a line) as well as Total Correct responses and Total Errors.

Two-Letter Cancellation (Ben-Yishay et al, 1978) is a task in which the subject is presented with a sheet of rows of letters, and is required to mark every instance of the two identified target letters. Both time and errors are recorded.

Choice Reaction Time (Van Zomeran, 1981) presents subjects with a panel of five buttons. Four of these are arranged in an arc at the top of the panel and the fifth is placed at the center bottom, equidistant from each of the upper four. Each of the top four contains a light, and the task requires keeping the bottom button depressed, monitoring the display, and quickly pressing any top button that lights. Analysis of performance data is divided into "decision time" (delay between lighting and releasing the bottom button) and "movement time" (interval between release of bottom button and press of lit button).

Rating Scale of Attentional Behaviours (Ponsford and Kinsella, 1988) is a rating scale which requires rating the subject on a number of attentional behaviors such as "slowness, distractibility, and attention to detail." Each item is rated on a scale from 0 - 4, with 0 representing "not at all" and 4 representing "always."

Videotaped clerical duties (Ponsford and Kinsella, 1988) involves videotaping the subject during a 30-minute period of performing a "clerical task." The videotape is subsequently scored with regard to the percentage of time the subject's eyes are focused on the task at hand.

PASAT, *Paced Auditory Serial Addition Task* (Gronwall, 1977), consists of taped presentation of 61 single-digit numbers with an interstimulus interval of 2.4 seconds; subjects are required to add each number to the number that immediately precedes it, speaking the sum aloud.

Spatial Relations, a subtest of the Woodcock-Johnson Psychoeducational Battery (Woodcock and Johnson, 1977), consists of presentation of a series of geometric figures each accompanied by a set of irregular shapes, some of which can be combined to form the geometric figure; subjects are required to identify the correct shapes on each item and to complete as many items as possible in three minutes.

Attention-to-task behavior (Wood, 1986) is a time-sampling procedure involving record-keeping during standard rehabilitation therapy sessions. Every two minutes, the observer notes whether the subject is "maintaining head posture and directing gaze toward the therapist during the therapy session." At the end of the session a %-time attending summary score is derived from the collected data.

Digit Span is a subtest of the WAIS-R (Wechsler, 1981) which is thought to reflect attentional performance. The subject is required to listen to increasingly lengthy series of digits read aloud by the examiner, repeating each series immediately after hearing it. After failing to correctly repeat two series of the same length, the procedure is repeated with the subject required to repeat each series in reverse.

Logical Memory is a subtest of the Wechsler Memory Scale (Wechsler, 1945) in which subjects are required to listen to two short stories read by the examiner. At the conclusion of each story, the subject must repeat it as completely as possible.

Rey Auditory-Verbal Learning (see Spreen and Strauss, 1991) is a standard test of verbal learning in which the examiner reads a list of common nouns to the subject, who recalls as many as possible immediately after hearing the list. The list is then repeated another four times, with the subject recalling as many as possible after each reading. Then a recognition test is provided, a trial consisting of a second list of words administered, and finally the subject is asked to recall the words from the original list.

Seidel Computerized Attention Test (SCAT) (Seidel, 1988) is a vigilance task in which the subject monitors a computer display as a series of letters is centrally presented one at a time. Each time the letter "X" is displayed, the subject is to depress the space bar of the computer keyboard. The rest of the keyboard is masked to minimize distraction. Onset to onset interstimulus interval is 1.5 seconds, with a total of 600 trials presented over a 15-minute period. The task is then repeated, with targets defined as X's that have been preceded by an "A". For both tasks target density is held at 15% (30 targets per 5-minute block of 200 trials), with performance data such as hits, false alarms, average reaction time, and standard deviation of reaction time reported for each 5-minute block as well as for each 15-minute task and for the two tasks combined. After performance of the A-X task, these data are reviewed with the subject.

Auditory Computerized Attention Task (ACAT) was developed specifically for the current study. It is an auditory analog of the SCAT X-task, presenting letters in a robotic voice using a speech synthesizer but otherwise consisting of the same characteristics as the X task (e.g., ISI = 1.5 seconds, stimulus duration = 200 msec., target density = 15%, duration = 15 minutes).

Figural Computerized Attention Task (F-CAT) was also developed for the current study and is a figural analog of the SCAT X-task. This task presents geometric figures at the same rate as the two computerized tasks described above, with the same parameters such as target density and presentation rate. Two to four geometric figures are presented in each trial, and all are identical within each trial. Between trials these figures vary not only in shape and number, but also in size, color, and position on the screen. The subject is required to press a spacebar on each trial consisting of squares.

APPENDIX B

Training Tasks Used in Attention Retraining Studies

React is a computerized reaction time task developed by Gianutsos and Klitzner (1981). The subject is required to monitor a blank screen, and to press a key as quickly as possible whenever a stimulus appears. The stimulus appears at unpredictable screen locations, and consists of a rapidly incrementing counter which the subject stops by responding. After eight trials are presented to the two hemifields, the task terminates and a data summary for speed and accuracy within the two fields is presented, as well as overall scores.

Search is another Gianutsos and Klitzner (1981) computerized task, requiring the subject to search an array to find a figure which matches a target figure presented in the middle of the array. Again, the subject is instructed to do so as quickly as possible, and speed and accuracy data are presented for the two halves of the screen as well as overall.

Red Square/Green Square is a computerized task devised by Ponsford and Kinsella (1988) requiring subjects to respond as quickly as possible by pressing a left-hand button when seeing a red square and a right-hand button when seeing a green square. In each series a total of 45 squares are shown sequentially, with interstimulus interval ranging from 1.5 - 2.5 seconds. Of the 45, 15 are red, 15 green, and 15 of various colors. After the series, mean response time and number of errors are displayed.

Spot the Letter is another Ponsford and Kinsella (1988) computerized task, this one requiring detection of letters among a series of 30 numbers and 15 letters displayed sequentially with interstimulus interval ranging from 1.5 - 2.5 seconds. Again, mean response time and number of errors are displayed.

Evens and Fives, the third Ponsford and Kinsella (1988) computerized task, requires detection of even numbers and multiples of five when presented with a series of 40 numbers at interstimulus intervals ranging from 2.5 - 3.5 seconds. Half of the 40 numbers are targets. The usual speed and accuracy data is displayed.

Attention Process Training package (Sohlberg and Mateer, 1986) involves presentation of a variety of tasks in a variety of contexts. These are too numerous to elaborate here, but include tasks ranging from simple auditory vigilance (e.g., pressing a buzzer when hearing the letter "J") to tasks with considerably more complex demands (e.g., quickly sorting playing cards by suit and turning over those with an "S" in their name). Tasks generally do not incorporate immediate (trial by trial) feedback, instead providing feedback at termination of each block of trials. As task performance improves, a graded approach to increasing difficulty is employed.

Auditory attention task (Wood, 1986) consisted of presentation of digits at the rate of six every five seconds, with 70 digits in a series. The subject was required to tap on a table every time three consecutive odd digits were presented. The same series was repeated throughout a given training session, providing a total session 30 minutes in duration. Token reinforcement was used, with subjects earning one token for each correct response and losing one for each incorrect response. Tokens were exchanged for prizes following each session.

Visual attention task (Wood, 1986) used a Possum Basic Skills Teaching Machine, which displayed two 4 x 4 matrices of symbols. The symbols of the two arrays were identical but arranged in different order. When one of the symbols of the left matrix was illuminated, the subject was required to guide a moving light to the matching symbol in the right matrix. This was accomplished by pressing a button to start the light moving across a row, pressing again to stop at the correct column, pressing again to start moving down the column, and pressing a fourth time to stop at the correct cell. Correct responses produced a tone, and incorrect responses a buzz followed by repetition of the trial. This task, too, was administered in 30-minute sessions and incorporated a token system of punishment in which one token was deducted from the subject's cache for each incorrect response.

Figural Computerized Attention Task (FCAT) was developed specifically for the present study. This BASIC computer program presented successive "frames" of geometric figures, requiring subjects to press a key each time the specified target appeared. Each frame was characterized by five features: shape, color, size, number, and position of figure on the screen. Targets, to be described below, were defined by a predetermined subset of these features.

There were three potential attributes for each stimulus feature (see Table A.1) and only one attribute of each feature was displayed at a given time, producing a pool of 243 unique frames.

Table B.1
Physical Characteristics of FCAT stimuli

FEATURE	ATTRIBUTES
Size	large, medium, small
Number	two, three, four
Color	green, purple, orange
Shape	square, cross, bar
Position	top, middle, bottom

Training was provided in 200-trial blocks, with 1 - 3 blocks per session. During task performance, interstimulus interval was slightly shortened after each detection of a target, lengthened an equal amount after each miss or false alarm, and was unchanged by correct rejections. Changes in interstimulus interval were accomplished by changing the number of cycles in the FOR/NEXT loop interposed between trials, with a single cycle added to produce longer interstimulus intervals or subtracted to produce shorter intervals.

The subject was presented the goal of increasing rate of stimulus presentation as much as possible. Initial interstimulus interval was set to permit ample time for subject responses, but was also set so that by about the last third of the session the interstimulus interval would encroach on the subject's reaction time and thereby demand faster responses.

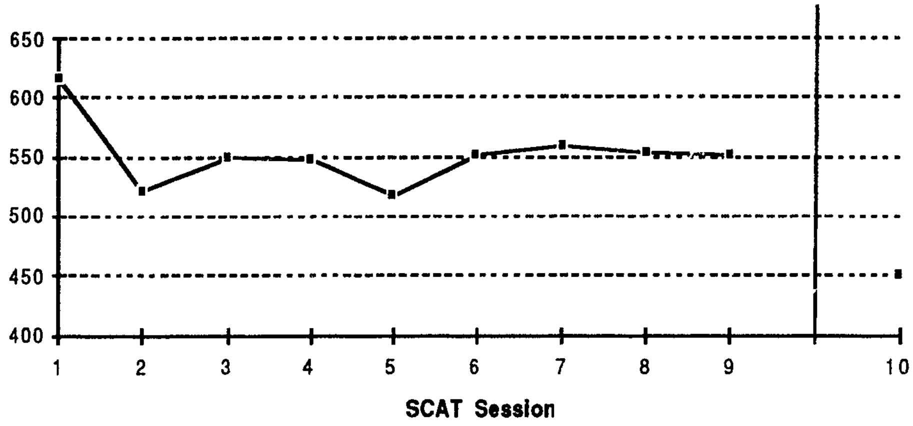
During task performance, auditory feedback was provided to indicate hits (a "beep"), misses, and false alarms (a "buzz"). After completion of each block, a summary score was presented indicating performance proficiency for that block (hits, false alarms, reaction time mean and standard deviation, and shortest interstimulus interval achieved). Because interstimulus interval was not constant but was linked to accuracy of performance, the validity of reaction time and accuracy measures was diminished. Shortest interstimulus interval for each block was felt to best reflect proficiency with the task.

The target was usually defined by a single feature: shape (specifically, squares). Some early subjects were administered targets becoming progressively more complex over sessions. This practice was discontinued with subsequent subjects since it seemed to offer no benefit to treatment.

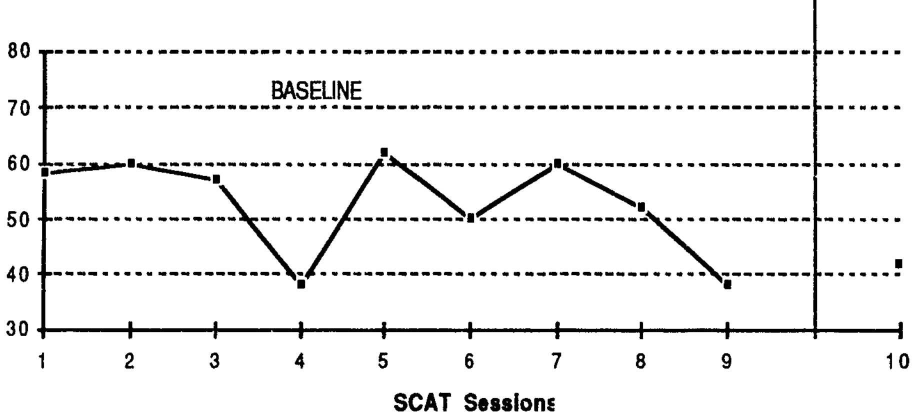
APPENDIX C

SCAT Pilot Data

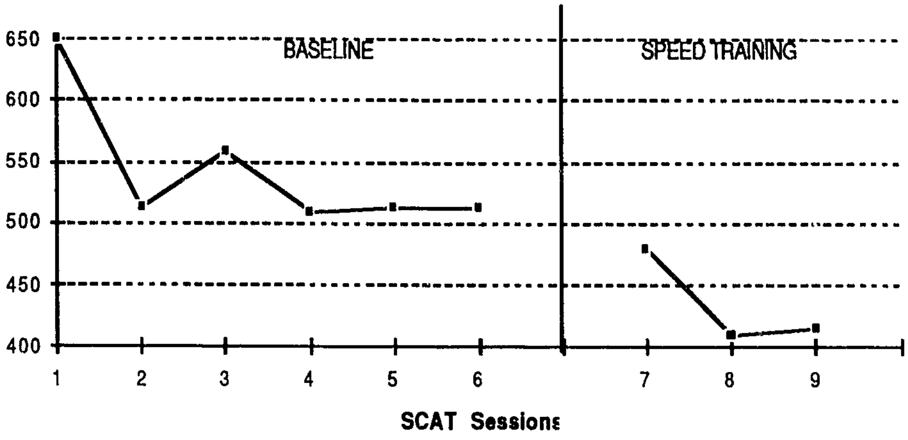
Pilot 1: SCAT Mean Reaction Time, in Millisecon



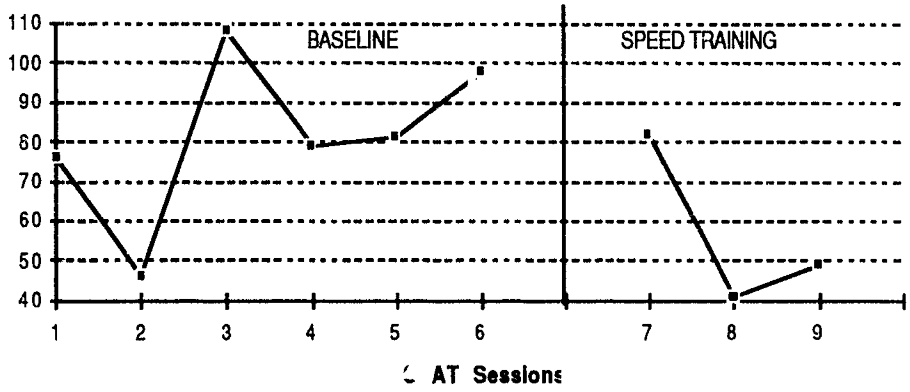
Pilot 1: SCAT Standard Deviation of Reaction Time, Milliseconds



Pilot 2: SCAT Mean Reaction Time, in Millisecon



Pilot 2: SCAT Standard Deviation of Reaction Time, Milliseconds



APPENDIX D

SUBJECT INSTRUCTIONS

Generalization Tasks: You will be (seeing/hearing) a series of (letters/shapes) on the computer. (Watch/Listen) carefully, because every time you (see/hear) (the letter 'X'/ squares) you need to press the spacebar as fast as you can. After we are through, the computer will tell us how fast you were and how many you got right. Make sure you pay careful attention because, although it seems like an easy task, it can be pretty difficult to keep your mind from wandering. Remember you need to press the bar as fast as you can without making mistakes.

Training Task: This is similar to some of the other tasks we've been doing, but there are some important differences. You will be seeing a series of shapes just like you are used to seeing, and just like before you need to press the spacebar as fast as you can whenever you see squares on the screen. This time, however, every time you press the bar when you're supposed to, you will hear a 'beep'. If you miss a square or press the bar when you're not supposed to, you will hear a 'buzz'. Also, every time it beeps, the computer will get a little faster and every time it buzzes it'll get a little slower. The idea is to see how fast you can get the computer to go.