Training Fluency in the Laboratory using Choice Reaction Time (CRT) Tasks: A Comparison of Instruction-based and Contingency-based Procedures

by

Diana Marie Grabavac
B.A., University of Victoria, 1991

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

MASTER OF SCIENCE

in the Department of Psychology

We accept this thesis as conforming
to the required standard

[redacted]

Dr. B.C. Goldwater, Supervisor (Department of Psychology)

[redacted]

Dr. L.E. Acker, Departmental Member (Department of Psychology)

[redacted]

Dr. J.A. Parsons, Outside Member (Counselling Services)

[redacted]

Dr. B. Howe, External Examiner (Department of Physical Education)

University of Victoria

Supervisor: Dr. Bram C. Goldwater

ABSTRACT

Practice on choice-reaction-time (CRT) tasks was used as a laboratory analogue to fluency (accuracy and response) training. The use of CRT tasks comprising well-learned stimulus-response relations allowed for the study of the acquisition of response speed without having to train response topography. In this systematic replication of Baron, Menich and Perone (1983), instruction-based and contingency-based procedures were compared as tools for promoting fluent responding on simple CRT tasks. Five female-undergraduate psychology students were trained first under instruction-based procedures, where they received response-by-response accuracy and end-of-trial accuracy and latency performance feedback. Next, all were switched into the contingency-based phase, where decreasing response latencies were progressively shaped. Here, they received both response-by-response and end-of-trial accuracy and latency performance feedback. In the last stage of training, monetary reinforcement was added to the shaping procedure. Following training, subjects completed a brief generalization phase, where the extent to which fluent responding generalized to non-training stimuli, that shared a history of control over the same responses as the training stimuli, was made. Contrary to Baron et al. (1983), but consistent with the CRT literature, relatively large decreases in latency were found under instruction-based training conditions. However, consistent with Baron et al., the introduction of the shaping procedure resulted in further decreases in latency of a magnitude greater than was predicted had subjects continued training under instruction-based conditions. Additional improvements were found during the shaping-plus-monetary-reinforcement condition. However, rather than promoting further decreases in response latency, the addition of monetary reinforcement appeared to promote more accurate responding, while latency changed little. Overall, contingency-based training procedures appeared to most effectively promote fluent responding. Finally, evidence of considerable generalization of fluent responding to functionally equivalent stimuli was demonstrated by all subjects.

Examiners:

[redacted]

Dr. B.C. Goldwater, Supervisor (Department of Psychology)

[redacted]

Dr. L.E. Acker, Departmental Member (Department of Psychology)

[redacted]

Dr. J.A. Parsons, Outside Member (Counselling Services)

[redacted]

Dr. B. Howe, External Examiner (School of Physical Education)

Table of Contents

Table of Contents (cont.)

Table of Contents (cont.)

Page

List of Tables

List of Figures

List of Figures (cont.)

Acknowledgements

I would like to begin by acknowledging and thanking my supervisor, Bram Goldwater, for being an excellent supervisor and mentor. It was a pleasure working with you!

Next, I would like to thank Tom Allen, who programmed SDT, Chris Darby, who provided technical support in the research lab, and Liping Zhang, who offered much assistance in the stats lab.

I would also like to thank Loren Acker, Carl Binder, Joseph Parsons, and David Polson for providing assistance at various stages of this project.

Finally, I would like to acknowledge and thank my parents, Kazimir and Evelyn Grabavac, for the guidance, support and encouragement they have given me for the past 26 years; with much gratitude, I dedicate this work to them.

Introduction

Reaction time has a rich history in psychology. Response speed, or its reciprocal, response latency, is a relatively easy variable to measure, and provides useful information about behavior.

In simple and choice reaction time studies, a series of stimuli are presented, and subjects are typically required to respond to each stimulus as quickly and as accurately as possible. Simple and choice reaction time (RT and CRT, respectively) tasks differ in that RT tasks involve a single stimulus and response, whereas CRT tasks involve two or more stimuli and responses. In early experimental psychology, reaction time was typically measured as an index of task "mastery" or of the complexity of the "inner processes" involved (Woodworth & Schlosberg, 1954). Researchers reasoned that the more completely a task had been mastered, the more quickly it could be performed. Also, the more difficult the task, or the more complicated the inferred underlying processes, the longer the reaction time would be.

In classical conditioning, response latency is frequently used as a measure both of the strength of a response and of the progress of the conditioning procedure (Catania, 1992, p. 43; Spence, 1956). Early in training, the conditioned response magnitude is relatively small and response latency is long. With more conditioning trials, the magnitude of the response increases and its latency decreases.

In (discriminative) instrumental conditioning, accuracy increases and response latency decreases as subjects gain more experience with the particular reinforcement contingencies in operation (Catania, 1984; Spence, 1956). As with classical conditioning, then, reaction time can be considered a measure of response strength or learning.

In recent years, cognitive psychologists have used response latency as an index of the "level of processing" required to complete a given task. Response latency, it is argued, increases when the inferred decision process requires greater central processing (Ellis & Hunt, 1993). Further, as the behavior is practiced and becomes "automatic," it is said that "central processing" requirements become negligible and reaction time decreases. The study of reaction time has also led to theories about "memory retrieval" processes. For example, Sternberg (1969) suggested that retrieval was a serial and exhaustive process, because response latency increased in a linear fashion as the length of the to-be-remembered list increased.

Researchers often report (e.g., Vernon, 1990; Welford, 1980, p. 356), under a variety of reaction time paradigms, reliable moderate-to-high correlations between response speed and performance on standardized intelligence tests. In fact, Vernon (1990) found that "speed of processing" accounted for up to 50% of the variance in some heterogeneous samples, making reaction time an important variable of study in human intelligence. Other researchers (e.g., Ackerman, 1988), however, report that these initial moderate-to-high correlations between intelligence and reaction time decrease when subjects train on the task. Ackerman also reported that on some tasks (e.g., CRT tasks), both between-subject and within-subject variability decreases with practice (p. 288).

Clinical neuropsychologists also measure and study reaction time (Lezak, 1983). Hicks and Birren (1970) reviewed a number of studies of the effects of various forms of brain damage on reaction time, and concluded that increased reaction time is one characteristic of brain damage. Further, they suggested that reaction time reliably distinguishes between some patient/non-patient populations (e.g., psychotics from nonpsychotics; learning disabled from nonlearning disabled; brain-damaged from nonbrain-damaged individuals), making reaction time an important variable in the diagnostic procedure (see also, Gaddess & Edgell, 1994). Welford (1980), too, reported that most psychopathological conditions are accompanied by slower and more variable response latencies, regardless of the type of task (RT or CRT), or the modality of the stimulus or response, and that increased response latency covaries with the severity of the condition (p. 356). Finally, response latency covaries with many factors in everyday life (e.g., fatigue; stress; gender; age), and is occasionally used as an index of some of these factors (e.g., fatigue; stress) (Welford, 1980; Woodward & Schlosberg, 1954).

Clearly, much useful information about behavior has been gained by investigating response speed as an indirect, observable measure of other variables. However, there has also been considerable interest in investigating response speed as an object of conditioning in its own right.

Response Speed as a Learned Behavior

Researchers studying verbal learning measure both accuracy and speed as performance criteria. They report that latencies continue to decrease after 100% accuracy has been reached (Judd & Glaser, 1969; see also Binder, 1993). Further, these researchers assert that learning beyond 100% accuracy can be measured in terms of faster responding, making response latency an important dependent variable and target behavior.

In mainstream education, slow, accuracy-intensive practice is the practice of choice for many reasons. For example, many educators believe that appropriate response speed develops with prolonged practice, and that by encouraging rapid, error-filled responding early in training, they are teaching poor study habits to their students that will be difficult to unlearn (Leonard & Newman, 1965). However, empirical evidence has challenged these assumptions. For example, Leonard and Newman (1965) reported that students were easily able to switch to a high accuracy (98%) training strategy, despite having first engaged in speeded practice (no accuracy criterion) on a five-choice CRT task. In addition, their performance was comparable to that of subjects who were required to maintain 98% accuracy throughout training. Similar to mainstream educators, teachers of learning disabled students strongly favour training students on "easy" tasks. With reading, for example, some educators advocate training the learners to read material where they can maintain 95-98% accuracy). These educators argue that successful completion of a given task builds the students' "self esteem" and increases their "motivation" to continue learning (Scott, Stoutimore, Wolking & Harris, 1990).

Precision teachers, a group of educational researchers and practitioners in the behavior analytic community, have also challenged the assumptions behind accuracy-intensive practice. Precision teachers assert that response speed is a property of behavior, and,therefore, appropriate response speed should be directly trained (Lindsley, 1990a; see also Binder, 1993). In fact, training, they assert, should continue until the task can be performed automatically or fluently. Behavioral fluency is defined as accuracy in combination with appropriate speed (Binder, 1987). Fluent performance is making the appropriate response "smoothly and without hesitation" (Binder, 1995, p. 3). Among the benefits produced by training to fluency, claim precision teachers, are greater retention of skill or learned material, higher resistance to distraction, and application (i.e., transfer of training) (Binder, 1987; Binder, 1988; Binder, 1993; Binder, 1995; Binder, Haughton, & Van Eyk, 1990; Johnson & Layng, 1992; Koorland, Keel, & Ueberhorst, 1990; Lindsley, 1990a). Eric Haughton (in Binder, 1995; also C. Binder, personal communication October 16, 1995) coined the mnemonic REAPS, which stands for Retention-Endurance-Application Performance Standards, to describe the additional benefits which result from training to fluency. Haughton suggested that "[precision teachers should] empirically determine what level of accuracy plus frequency is required to produce behavior that remains in the effective repertoire (retention), resists fatigue and distraction (endurance), and transfers to new situations and more complex behavior (application)" (Binder, 1995 p. 3). Precision teachers regularly report educational gains made by their students that are well in excess of those typically found in mainstream education. For example, Kent Johnson, director of Morningside Academy, reported that child learners generally complete greater than two grade levels per academic year, whereas adult learners typically complete greater than two grade levels per month or approximately 19 hours of instruction (Johnson & Layng, 1991; see also Binder, 1988; Lindsley, 1990b). Such educational gains are, clearly, well above the U.S. Government standard of one grade level per 100 hours of instruction (Johnson & Layng, 1991). Furthermore, Scott et al. (1990) reported that learning disabled students made larger educational gains when emphasis was shifted away from maintaining high accuracy on relatively easy tasks, to training on more challenging tasks (e.g., more difficult reading passages), where rate-pacing procedures were also implemented to increase response rate. Also, Johnson and Layng (1991) found that students diagnosed with attention-deficit disorder (ADD), greatly increased their "attention span" after extensive endurance training on a number of different tasks.

Frank Logan and his colleagues, who applied micromolar behavior theory to performance training, assert, also, that part of what needs to be learned or conditioned during training is response speed. (Logan, 1956, Logan, 1963, Shiffrin and Logan, 1965; Rogers, 1968). For convenience, these researchers will be referred to as micromolar behavioral researchers throughout the rest of this paper. Micromolar behavior theory treats quantitatively different behaviors (e.g., slow versus fast responses) as qualitatively different responses, rather than as different strengths of the same response (Logan, 1965). That is, rather than being an indirect measure of the strength of a response, response speed is a defining characteristic of the response. As a result, both accuracy and appropriate response speed must be trained if they are ultimately to occur under non-training conditions. In support of these assertions, Shiffrin and Logan (1965) reported that upon training two groups of subjects on the same paired-associate learning task, but under different response speed conditions (i.e., immediate versus delayed responding), subjects in the immediate-responding group produced significantly more correct responses when both groups were required to produce as many responses as possible during a two minute test phase. These results suggested that subjects in the fast group had an advantage over subjects in the slow group during speeded-testing conditions, because, during training, the former were required to respond quickly. Presumably, although this was not specifically tested, subjects in the slow group would have had an advantage over subjects in the fast group had the test required delayed responding. Similarly, Rogers (1968) was interested in whether changing stimulus-presentation and response speed requirements affected performance on paired-associate learning tasks. In this experiment there were six groups of subjects. Stimuli were presented to each group either quickly or slowly (i.e., after an inter-stimulus interval of 2 s or 5 s, respectively). In addition, subjects were required to respond either "immediately" or after a short delay (i.e., .8 s or 3 s later, respectively). During the test phase, three new paired-associate lists were presented, and either presentation speed, response speed or both rate parameters changed for 75% of the subjects in each group. Rogers (1968) reported that changing either or both of the rate parameters produced a decrement in performance. These data suggest that performance is best at practiced speeds, and that changing speed components increases the difficulty of the task. Furthermore, these data provided support for the contention that training task-appropriate response speed is necessary and important.

According to precision teachers and micromolar behavioral theorists, then, mainstream educators may be doing students a disservice when the desired behavioral outcome is different than the trained behavior. For example, the successful development of speed reading and touch typing require lots of speeded practice, not just lots of practice (Logan, 1963). Also, learning basic arithmetic to fluency is essential for students to be successful at learning higher math operations (Binder, 1993; Gagné, 1983; Haughton, in Johnson & Layng, 1991). If students are unsuccessful at achieving fluency at the basic levels, they will also be unable to learn to perform at higher levels (Johnson & Layng, 1991).

Training on a musical instrument provides another example of the importance of learning response speed (Binder, 1995). Students are not only trained to play each note of a musical score accurately, but appropriate response speed is also taught as part of what needs to be learned to correctly play each note. Indeed, learning the tempo of a musical score is as important as learning the notes themselves.

Precision teachers and micromolar behavioral researchers agree that training response speed is very important. They do, however, disagree over how important it is to have an established accuracy criterion early in training.

According to Ogden Lindsley (1990a), founder of precision teaching, the early aim of precision teachers was "to introduce rate of response with standard, direct, continuous and self-recording to public school classrooms...because laboratory research had proven rate to be more sensitive [to learning variables] than percent correct and other less direct measures of behavior (e.g., percent of time observed on task)" (p. 7). Lindsley (1990b) advocates emphasizing speed over accuracy during training, and has reported that the more challenging "error-filled" (i.e., high speed) learning has resulted in better overall performance by his students, relative to traditional accuracy-oriented curricula. Furthermore, his students reported finding the "error-filled" learning more "enjoyable" than traditional, accuracy-based instruction. Among his many achievements, Lindsley, along with Stephen Graf, designed a flashcard training procedure called SAFMEDS (Say-All-Fast-Minute-Each-Day-Shuffle) to help students achieve high fluency on various tasks (Potts, Eshleman, & Cooper, 1993). According to McGreevy (in Potts et al., 1993), students are required to answer as many flashcards as possible in a given amount of time (e.g., one-minute timings may be used), and then count up the number of incorrect and correct responses. Typically, the number of incorrect responses far exceeds the number of correct responses early in training. Training continues until students achieve some fluency goal (e.g., 40 correct-responses-per-minute). Note that students are not penalized for making incorrect responses; rather, errors are considered to be "opportunities" for further learning (McGreevy, in Potts et al., 1993). Lindsley (in Potts et al., 1993) suggested further that training fast responding leads to greater performance gains relative to more mainstream training techniques. Binder (C. Binder, personal communication, October 16, 1995) explained further that while precision teachers generally do not work to eliminate [seemingly] "random" errors, clearly, there is a need to decelerate "systematic" errors. However, too much emphasis on accuracy impedes rate building (see also Lindsley, 1990). Interestingly, Eric Haughton (C. Binder, personal communication, October 16, 1995) observed that once his students achieved around half of their fluency aim, it became easier to decelerate their error-rates. In summary, precision teachers encourage high rates of responding from the onset of training, and chart both rate correct and rate incorrect for each student (Lindsley, 1990b). If necessary, using the student's chart as their guide, precision teachers implement procedures to eliminate persistent errors. Otherwise, if rate correct continues to increase, precision teachers tend to focus less on errors than on rate (C. Binder, personal communication, October 16, 1995). Training continues until fluent performance has been achieved and Retention-Endurance-Application Performance Standards (REAPS) are met.

Whereas precision teachers support training appropriate response speed from the onset of training, micromolar behavior theorists (viz., Shiffrin and Logan,

1965) suggest that with tasks involving difficult responses, initial, slow practice is often essential. Here, presumably, optimal response speed is gradually shaped in order to maintain high accuracy. Further, it is presumed that without an initial minimum accuracy criterion, fluent performance would not result from training. In contrast to precision teachers, then, micromolar behavior theorists suggest that sometimes early emphasis should be placed on accuracy alone. However, micromolar behavior theorists provide no evidence to support this claim, nor do they attempt to define which types of tasks require initial, accuracy-intensive practice.

Despite mainstream educators' preference for accuracy-oriented curricula, the results reported by precision teachers and micromolar behavior theorists illustrate the importance of training response speed (i.e., behaviors to fluency). What remains unclear, however, is which procedure most effectively promotes fluent responding.

Training fluency on tasks involving simple discrimination

One method of studying the acquisition of fluency is to simplify the discriminations involved. The present study employed CRT tasks that involve relatively simple discriminations (e.g., pressing one of three adjacent keys in response to the presentation of one of three numerals on a computer screen). Simple CRT tasks were chosen because they minimize both the variance arising from errors in typing words, and from the acquisition of initial stimulus control of each stimulus over the correct response. A number of CRT studies (e.g., Ackerman, 1988; Hale, 1969; Kristofferson, 1977; Pashler & Baylis, 1991; Rabbitt & Banerji, 1989; Schneider & Fisk, 1984) have demonstrated clear improvement in response latency as a result of practice. However, because of procedural differences in duration of training, feedback contingencies, instructions, monetary incentive, to name a few, it was difficult to determine which variables contributed to fluent performance. Rabbitt & Banerji (1989), for example, used a three-choice (numbers) CRT task, stressed speed only, paid subjects a small amount for participating but did not provide any performance feedback, and ran subjects for a minimum of 10,000 trials over five days. Schneider & Fisk (1984), on the other hand, used a three-choice (categorization) CRT task, stressed both speed and accuracy, provided both trial-by-trial and cumulative accuracy feedback, paid subjects for participating, and ran subjects for 1420 trials over three days. Despite the many differences between these two studies, both research teams reported a similar decrease in mean latency correct (approximately 18%) over training.

Baron, Menich and Perone (1983) observed that one important difference between human and animal reaction-time research is that animal researchers use contingency-based procedures (i.e., the target response is differentially reinforced), whereas human researchers use rule-governed procedures (i.e., giving verbal instructions) to promote rapid responding. As a result, Baron and his colleagues conducted a series human reaction-time experiments that examined the effects of directly shaping response speed. Since Baron et al. (1983) most clearly addressed the issue of the necessity of directly training response speed, it will now be described in detail.

Baron, Menich, and Perone (1983). A conditional discrimination, matching-to-sample task was used in this experiment. This task was essentially a 2-choice CRT task, in that after the target stimulus was presented, two further stimuli (one target and one distracter) were presented, and subject were required to press the response key paired with target stimulus. Prior to the presentation of each target stimulus, subjects were instructed to choose the comparison stimulus that either matched (choose "same") or did not match (choose "different") the target stimulus. Subjects practiced initially under a baseline condition, where they earned 1.75 cents contingent upon every correct response, to a maximum of $2 per hour. Baron et al. (1983) informed subjects that they would have to respond quickly in order to maximize earnings. However, the authors also informed subjects that whereas unlimited time was available for responding during baseline, time limits would be imposed during training. That is, during training, subjects earned monetary reinforcement contingent upon only those correct responses that met the latency criterion for reinforcement. Baron et al. chose, as the latency criterion, the number of milliseconds on the immediately preceding training session that separated the fastest 80% of responses from the slowest 20%. An adjustment was made to a subject's latency criterion, provided 90% of those responses made within the time limit on the previous session were correct. However, if no change was made after three sessions, the latency criterion was automatically adjusted, provided that the new latency criterion was more stringent than the existing one. Baron et al. used a "modified" latency shaping procedure, in that the latency criterion could only be made increasingly stringent as training progressed. Subjects completed approximately 900 trials (a trial being a single stimulus presentation and response) during baseline and around 1500 trials during training. Following training, subjects returned to baseline conditions (500 to 1200 trials), and then completed a second brief training phase (500 to 600 trials). Conditions were reversed to determine how the removal and the reintroduction of the shaping procedure might affect response speed. In addition to the monetary reinforcement earned during training, subjects were paid $2 per hour of participation, contingent upon completion of the experiment. Subjects committed themselves to participate for 40 hours, which were scheduled over a period of 2-4 weeks.

Baron et al. (1983) reported that although response speeds increased during initial baseline, greater relative improvements occurred when the time limit contingencies were in place. When baseline was reintroduced, subjects continued to respond at levels comparable to those found during latter training. This suggested that, at least in the short term, speeded responding could be maintained in the absense of the shaping procedure. Finally, Baron et al. reported that subjects made no further improvements during the second shaping phase. Overall, these results suggest that the contingency-based procedure was superior to the rule-governed procedure in facilitating the development of speeded responding, at least with CRT-type tasks. Unfortunately, however, instructional (S^D) and reinforcement (S^{r+}) effects may have been confounded, in that provision of reinforcement for slow responses during instruction-based training may have inadvertently cued subjects to respond slowly, despite instructions to the contrary. As a result, these results must be interpreted with caution.

Preliminary research. In preliminary work done in the present researchers laboratory, simple discriminations were trained, and emphasis was placed almost exclusively on speed of responding. Four subjects trained on three-choice CRT tasks, using a computer program called Speeded Discrimination Training (SDT), and received response-by-response and end-of-trial accuracy and latency feedback. Two unique, three-choice CRT tasks, one a numbers task, the other a categories task, were created for this study. The digits "7," "8," and "9" served as numbers stimuli, and five exemplars each from "animals," "body parts," and "occupations" served as categories stimuli. Two subjects trained on each task. With the numbers task, subjects had to press the key paired with the digit presented. With the categories task, subjects had to press the key paired with the category that matched the exemplar presented. During each 30-second trial, subjects were instructed to respond to as many stimulus presentations as possible. A fixed response-stimulus interval (RSI) of 500 ms was used in this study. At the end of every trial, the latency criterion was automatically readjusted, regardless of the number of errors subjects made. The number of milliseconds which separated the fastest 75% of responses from the slowest 25% was chosen as the latency criterion for the next trial. During each session, subjects completed 30 or 40 trials, and training lasted for up to 25 days.

The results of this experiment were mixed. The performance of two subjects (one from each task) improved across training. However, these improvements were small relative to those typically reported in the CRT literature. That is, accuracy was stable or increasing, but relatively low, while latency decreased. The performance of the remaining two subjects, in contrast, was highly variable throughout training, and there was little evidence for improvement.

Several procedural changes were made part way through training. First, it became apparent that too much emphasis had been placed on speed when three of the four subjects were responding regularly with latencies below 100 ms. Merkel (in Keele, 1973, p. 73) reported that a minimum of 364 ms were required to accurately respond on a three-choice CRT task. This suggested that the task had become a RT task, in that the properties of the various stimuli did not appear to be controlling responding. As a result, the instructions were changed to "maximize the number of correct-and-fast responses per trial," thus greater emphasis was placed on accuracy. <u>SDT</u> was also modified to reject responses with latencies less than 200 ms. Second, the latency adjustment procedure was changed from a trial-by-trial to a session-by-session basis. The latency criterion was now calculated in the following way. For each session, mean latencies per group of five trials (six or eight per session) were ordered, and the second longest value was chosen as the new latency criterion. However, two subjects continued to show no improvement. Late in training, monetary reinforcement, contingent upon target performance, was added to the shaping procedure of those subjects who were training on the numbers task. The result of this manipulation was that both subjects became more "eager" to continue training on the task. Early in training, one of the numbers-task subjects (the one who showed no improvement) had complained that the number of trials per session was too high (i.e., training was "too fatiguing"). As a result he was allowed to complete 10 fewer trials per day than the other subjects. However, when monetary reinforcement was introduced, this subject asked to return to completing 40 trials per session. Furthermore, this subject's performance became less variable during this phase. In summary, given all the methodological changes made during this experiment, the most appropriate conclusion that could be drawn was that further investigation was both necessary and worthwhile.

Fluency training, functional equivalence and generalization

As discussed earlier, training to fluency on a given task is reputed to provide several important benefits to the learner, beyond proficiency on the task itself. Again, the benefits reported to accompany the development of fluency include greater retention, longer endurance, and improved transfer of training. As such, an investigation of the extent to which the training of speeded responding generalized beyond the training environment seemed to be a very meaningful and worthwhile exercise. This became the second goal of the present experiment.

Two or more stimuli are functionally equivalent when they exert discriminative control over an identical response (Catania, 1992; Hall & Chase, 1991; Sidman, 1986). For example, if a child learns to say "bird" when shown a picture of a bird and the word "b-i-r-d," the picture of the bird and the word "b-i-r-d" are functionally equivalent stimuli, because they both exert stimulus control over the same response (i.e., saying "bird"). Hall and Chase (1991) add that because members of this set of stimuli control the same response, they are substitutable for each other when trying to evoke the target response. Here, then, either presentation of a picture of the bird or the word "b-i-r-d" would be appropriate when prompting the child to say "bird." Functional equivalence may be a result of direct training, like above, or involve emergent relations (Hall & Chase, 1991, p. 116).

Catania, Horne, and Lowe (1989) provided evidence, collected on one subject, for transfer of function among members of an equivalence class. They first trained a different response rate to each of two unique stimuli (respond quickly to STAR symbol; respond slowly to TREE symbol). Next, they established arbitrary matching-to-sample relations between these stimuli and two novel stimuli (i.e., matching of STAR to WORM; and, matching of TREE to BLOCK). During the test of rate transfer, they found that response speed generalized to the novel stimulus. That is, the subject responded quickly to WORM and slowly to BLOCK.

Vaughan (1988) arbitrarily divided 40 slides into two sets, and trained pigeons to respond to half of them (Set A), while not responding to the other half (Set B). He was interested in whether, when the reinforcement contingencies were reversed for a subset of slides in each set, responding to the remaining slides would change appropriately in the absence of further training. Vaughan reported that, although not so initially, as training progressed and reinforcement contingencies were periodically reversed, subjects were able to detect the change, presumably after the first few stimulus presentations, and respond appropriately to the remaining stimuli. In sum, the response trained to some members of two functionally equivalent classes of stimuli generalized to the remaining members of each class without further training.

Similar to Catania et al. (1989) and Vaughan (1988), the present researcher was interested in whether a newly-trained response (here fluent responding) would generalize to non-training stimuli which were functionally equivalent to the training stimuli. Although considered secondary in this experiment, the importance of this generalization test should not be understated. For many years, developmentalists and linguists have criticized behavior analytic theory for failing to explain language acquisition, particularly the emergence of "novel" verbal behavior (Hall and Chase, 1991). However, by conducting experiments like the ones described earlier, equivalence researchers are beginning to gain an understanding of the environmental determinants of language acquisition. As such, the data collected in the present study may make a valuable contribution not only to the CRT and learning literatures, but also to the equivalence literature.

The Experiment

The primary goal of the present research was to define an effective and efficient procedure for training speeded responding with high accuracy (fluency) on three-choice CRT tasks. Of interest, also, as a secondary goal, was the degree to which training effects generalized to non-training stimuli that share a history of control over the same response classes as the training stimuli (i.e., functionally equivalent stimuli).

In terms of the primary goal, the literature provides no clear indication of the variables that most effectively facilitate the development of fluency. Recall that mainstream educators focus almost exclusively on accuracy. Precision teachers suggest that emphasizing speed, or speed in combination with accuracy, from the onset of training, produces benefits beyond simple fluency on the training task. However, they offer little empirical data to support these claims, although this is changing. Like precision teachers, micromolar behavior theorists assert that training appropriate response speed is necessary, but suggest that, at least with some tasks, initial accuracy-intensive practice is critical. However, they, too, offer little empirical evidence to corroborate the claims they make. Work done in the present researcher's laboratory, where only speed was emphasized during initial training, produced very mixed results. Only two of four subjects demonstrated decreasing latencies during training, and their accuracies were low relative to values reported in the CRT literature. Baron et al. (1983) suggested that directly shaping rapid response speed was necessary for subjects to approach their maximum speed. Unfortunately, however, reinforcement and instructional effects may have been confounded in the initial training phase of their study, making the results of their comparison between instruction- and contingency-based procedures inconclusive. Finally, with the CRT literature, where methodology is typically instruction-based (although some contingency-based procedures, like monetary reinforcement, often play supplementary roles), no clear indication of optimal training conditions has emerged, because methodologies remain fairly unique across studies. What should be noted, however, is that most CRT studies imposed some kind of accuracy criterion on subjects early in training. In light of the above incertitude, investigating the effects of some of the seemingly important operant variables (e.g., shaping response speed; using monetary reinforcement) seemed a worthwhile exercise. Such research would allow us to begin identifying those variables which most effectively and efficiently promote the development of fluent responding on CRT-type tasks.

In terms of the primary goal, then, the present experiment was designed to answer the following questions:

1. Does training with an operant procedure (i.e., speed- and accuracy-contingent performance feedback; shaping) produce an

improvement in performance over typical, instruction-based CRT training?

2. Late in training, does the introduction of monetary reinforcement, contingent on response latency, result in even better performance?

As mentioned earlier, additional benefits are often reported to accompany the acquisition of fluency. That is, when subjects become fluent on a given task, researchers often report that they are able to continue working on the task for longer periods of time, retain the skill longer, and more easily learn complex tasks of which the fluent skill is a component part. Of interest, at present, is the extent to which the present training procedure promotes generalization of fluent responding to non-training stimuli which share a history of control over the same responses as the training stimuli. In terms of the secondary goal, then, the present experiment addressed the following question: does speeded responding generalize to non-training stimuli which are functionally equivalent to the training stimuli?

Method

Subjects.

Five female-undergraduate psychology students from the University of Victoria participated in this experiment. Three subjects participated as partial fulfilment of course requirements for an independent research project. These subjects were recruited from an advanced behavioral psychology course. Subject 1 was 43 years old, subject 2 was 24 years old, subject 3 was 26 years old, and all were in their third year of studies. The remaining two subjects were recruited from an unrelated fluency research course, and were paid $4 per day for attendance and participation. Subject 4 was 32 years old and in her second year of studies, and subject 5 was 22 years old and in her fourth year of studies.

Apparatus & Materials.

Speeded Discrimination Training A computer program called Speeded Discrimination Training (SDT) was created to meet the methodological demands of the present experiment. Most generally, SDT was designed to present stimuli on a computer screen, record response accuracy and latency, provide performance feedback to subjects, and store the data for later analysis.

The following is brief description of how SDT works. First, the session parameters (a subject's training or test information) are entered, and the training or test session begins. During each trial, stimuli are presented randomly, one at a time on the computer screen, and the subject learns to press the key on the 3-key response box which has been paired with each stimulus. After each response, subjects receive brief auditory (accuracy alone or accuracy & latency) performance feedback, followed by a brief response-stimulus interval (RSI), where the screen is blank. Following the RSI period, the next stimulus is presented, and the procedure is repeated until the trial is over. At the end of each trial, performance feedback for that trial and the preceding four trials is flashed onto the computer screen for the subject's perusal. The subject must then press the "enter" key on the computer keyboard, which sets up a new trial. After the subject completes a predetermined number of training trials or a given amount of time passes, SDT signals the end of the session, and the data are stored to file. At the end of the training session, a subject's session parameters can be updated.

Two additional features of SDT merit discussion. First, SDT does not allow subjects to correct their errors before proceeding to the next stimulus. This feature frees up more time per trial for speeded response training, and eliminates the possibility that an error-correction requirement might impose an implicit accuracy criterion on subjects, which may influence task performance. Second, SDT was programmed to print all per-trial and end-of-session performance data, allowing the experimenter to constantly monitor each subject's progress.

Performance measures collected by SDT. SDT was designed to collect or compute the following measures. First, accuracy as percent correct (% C) is scored. Second, a measure which reflects both accuracy and speed as percent-correct-and-fast (% C&F: the percentage of responses which are both correct and meet the latency criterion) is calculated. Third through seventh, SDT collects four reaction-time measures: mean and median latency correct and mean and median speed correct. Eighth and ninth, the standard deviations (SD) of the latency and speed measures are computed. Tenth, rate-correct-per-minute (RCPM) is calculated. SDT computes this statistic by dividing the total number of correct responses on each trial by the sum of the latencies for all correct and incorrect responses. In addition, SDT records the distribution of latencies for correct responses at the end of each training session. Collection of these data is useful, because it provides information about changes in the distribution of latencies across sessions, days or even phases.

SDT collects or computes all performance measures after every trial, and all measures are available to the experimenter on a per-trial basis, with the exception of the distribution of latencies for correct responses, which is only available on a per-session basis.

Experimental Task: Stimulus-Response Pairs. The training and generalization tasks chosen for this experiment were simple, three-choice, numbers and categories CRT tasks. For the numbers task, the training stimuli were "zero," "four," and "nine" in either uppercase or lowercase form. During the generalization phase, the numbers stimuli were the same numbers presented in the non-training case-size. For the categories task, five exemplars (see Table 1) from each of three semantically unrelated categories were chosen as stimuli. The exemplars, selected from a list of category norms (Battig & Montague, 1969), were between four and six letters in length, and were considered to be highly familiar to native English speakers (Kučera & Francis, 1971). Similar to the numbers task, categories-task stimuli were presented in either uppercase or lowercase during training, and the non-training case-size was used in the generalization phase.

Trials and sessions. Each training and generalization trial lasted 20 seconds, and consisted of as many stimulus presentations and responses to stimuli

Table 1

Category task stimulus-response pairs

Animals	Colours	Sports
deer	blue	golf
tiger	green	track
horse	white	tennis
sheep	purple	hockey
rabbit	yellow	soccer

as each subject could make during that period of time. A fixed RSI of 300 ms was used. Subjects completed three, 15-trial sessions per day, and trained five times per week for between six and eight weeks. Subjects were free to take short breaks whenever necessary, so long as they were able to complete the required amount of training in the allotted time (45 minutes per day).

General Procedure

The experiment was divided into two parts: speeded discrimination training and a test of generalization of speeded responding to functionally equivalent stimuli. Speeded discrimination training was divided into three different phases. Subjects trained first under instruction-based conditions (Phase A). Next, they trained under shaping of speeded responding conditions (Phase B), and, finally, they trained under shaping-plus-monetary-reinforcement conditions (Phase C). A multiple-baseline-across-subjects design (i.e., subjects trained under each condition for different amounts of time) was used in this experiment to preclude the possibility that some unidentified variables, correlated with time, rather than the manipulation of the independent variable, were responsible for changes in the level of performance across training. Also, two subjects experienced reversals (i.e., ABAB) after Phase B. Subjects N02 and C02 completed several additional sessions under instruction-based conditions to assess what effect the removal of the contingency-based procedure would have on their performance. Finally, in preparation for the final training phase (Phase C), subjects N02 and C02 trained further under Phase B conditions.

Assignment of subjects to training tasks. A quasi-random procedure was used to assign subjects to training task. That is, after two conditions had been met, subjects were randomly assigned to one of two training tasks. First, because the numbers task was considered the easier of the two tasks, involving only three stimuli compared to 15 for the categories task, it seemed prudent to have only two subjects train on it, while the remaining three subjects trained on the categories task. Second, because the paid subjects were concurrently involved in an unrelated fluency experiment, it seemed possible that their research history could affect their data. In other words, different "subject expectations" from those of the non-paid subjects might have been in place due to this research history. As such, the two paid subjects were each assigned to different training tasks. This was done so that if their data had to be removed from analysis, data would still have been collected on both training tasks. Which paid and non-paid subjects trained on the numbers task was determined randomly, and the remaining subjects were assigned to the categories task. The following assignment of subjects to task was used: subjects one and five became (number) subjects N01 and N02, respectively, while subjects two, three and four became (categories) subjects C01, C03, and C02, respectively. Furthermore, subject N01, C02 and C03 trained in lowercase, while subjects N02 and C01 trained in uppercase.

Pretest. Before training began, a test was given to ensure that the training and generalization stimuli were functionally equivalent under non-speed conditions. Here, both training and generalization stimuli were presented randomly, and subjects were instructed to respond by pressing the response key paired with the stimulus with 100% accuracy. Subjects received both response-by-response and end-of-trial accuracy feedback. That is, after each response, the response key that the subject pressed was flashed onto the computer screen, and the subject heard one of two auditory tones. A "beep" was contingent upon a correct response, while a "buzz" was contingent upon an incorrect response. Also, at the end of the trial, accuracy for that trial and the preceding trials was displayed on the computer screen. The subjects completed only three trials during this phase.

Phase A - Instruction-based training: This procedure was designed to closely approximate the methodology most typically found in the CRT literature, where neither speed-contingent feedback for each response nor shaping were employed.

In this initial training phase, subjects were instructed to respond as quickly as possible, while keeping their accuracy at or above 90%. Response-by-response feedback identical to that of the pretest was used during this phase. At the end of each trial, accuracy and median latency correct scores for that trial and the preceding four trials were displayed on the computer screen. Subjects were presented with median latency correct, rather than the more traditional mean value, because given the skewed nature of reaction time data the median score was considered to be more stable than the mean, which is sensitive to extreme scores. To maximally encourage rapid responding and avoid the confound between reinforcement and instructional effects suspected to be operating in the Baron et al. (1983) study, subjects were further instructed to always be trying to reduce their median latency correct scores while keeping accuracy at or above 90%. Also, the instructions were regularly read to subjects, or, alternatively, subjects were asked to verbalize them. In addition, the instructions were posted on a card located adjacent to the computer screen for the subjects perusal. Furthermore, trial-by-trial and end-of-session data were printed on a printer located outside the training chamber so the experimenter could monitor performance, and subjects were informed that they were constantly being monitored. Finally, at the end of each session, subjects were required to graph their grand RCPM score, and these graphs were displayed in the lab for all subjects to see. One could argue that these factors added an "experimenter demands" component to the procedure. Subjects trained in this phase for between seven and 15 days.

Phase B - Shaping speeded responding: During this phase, progressively faster (accurate) responding was shaped using a procedure based upon that found in Baron et al. (1983). Subjects were instructed to work as quickly as possible while keeping their percent-correct-and-fast score at or above 85. Here, as training progressed, only correct responses that were faster than or as fast as the progressively decreasing latency criterion were reinforced. Subjects received the following types of auditory, response-by-response performance feedback: a "beep" was contingent upon a response which was correct and met the latency criterion; no tone was contingent upon a response which was correct but did not meet the latency criterion; and, finally, a "buzz" was contingent upon a response which was incorrect. As during baseline, the key pressed immediately after the presentation of the stimulus was displayed on the screen. At the end of each trial, percent-correct-and-fast and median latency correct scores for that trial and the preceding four trials were displayed on the computer screen. Subjects trained in this phase for between 11 and 21 days.

Procedure used to adjust latency criterion. The following procedure was used to adjust the latency criterion on all but the first Phase B training sessions. As discussed earlier, subjects were required to meet a minimum of 85% C&F on each trial. At the end of the training session, if the subject achieved 90% C&F or greater, then her latency criterion was made more stringent. If, alternatively, the subject was unable to maintain a minimum of 85% C&F, then her latency criterion was made less stringent. Otherwise, the criterion was not changed. The adjustment was made by appealing to the distribution of latencies for correct responses. The difference, in milliseconds, between the 40th and 50th percentile of this distribution was chosen as the number of milliseconds to subtract from or add to the current latency criterion. Adjustments were made on a session-by-session basis.

The latency value at the 90th percentile of the distribution of latencies for correct responses on the last instruction-based training session served as the initial latency criterion. The latency criterion was adjusted on either or both of the second and third training sessions each day. Finally, the latency criterion on session 1 was always the same as that used during the final training session on the previous training day.

Phase C - Shaping-plus-monetary-reinforcement. To establish whether the addition of monetary reinforcement would further enhance training effects, subjects completed the third and final phase of training. Here, subjects were told that they would earn 5 cents for each trial where they scored 85-89% C&F; 15 cents for % C&F scores between 90 and 94; and 30 cents for scores of 95% C&F or greater. Subjects could also earn a bonus of 30 cents on every trial where they beat their highest to date rate-correct-per-minute (RCPM) score. Subjects were required to fill out an earnings sheet during this phase. Subjects trained in this phase for between two and eight days.

Generalization test. To determine whether training effects generalized to functionally equivalent, non-training stimuli, subjects completed a generalization test. This final phase consisted of 28 trials, and was divided into two sessions. The odd trials were identical in every way to the previously completed training trials. The even trials, however, differed from the odd trials in that subjects practiced speeded responding with the generalization stimuli. Recall that the generalization stimuli were the training stimuli in their non-training case-size. The latency criterion used during the test phase was the final criterion used during training.

Results and Discussion: Training Phases

Before considering the data, certain features of the present methodology merit discussion. During the initial meeting with the experimenter, subjects completed one session under instruction-based training conditions. This initial training session, referred to as "P" for practice, was given to allow task-relevant stimuli (e.g., training stimuli; instructions) to acquire control over responding, while, at the same time, minimizing the control exerted by irrelevant stimuli (e.g., physical characteristics of the lab; background noises). Similarly, subjects completed only one session on the initial shaping day to keep conditions across phases comparable, and to allow subjects to become familiar with novel training instructions and performance feedback. Subjects, however, completed two sessions on the first shaping-plus-monetary-reinforcement training day, because training conditions had not changed significantly relative to Phase B, and it was desirable to collect as much data as possible over the few remaining training days.

To provide as global a picture as possible of each subject's performance, daily accuracy, mean latency correct, SD of latency correct, and RCPM scores were computed and graphed in the following manner. Accuracy (% C), mean latency correct (MLC) and SD of latency correct (SDLC) were presented on one graph to illustrate individual trends across each performance measures within and across training sessions. This presentation was informative in cases where one of either accuracy or mean latency correct were relatively stable while the other measure changed. However, as a result of the well-established trade-off between speed and accuracy, all other things being equal, when decrements occurred in both performance measures, it was not possible to separate the amount of decrease in latency attributable to loss in accuracy from that characteristic of a genuine practice effect. A post-hoc statistic called "latency on 100% accuracy trials" was created to allow the experimenter to overcome the ambiguity in the accuracy and latency data. This statistic was computed in the following manner. For each day, an average latency score was computed on only those trials where the subject scored perfect accuracy. In this way, accuracy was held constant, allowing the experimenter to study the trends in the latency data within and across training phases. Latency on 100% accuracy trials, however, had one inherent problem. As the number of 100% accuracy trials was variable across days and within and across training phases, there was some question as to the representativeness of this statistic. The RCPM statistic, on the other hand, appeared to offer a way of analyzing both latency and accuracy data, as it was computed from both performance measures, without the loss of any potentially important data, or question as to the representativeness of the statistic. In the end, all performance measures were graphed and studied for each subject. Since parallel effects were seen across the different measures, only the RCPM data will be discussed in detail. However, on occasion, some of the other performance measures will be referred to. The reader can locate the graphs of the remaining performance measures, in the same order as the data are discussed in the next section, in Appendix A.

Included on all graphs are the first day of each training phase, where fewer than three sessions were completed. It seemed more appropriate to include these data, given this qualification, rather than to include them without qualification or exclude them all together. They are, after all, important data. Also, the reader should note that daily performance measures were computed by taking an average of the scores obtained on each of three, daily training sessions. As such, the SD scores presented are, in fact, means of SDs rather than true SDs.

Numbers Task

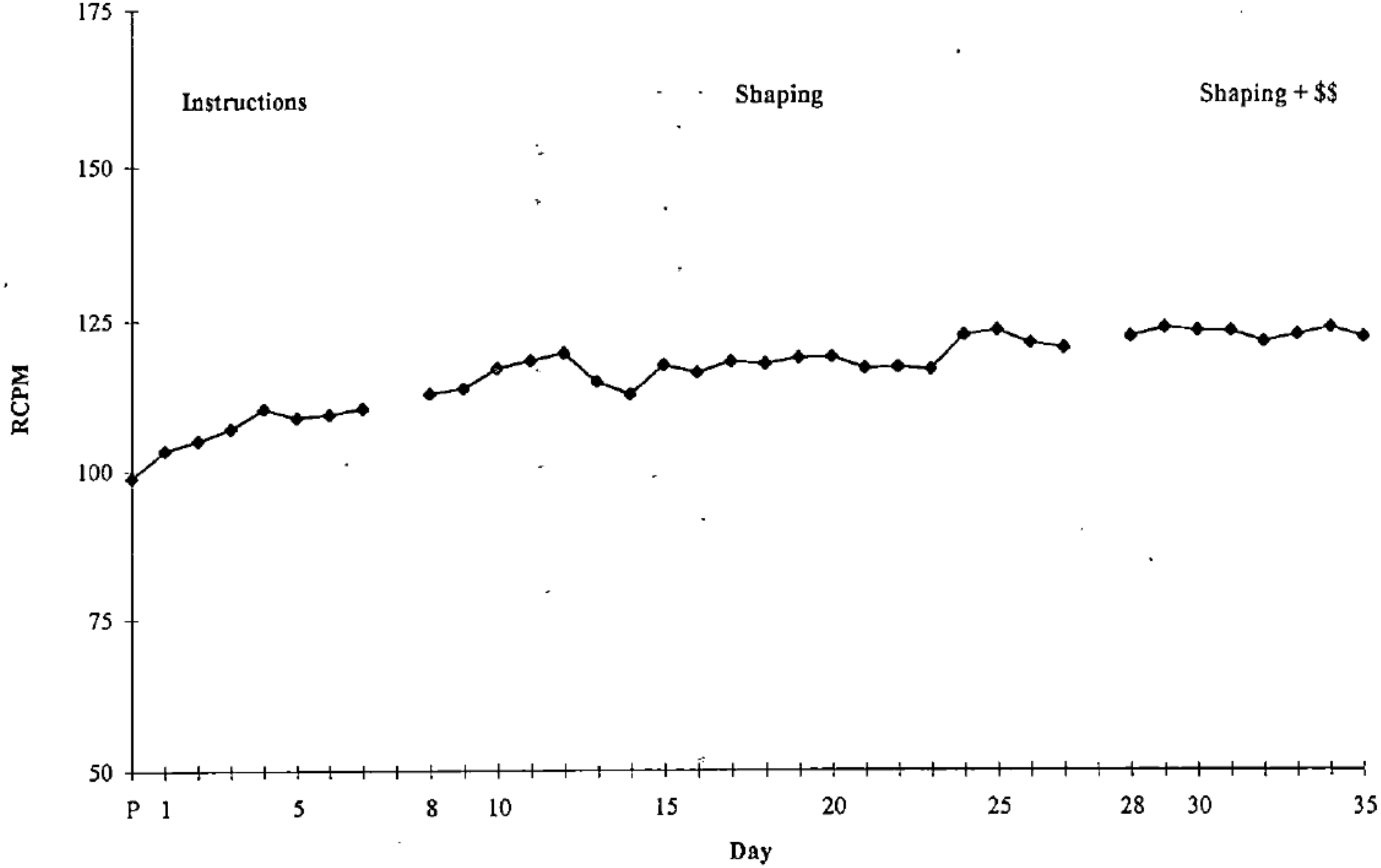
Subject N01.

Pretest. The subject scored 100% on all three trials, therefore functional equivalence between training and generalization stimuli was demonstrated under non-speed conditions.

Phase A - Instruction-based training. Subject N01's daily RCPM data are located in Figure 1. From the initial training session (P) to the end of the first full day of training (1), her RCPM increased by almost 5 correct-responses-per-minute (for convenience, r/m will be used throughout the remainder of this paper), from 98.7 to 103.3 r/m. Performance continued to improve over the next three training days, and by the end of day 4, RCPM had increased to a high of 110.3 r/m. Over the remainder of this initial training phase, RCPM varied between 108.8 and 110.3 r/m. Over the seven full days of instruction-based practice, RCPM had increased by 11.7 r/m. Similar to data reported by CRT researchers, greatest relative improvement occurred very early in this training phase (i.e., from session P to the end of the first training day).

Phase B - Shaping speeded responding. Subject N01's shaping data were by far the most unique and interesting data collected in this experiment. When shaping was introduced, RCPM was 112.8 r/m, an increase of 2.4 r/m over the terminal Phase A score. Over the next four training days, RCPM showed a consistent, increasing trend, such that by day 12, it had increased to 119.5 r/m. Note that the data collected during latter Phase A suggested, if anything, that performance was, perhaps, levelling off, yet when shaping was introduced, performance improved almost as much during the initial four days as it did over the entire instruction-based phase. During this period, the latency criterion was reduced from 577 ms to 533 ms. More precisely, an average of 10 ms, spread over two training sessions, were subtracted from the subject's latency criterion

Figure 1. Subject N01's daily rate-correct-per-minute (RCPM) data. Each data point (except the first of each phase) represents three 15-trial sessions.


175
150
125
100
75
50
RCPM
Instructions
Shaping
Shaping + $$
P
1
5
8
10
15
20
25
28
30
35
Day

each day.

On day 13, the importance of being able to both increase and decrease the latency criterion was well illustrated. On session 1, the subject scored 90% C&F, suggesting that the present latency criterion could be made more stringent. As a result, a further 8 ms were subtracted from the latency criterion, making the task, once again, more difficult. This amount of change, which had been easily mastered on previous training days, was suddenly too large, and resulted in poor performance by the subject. Her % C&F score, for example, decreased to 78, and this was the first time it had fallen below the minimum requirement of 85. Furthermore, the subject, for the first time, reported finding the task "frustrating." Such emotional responses typically accompany training procedures where insufficient reinforcement is available (Martin & Pear, 1988, p. 60). Subtracting 8 ms from the latency criterion had apparently made the task too difficult for the subject, thus a reduction in the stringency of the latency criterion appeared necessary.

At this point, the experimenter modified the latency adjustment procedure such that the criterion could never be made less stringent than the most recent criterion mastered by the subject. This seemed reasonable given the subject's prior success at that latency criterion. The following is an illustration of the modified latency-adjustment procedure. Given the following information: the overly-stringent latency criterion was 535 ms, the previously-mastered criterion was 550 ms, and the criterion produced by the original latency-adjustment procedure was 540 ms, the experimenter would have adjusted the latency criterion to 540 ms. However, if the original latency-adjustment procedure produced a criterion of 555 ms, then the criterion would have been adjusted to 550 ms. As a result of this change in procedure, subject N01's latency criterion was returned to 533 ms on session 3. Subject N01's performance during session 2 (not observable due to data averaging), under this once manageable latency criterion, was, however, surprisingly poor. That is, her % C&F score was only 82, which was far below the 90 she had obtained on session 1, under that same latency criterion. Once again, the subject reported that she found the task frustrating. Returning to the RCPM graph, subject N01's data clearly illustrated that performance had declined. In fact, by the end of day 13, RCPM had decreased by 4.8 r/m, to 114.7 r/m, a result of both decreasing accuracy and increasing response latency.

Training was cancelled for the next three days, because of the reading-week break. Given the subject's poor performance on day 13, and the long break from training, the experimenter decided to further decrease the stringency of the latency criterion on day 14, to give the subject experience, once again, with a high-reinforcement training session. After all, failing to do so might have resulted in further weakening of the behavior. The latency criterion chosen was 541 ms, a criterion that subject N01 had previously mastered in two sessions. This time, however, her initial % C&F score was 85, which was 2% C&F less than when she first encountered this latency criterion on the third session of day 11. Subject N01's performance failed to improve over the next two sessions, and by day 14, her RCPM had declined to 112.6 r/m, which was marginally lower than her initial Phase B RCPM score.

Over the next few training days, performance slowly began to recover, as % C&F fluctuated between 88-90. On day 17, subject N01 scored 90% C&F in session 1. Instead of immediately adjusting the criterion, the experimenter chose to continue the subject with this same latency criterion, again, in an attempt to increase her exposure to a highly reinforcing training sessions. However, the subject's % C&F score fell below 90 on the very next session.

On session 2 of day 18, five milliseconds were added to the criterion (546 ms) to make it even less stringent. A further five milliseconds were added on session 3 (551 ms), when recovery of rapid responding was simply not forthcoming, and because the subject continued to report elevated levels of frustration. When the latency criterion was adjusted to 551 ms, the subject scored 91% C&F. After a further four sessions of practice with this latency criterion (% C&F scores varied between 87-90), the experimenter decided to resume increasing the stringency of the latency criterion. Although, the subject did not score 90% C&F or above on all sessions, it had become evident that multiple, consistent sessions with % C&F scores above 90 might only be achieved by significantly decreasing the stringency of the latency criterion. Instead, the experimenter chose to begin slowly increasing the stringency of the latency criterion, using a further-modified latency-adjustment procedure. Here, when the subject scored 90% C&F or greater, only very small (i.e., 5 ms) adjustments were made to the criterion.

Over the next seven training days, small adjustments to the latency criterion were made, typically every other day. The final latency criterion was 529 ms, which was 4 ms less stringent than the criterion that was in place when the behavior was initially disrupted.

Despite the difficulties encountered by this subject, RCPM began to slowly recover, after reaching the low of 112.6 r/m on day 14. From days 15 through 23, subject N01's RCPM scores fluctuated between 116-119 r/m. On day 24, RCPM increased to a new high of 122.4 r/m, a result of an increase in accuracy, which had been on a progressively declining trend since day 16. RCPM increased again on day 25, to 123.2 r/m, this time an apparent result of decreasing response latency, and then decreased slightly over the remaining two training days, finishing the phase at 120.3 r/m.

<u>Assessing the effects of the shaping procedure using logarithmic-transformed RCPM data.</u> Up to this point, only simple, visual inspection of trends across various performance measures has been used to determine whether subjects attained any benefit by switching from instruction-based to a contingency-based procedures. However, according to the power law of learning, which applies to reaction time data, large practice effects develop early, followed by progressively smaller performance gains as training progresses (Newell & Rosenbloom, 1981). What this means is that as training progresses, a subject makes progressively smaller improvements in response speed until, presumably, he or she meets his or her physiological limit. Given that subjects are expected to

make progressively smaller performance gains as training continues, it is difficult to compare the relative effectiveness of different training procedures, using visual inspection of the raw, non-linear data as the sole method of analysis. Clearly, it is not appropriate to predict comparable improvements in performance across training phases, because most improvement will necessarily have occurred very early during the initial, instruction-based training phase. The difficulty, then, lies in determining how much improvement the subject would have been expected to make during the contingency-based phase, had she simply remained under instruction-based conditions, rather than switching into the contingency-based phases. Of interest, was whether the subject's improvement during shaping was of a magnitude greater than was expected under continued instruction-based practice. Given the data in their present form, analysis via visual inspection alone could not provide a satisfactory answer to this question, because practice effects were confounded with the order of the training phases.

Baron et al. (1983) chose to convert response latencies to speeds (one being the reciprocal of the other), on the assumption that "increasing weight should be given to latency changes as latencies decreased" (p. 279). Although they seemed to acknowledge the fact that treatment and duration of practice were confounded, their choice of transformation did not entirely remove or adequately control for the confound (M. Masson, personal communication, September 14, 1995).

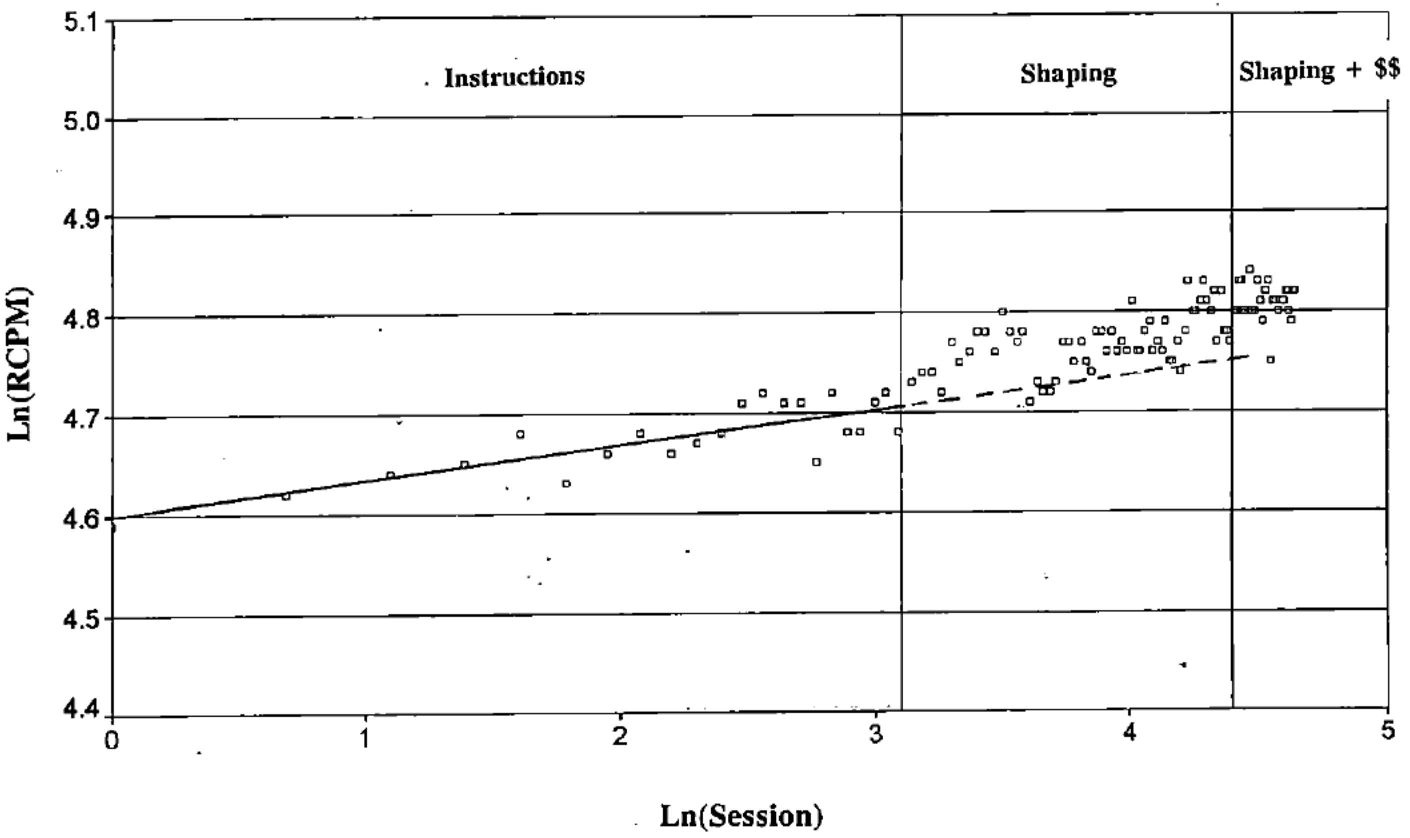
Newell and Rosenbloom (1981) described an interesting and useful relationship that exists between various performance measures and practice time (e.g., training day, session or trial, etc.), when logarithmic transformations of each measure are taken, and these transformed measures are then graphed. Reaction time is studied most often, but see Stevens and Savin (1962) for examples of several, non-reaction time performance measures which behave similarly. Logarithmic transformations compress differences between larger numbers, resulting in the linear transformation of once non-linear data. According to Kadlec (H. Kadlec, personal communication August 18, 1995), log-log transformations are commonly used to make non-linear data linear. If the present non-linear data were made linear, it would be possible to make an assessment of the relative contribution of the contingency-based procedure. This assertion is based upon the premise that if training under the contingency-based procedure has had no effect upon performance, then the resultant data should follow the same linear trend as that collected under instruction-based conditions. If, alternatively, the contingency-based procedure has affected performance on the training task, then the resultant data should deviate from the linear trend which describes the instruction-based data. In summary, one way of assessing the relative contributions of shaping, would be to fit a regression line through the linearly-transformed, instruction-based data, project this regression line through the linearly-transformed contingency-based data, and assess whether the shaping data meaningfully deviate from the instruction-based regression line.

A procedure similar to that described above was used in the present experiment. To assess whether performance during shaping differed from that which was predicted under continued instruction-based practice, logarithmic transformations of each subject's, per session, rate-correct-per-minute (RCPM) data were made and plotted against the logarithmic transformation of the respective session number. Next, a regression line was fit through the instruction-based data, and provided that the regression line fit the data moderately well to well, the regression line was projected through the shaping data. The goodness-of-fit of the regression line was determined by looking at the adjusted, squared correlation (which will be referred to as adjusted $\underline{r}^2$ throughout the rest of this paper) associated with each regression line. The adjusted $\underline{r}^2$ indicates the proportion of variation in the dependent variable accounted for by the independent variable (Tabachnick & Fidell, 1989). For example, an adjusted $\underline{r}^2$ score of 90 indicates that 90% of the variability in the dependent variable (i.e., natural logarithm of RCPM scores) is accounted for by the independent variable (i.e., natural logarithm of the session number). In the present experiment, regression lines with adjusted $\underline{r}^2$ values of 85 or greater were considered to fit the data well, while regression lines with values below this, provided they were above 50, were considered to fit the data only moderately well. Upon completion of these important preliminary analyses, an assessment of whether the shaping data deviated meaningfully from the instruction-based data was made. The natural logarithmic transformation was chosen here, however, other logarithmic transformations could have been used as well (Newell & Rosenbloom, 1981).

The RCPM data were chosen for this analysis for several reasons. First, because RCPM scores are computed from both accuracy and latency information, they are arguably the most representative performance measures available, and, thus, the most appropriate measure for this analysis. Second, when regression lines were fit through the transformed, RCPM data, the fit of these lines was generally very good (i.e., typically between 80-90% of the variance in the data was explained by the perspective regression line). Third, as alluded to earlier, Stevens and Savin (1962) successfully completed log-log transformations of non-reaction time data. Some of the measures they studied were similar to the present rate measure (e.g., number of correct responses per minute or hour; number of correct responses across trial, where trial length is a constant), which suggested that the RCPM measure could appropriately be used. This analysis was completed using session RCPM scores rather than daily RCPM scores to allow the data collected on the first day of each phase to be used in the analysis, and to increase the number of data points (i.e., three RCPM scores per day were used rather than just one), to fit the most accurate regression line possible.

Subject N01's Transformed RCPM data. Subject N01's log-log transformed, RCPM data are located in Figure 2. First, the adjusted r^2 was only 64.9, therefore these data were only moderately well fit by the regression line. Subject N01 trained least under instruction-based conditions, and thus fewer data points were available for the analyses relative to other subjects. This may have

Figure 2. Subject N01's per session, transformed (natural logarithm) rate-correct-per-minute (RCPM) data, plotted against the transformed (natural logarithm) session number. Each phase is separated by a vertical line, and each data point represents fifteen 20-second trials of practice. A regression line was fit through the Phase A (instruction-based) data, and projected (dotted line) through the Phase B (shaping) data.


Instructions
Shaping
Shaping + $$
Ln(RCPM)
5.1
5.0
4.9
4.8
4.7
4.6
4.5
4.4
0
1
2
3
4
5
Ln(Session)

been responsible, in part, for the moderate goodness-of-fit score. Notice that prior to the introduction of shaping, subject N01's transformed RCPM scores fell both above and below their regression line, thus the regression line appeared to adequately describe terminal, Phase A performance. However, upon introduction of shaping, her transformed RCPM scores were consistently above (and by the 5th session were meaningfully above) this regression line. These data suggested, then, that greater improvement in performance than that which would have been predicted under continued instruction-based practice resulted when shaping was introduced. The largest, initial improvement in performance under shaping was found with this subject, although this, too, may have resulted from her having been exposed the least to Phase A conditions. On day 13, when the latency criterion was made too stringent, subject N01's performance fell sharply. Here, for the first time, her transformed RCPM scores approached the instruction-based regression line. More explicitly, her RCPM scores clustered around the Phase A regression line for the first five sessions, and then slowly, they began to occur above the line, although not to the same degree as that found initially. Near the end of the shaping phase, subject N01's performance approached the level attained earlier in shaping, however performance was not consistently high as it had been during initial shaping.

To summarize, when shaping was introduced, the subject demonstrated that further improvement was possible. Initially, she improved more quickly than would have been expected had she remained under the instruction-based procedure. However, when too large of an adjustment was made to the latency criterion, the behavior was partially lost and required retraining. By the end of the shaping phase, the subject was able to achieve a nearly comparable level of performance to that found earlier during the phase. Most importantly, this subject's performance illustrated the importance of having a latency adjustment procedure which allowed the stringency of the latency criterion to be either increased or decreased, depending upon the subject's performance. Having completed a more thorough analysis of the shaping data, by studying both subject N01's raw and transformed RCPM data, it was now possible to move on to looking at the effects of adding monetary reinforcement to the shaping procedure.

Phase C - Shaping-plus-monetary-reinforcement. When monetary reinforcement was introduced, the subject's accuracy increased immediately. This resulted in a slight increase in RCPM, which had increased to 122.1 r/m. Over the first two training days of this phase, subject N01's % C&F scores were consistently above 90, thus the stringency of the latency criterion was increased on all sessions where adjustments could be made. Note that this subject, who had experienced great difficulty when adjustments were made to the latency criterion during Phase B, was now having little difficulty meeting the constantly changing latency criterion. Also, during this time, the subject reported that the task had become "enjoyable." Surprisingly, however, a clear increase in RCPM did not appear to accompany the higher % C&F values. This may have occurred for two reasons. First, while accuracy increased, the subject's mean latency correct scores appeared slightly longer, suggesting that the additional responses that met the latency criterion were not faster than average, but, rather, may have just been meeting the latency criterion, causing the average mean latency correct score to increase slightly. Indeed, the latency on 100% accuracy trials data suggested that performance had slowed over these initial Phase C training days. Also, given that each training trial was only 20 seconds long, a seemingly large increase in % C&F might have been the result of only a single additional response meeting the latency criterion. Subject N01 achieved her highest RCPM score on day 34, at 123.6 r/m, which was only negligibly higher than her day 25 (Phase B) value. Most often during this phase, her RCPM score varied between 121.3 and 123 r/m. Thus when looking at the raw RCPM data, little improvement appeared to have occurred as a result of introducing monetary reinforcement to the shaping procedure. Turning, briefly, to the transformed RCPM data, visual inspection of the shaping-plus-monetary-reinforcement data suggested slight improvement over latter performance under shaping. Notice the continued increasing trend in combination with much less variable performance across sessions, as illustrated by the increased clustering of shaping-plus-monetary-reinforcement data points.

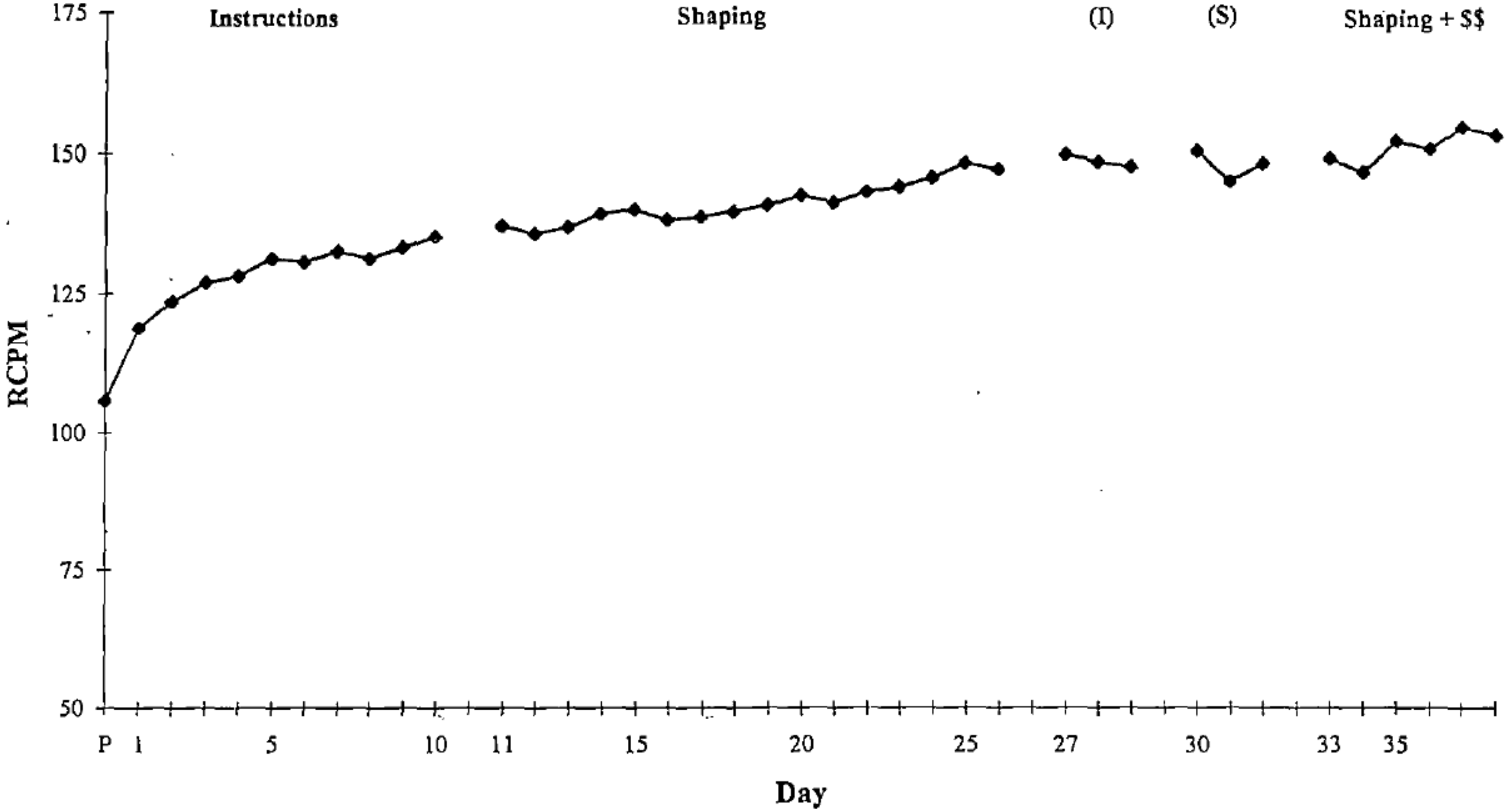
Summary of subject N01's training data. To summarize, large initial improvements in performance occurred early in training, under instruction-based conditions. When shaping was introduced, subject N01 not only continued to improve, but she improved more than was predicted under continued instruction-based practice. Performance was disrupted on day 13, when the latency criterion was made too stringent, and, thereafter, the subject struggled to re-establish her high level of performance. Despite the disruption in performance, subject N01's transformed RCPM scores remained almost consistently above their predicted value. When monetary reinforcement was introduced, subject N01 began responding more accurately, and because response latency was unaffected by the change, her % C&F scores improved. Although not clear in the raw RCPM data, subject N01's transformed RCPM data hinted at continued improvement under shaping-plus-monetary reinforcement conditions.

Subject N02

Pretest. The subject scored 100% on two of the trials and 94.4% (i.e., made one error) on the third, therefore it seemed reasonable to conclude that functional equivalence had been demonstrated between the training and generalization stimuli.

Phase A - Instruction-Based Training. Referring to Figure 3, subject N02's initial, session P RCPM score was 105.6 r/m. By the end of the first full training day, her RCPM score had increased by 13 r/m to 118.6 r/m, with only a 2% decrease in accuracy. Subject N02's RCPM continued on an increasing trend over the next 4 training days, and by day 5 had climbed to 131.3 r/m. Subject N02's RCPM continued to increase, although to a much lesser degree, during the later half of this phase, and on the final training day, reached a high of 135.1 r/m. Similar to subject N01, then, greatest improvement occurred very early in training, although contrary to subject N01, subject N02 appeared to continue to improve

Figure 3. Subject N02's daily rate-correct-per-minute (RCPM) data. Each data point (except the first of each phase) represents three 15-trial sessions.

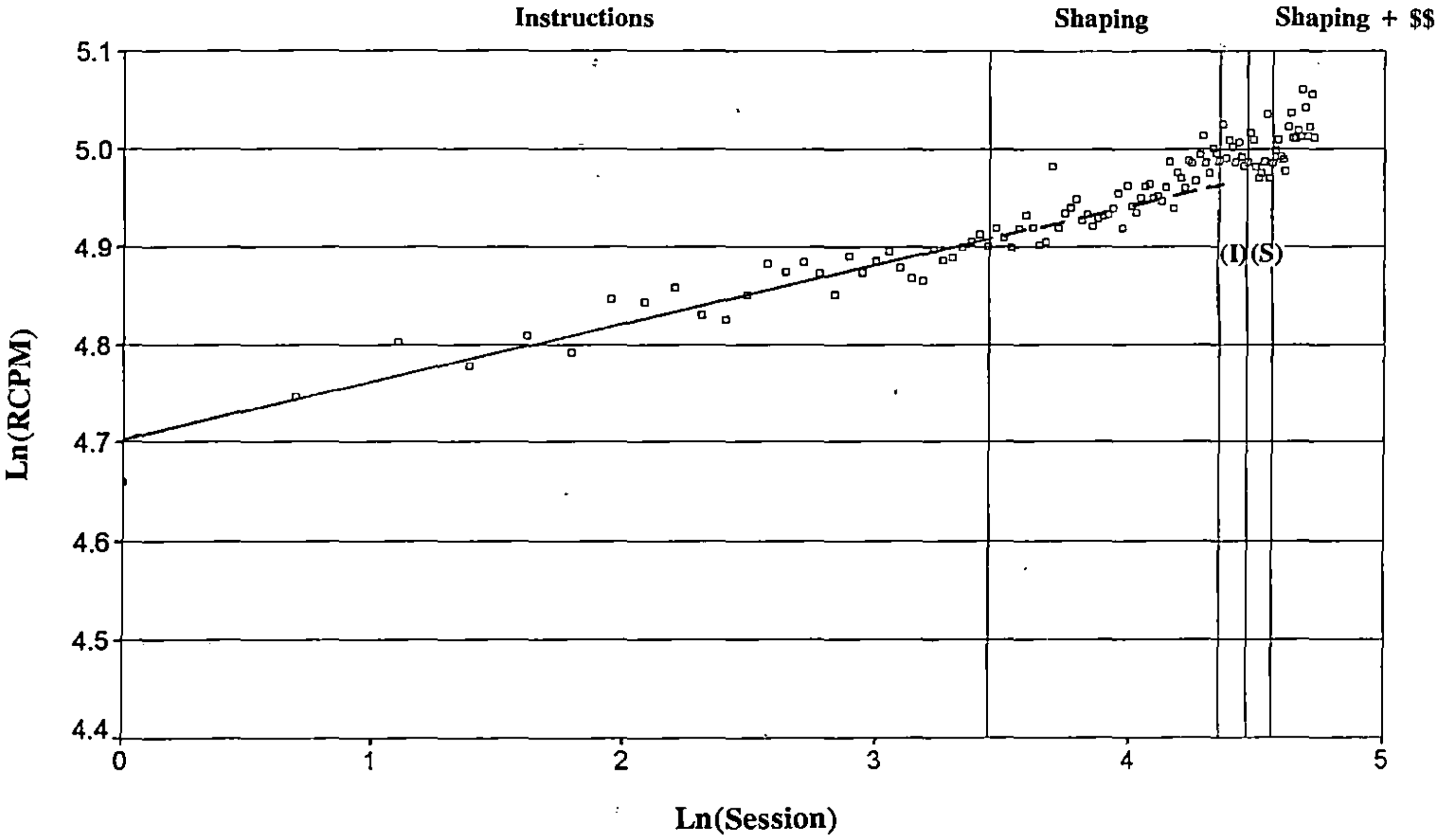

Instructions
Shaping
(I)
(S)
Shaping + $$
RCPM
175
150
125
100
75
50
P
1
5
10
11
15
20
25
27
30
33
35
Day

throughout the entire training phase.

Phase B - Shaping speeded responding. When shaping was introduced, RCPM increased slightly, to 136.9 r/m. Over the next couple of days, RCPM changed little, despite a 9 ms increase in the stringency of the latency criterion. However, on day 14, RCPM increased by almost 3 r/m, finishing at 139.2 r/m. Over the next four training days, RCPM varied between 138 and 139.5 r/m. On day 19, subject N01's RCPM surpassed the 140 r/m mark. Over the remaining seven training days, RCPM continued to increase steadily, reaching a high of 148.1 r/m on the second to last day, and finishing at 147.1 r/m. Subject N02's % C&F score never fell below 85, thus allowing experimenter to use the original latency adjustment procedure. On average, the stringency of subject N02's latency criterion was increased by between five and 10 ms every day or so.

Subject N02's transformed RCPM data. Subject N02's transformed RCPM data is located in Figure 4. First, note that the adjusted r^2 value associated with the instruction-based regression line was 88.84, thus these data appeared to be well fit by the regression line. When shaping was introduced, consistent with her raw RCPM data, subject N02's transformed RCPM data suggested continued improvement, though not of a magnitude greater than predicted under continued instruction-based practice. In fact, for the most part, subject N02's early shaping data appeared well fit by the regression line which described the earlier, instruction-based data. Around day 22, however, subject N02's transformed RCPM scores were occurring exclusively above the line, and during the latter

Figure 4. Subject N02's per session transformed (natural logarithmic) rate-correct-per-minute (RCPM) data plotted against the transformed (natural logarithmic) session number. Each phase is separated by a vertical line, and each data point represents fifteen 20-second trials of practice. A regression line was fit through the Phase A (instruction-based) data and projected (dotted line) through the Phase B (shaping) data.


Instructions
Shaping
Shaping + $$
(I)
(S)
Ln(RCPM)
5.1
5.0
4.9
4.8
4.7
4.6
4.5
4.4
0
1
2
3
4
5
Ln(Session)

third of this training phase, her performance appeared to be meaningfully better than was expected under continued instruction-based practice. Thus although shaping had little initial effect on performance, it appeared to be responsible for better-than-predicted performance during the latter part of this training phase.

Return to instruction-based training. Instruction-based conditions were reintroduced for eight sessions to see what effect the removal of the latency criterion would have on the subject's performance. On the initial training session of this phase, subject N02's RCPM increased to its highest level ever at 152.2 r/m, and then returned to a level more comparable to terminal shaping performance, at 147.1 r/m. This was well illustrated in both the raw data, where the day average was 149.7 r/m, her personal best to date, and in the transformed RCPM data, where her initial data point was much higher than subsequent data points. Thereafter, subject N02's RCPM appeared to be on a gradually declining trend, suggesting that, in the absence of the shaping procedure, additional training might have led to further decrements in performance. Although caution must be exercised when attempting to draw definitive conclusions based upon these reversal data, because of the brevity of the phase, it was interesting and worthwhile to speculate. After all, these data, too, offer potentially valuable information about the relative effectiveness of the shaping procedure.

Return to shaping speeded responding. To prepare the subject for Phase C and establish whether any further performance gains could be made, shaping was briefly reintroduced. The initial latency criterion chosen was the last criterion used prior to the reintroduction of instruction-based training. This appeared to be an appropriate latency value to use, because subject N02 had scored 89% C&F previously under that latency criterion, and, again, the experimenter was interested in re-establishing the subject's high level of performance in preparation for the final training phase. Initially, the subject appeared to have improved over her latter Return-to-Phase-A performance, as her initial RCPM score was 150.9 r/m, the second highest to date. However, similar to the previous phase, subsequent performance fell below (late) initial-Phase-B and return-to-phase-A levels. This trend was also evident in the transformed RCPM data, where a decrement in performance was clearly illustrated. It is difficult to speculate on why a decrement in performance resulted from the reintroduction of shaping. The latency criterion was not changed for the first seven sessions of this training phase, and subject N02's % C&F scores always remained above 85, suggesting that the latency criterion was manageable for the subject. Furthermore, this subject did not suggest that the task had become frustrating. However, she did say that she "believed her RCPM score had gotten as high as it was going to get." It was, once again, inappropriate to speculate further because of the short duration of this training phase.

Phase C - Shaping-Plus-Monetary-Reinforcement. Similar to Subject N01, when monetary reinforcement was introduced, rather than responding more quickly, subject N02's accuracy increased by 2%, which resulted in a slight increase in RCPM over the terminal, Return-to-Phase-A value. Performance on the second day of this phase was relatively poor, as subject N02's RCPM decreased to 146.4 r/m. However, on the third day of this phase, performance improved such that subject N02's RCPM was 152 r/m, a new high. On this day, her latency criterion was decreased by 9 ms. Throughout the remainder of this phase, performance remained strong, and subject N02 reached a new personal-best RCPM score of 154.3 r/m on the second to last day of training. Her final RCPM score was (a very respectable) 153 r/m.

Overall summary of subject N02's training data. To summarize, greatest improvement occurred very early, under instruction-based conditions, followed by continued, though more gradual improvement, throughout the rest of this training phase. Initially, shaping did not appear to affect performance in a meaningful way, although accuracy did increase slightly. However, near the end of shaping, subject N02 was performing consistently above the level predicted under continued instruction-based practice. Also, there was some suggestion that the removal of the shaping procedure might have, in the long-term, resulted in decrements in performance. However, after a few training sessions under return-to-shaping conditions, performance, again, began to decline. Unfortunately, these reversal sessions appeared to be too brief to allow definitive conclusions to be drawn from the data. When monetary reinforcement was introduced, the subject initially responded more accurately, which resulted in only a small increase in RCPM. However, by day three of this final training phase, performance was, once again, improving, and the subject reached a high of 154.3 r/m on the second-to-last day of training.

Category Subjects

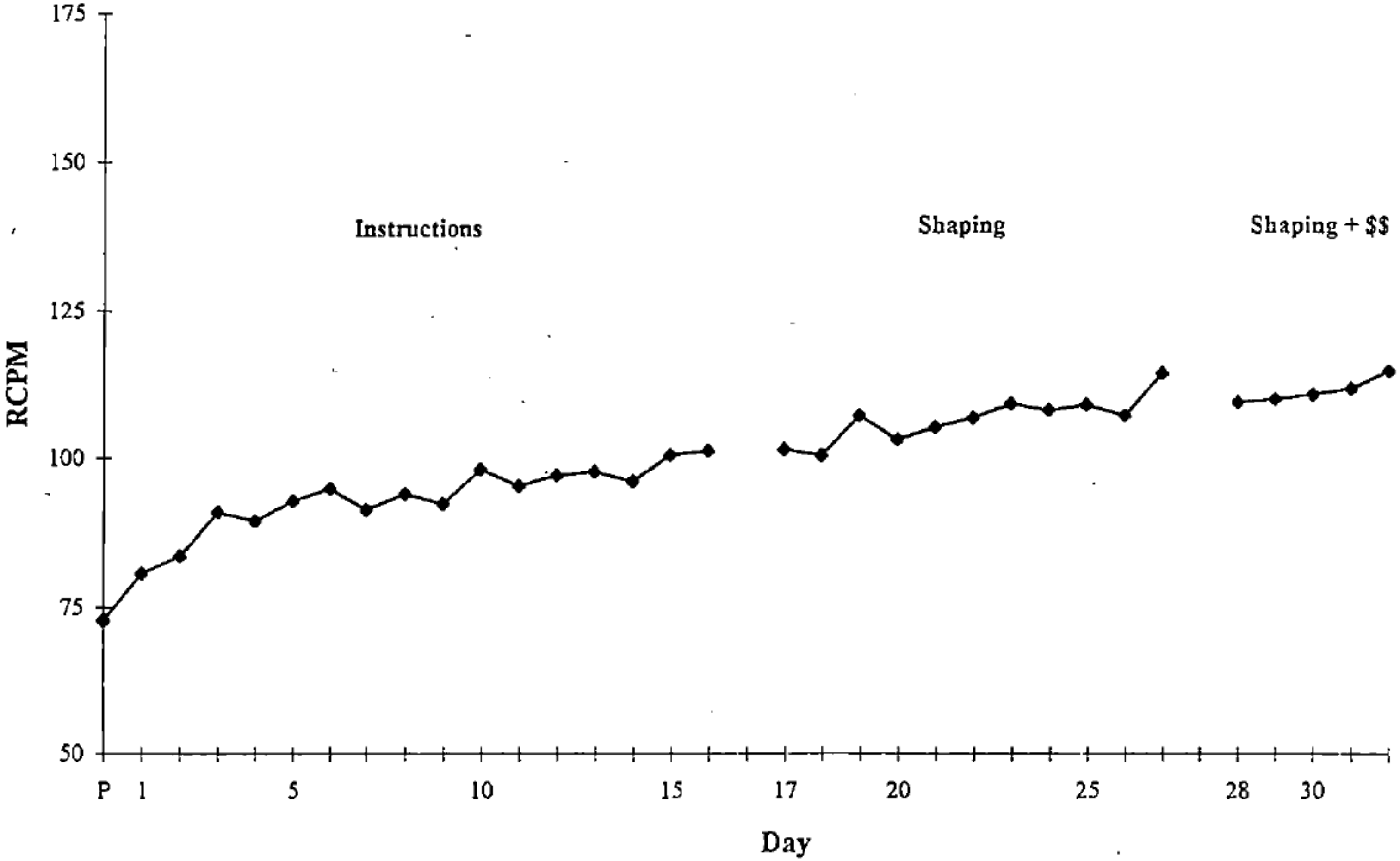
Subject C01.

Pretest. On the first two trials, the subject scored 100%, while on the final trial, she scored 92.3% (i.e., one error was made). Again, given only one error, it seemed appropriate to conclude that functional equivalence of the training and generalization stimuli had been demonstrated.

Phase A - Instruction-based training. As illustrated in Figure 5, subject C01's initial (day P) RCPM was 72.8 r/m. By the end of the first full day of training, her RCPM had increased by almost 8 r/m, to 80.6 r/m. Over the next two training days, subject C01's RCPM score increased a further 10 r/m, finishing at 90.9 r/m. From day 4 until the end of this initial training phase, the subject's RCPM continued to improve, though much more gradually, reaching a high of 101.1 r/m by the end of this training phase. Similar to the previous two subjects, greatest improvement occurred very early in training. Furthermore, like subject N02, subject C01 continued to improve, although more gradually, over the remainder of this initial training phase.

Phase B - Shaping speeded responding. When shaping was first introduced, subject C02's RCPM was comparable to terminal Phase A values. However, by the third day of shaping, gradual improvement was, once again, evident, as subject C02's RCPM score increased to 107.1 r/m. RCPM decreased to 102.9 r/m on the next training day, and, thereafter, began increasing, once

Figure 5. Subject C01's daily rate-correct-per-minute (RCPM) data. Each data point (except the first of each phase) represents three 15-trial sessions.

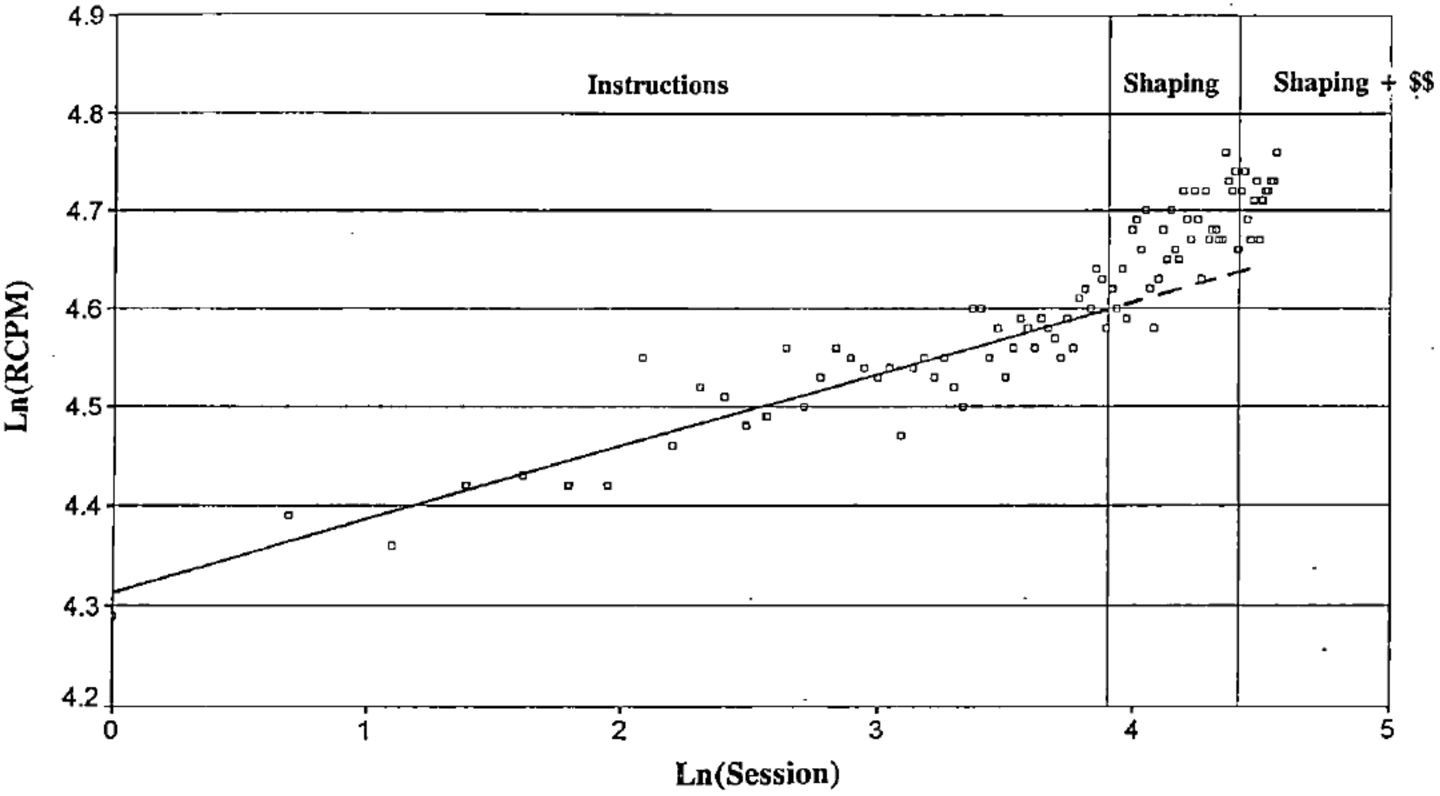

175
150
125
100
75
50
RCPM
Instructions
Shaping
Shaping + $$
P
1
5
10
15
17
20
25
28
30
Day

again. From days 23-26, RCPM varied between 107 and 109 r/m; on day 27, RCPM increased by 7 r/m, to 114.2 r/m. The reader will note that this subject's accuracy was relatively stable across this phase, therefore increasing RCPM was, for the most part, a result of decreasing latency.

Subject C01's transformed RCPM data. Early in the shaping phase, subject C01's transformed RCPM data (see Figure 6) were initially well fit by the instruction-based data regression line (adjusted $\underline{r}^2$ was 82.84%). Soon, however, her performance rose to well above the predicted level, and unlike subject N01, subject C01 continued to perform well above predicted levels throughout the rest of this training phase. The shaping procedure thus appeared to be superior to the instruction-based procedure in facilitating fluent responding, at least with this subject.

Phase C - Shaping-plus-monetary-reinforcement. Unlike with the previous two subjects, when monetary reinforcement was first introduced, subject C01's accuracy did not increase. Rather, subject C01's RCPM decreased from her personal best of 114.2 r/m to 109.4 r/m, a value more comparable to scores collected on most of the latter shaping day scores. The subject reported that filling in the earnings sheet after every trial was "disruptive." This might have been partially responsible for the early decrement in performance. She was subsequently allowed to fill out the earnings sheet at the end of each session. Throughout the remainder of this phase, accuracy was relatively stable (between 91.3-92.7%), and RCPM began increasing. This improvement was similar in

Figure 6. Subject C01's per session, transformed (natural logarithmic) rate-correct-per-minute (RCPM) data plotted against the transformed (natural logarithmic) session number. Each phase is separated by a vertical line, and each data point represents fifteen 20-second trials of practice. A regression line was fit through the Phase A (instruction-based) data and projected (dotted line) through the Phase B (shaping) data.


Instructions
Shaping
Shaping + $$
4.9
4.8
4.7
4.6
4.5
4.4
4.3
4.2
0
1
2
3
4
5
Ln(RCPM)
Ln(Session)

nature to that found during the latter half of the previous training phase. Subject C01 reached a new RCPM high score of 114.6 r/m on the last training day. Overall, subject C01 showed little benefit from completing this final training phase.

Overall summary of subject C01's training data. To summarize, large, initial practice effects developed over the first several days, followed by continued though more gradual improvement throughout the remainder of training Phase A. When shaping was introduced, subject C01 did not respond differently to the CRT task than would have been predicted by her instruction-based data. However, by the third shaping day, her performance was almost consistently well above the level predicted by her former data, suggesting that the shaping procedure was superior to the instruction-based procedure for facilitating fluent responding. Finally, the addition of monetary reinforcement to the shaping procedure appeared to disrupt the subject's initial performance. However, by the end of the phase, performance had recovered, and the subject scored a personal best RCPM score of 114.6 r/m. The subject suggested that filling out the earnings sheet on a trial-by-trial basis, and not the addition of monetary reinforcement, was responsible for the initial decrement in performance.

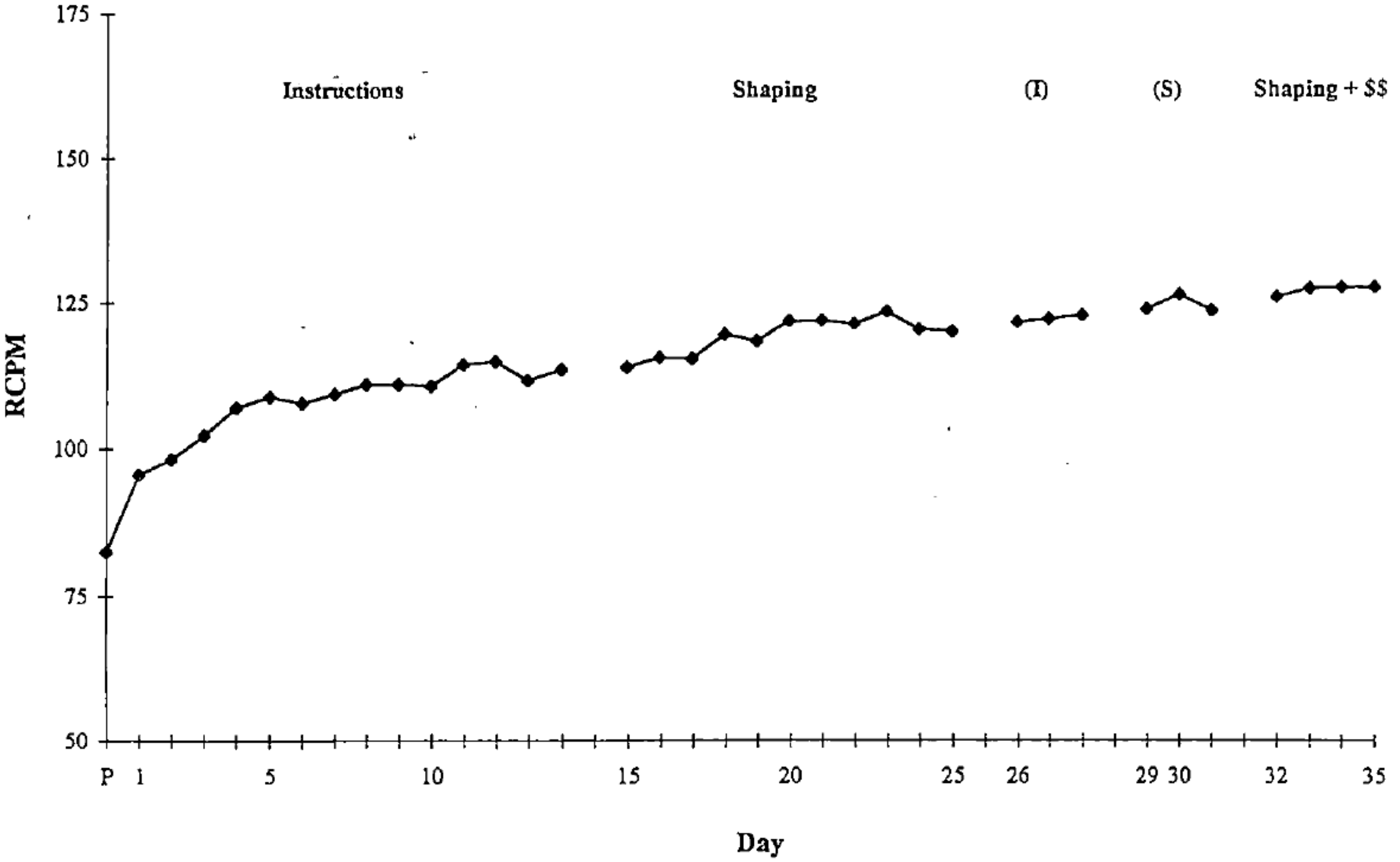
Subject C02

Pretest. Similar to the previous two subjects, subject C02 scored 100% on two of the three trials. However, she scored 87.5% (2 errors were made) on the final trial. Given that two of her trials were error-free, and, overall, she responded correctly to 41 out of 43 stimuli, it seemed reasonable to conclude that functional equivalence had been demonstrated between training and generalization stimuli.

Phase A - Instruction-based training. As illustrated in Figure 7, subject C02's initial RCPM score was 82.6 r/m. At the end of the first full day of training, RCPM had increased by 13 r/m, a result of a large decrease in response latency. Steady though more gradual improvement occurred over the next four training days, resulting in an increase in RCPM to 108.7 r/m. Over the next five training days, RCPM appeared fairly stable at around 110 r/m. Then on training days 11 and 12, subject C02's RCPM increased to 114.2 r/m. Her RCPM finished at 113.3 r/m on the last day of instruction-based practice. Similar to subjects N02 and C01, then, large performance gains were made by subject C02 very early in this initial training phase, followed by consistent though more gradual improvement throughout the remainder of Phase A.

Phase B - Shaping speeded responding. When shaping was introduced, RCPM became only marginally higher relative to terminal Phase A values, at 113.7 r/m. Similarly, over the next couple of training days, RCPM increased by only 2 r/m. Thus it appeared that shaping had little initial affect upon performance. More noticeable improvements were made on day 18, when RCPM increased from 115.5 to 119.3 r/m, and again on day 20, where subject C02 was now responding at 121.7 r/m. Subject C02 reached a personal best RCPM score of 123.3 r/m on day 23. However, over the

Figure 7. Subject C02's daily rate-correct-per-minute (RCPM) data. Each data point (except the first of each phase) represents three 15-trial sessions.

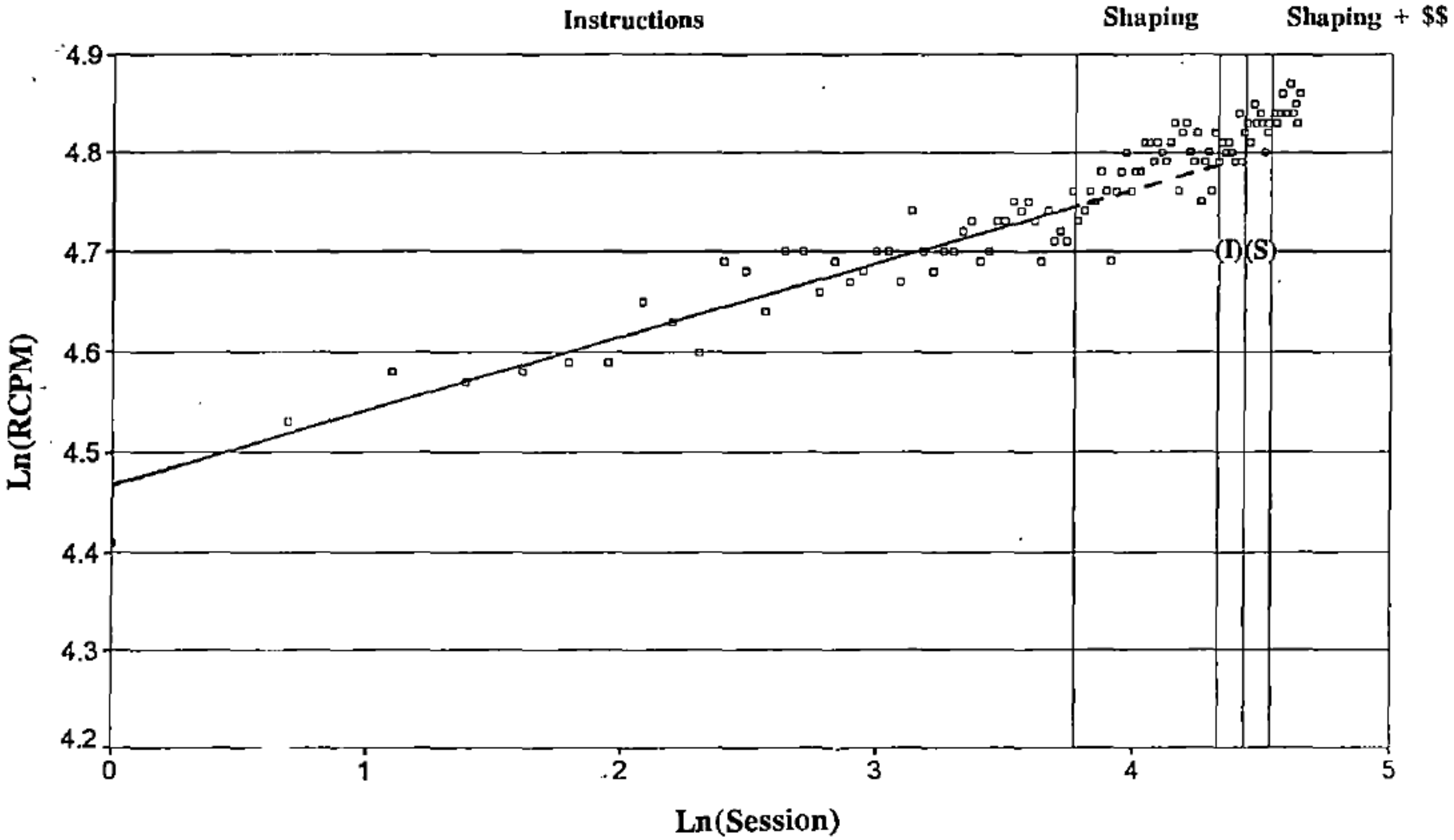

Instructions
Shaping
(I)
(S)
Shaping + $$
RCPM
175
150
125
100
75
50
P
1
5
10
15
20
25
26
29
30
32
35
Day

remaining two shaping days, RCPM was on a declining trend, finishing at 119.8 r/m on day 25. This decrease in performance during latter shaping appeared to be a result of increasing response latency. The latency criterion, which had been changed two days earlier, did not appear to be overly stringent, because on the first full day of training using that particular latency criterion, subject C02's % C&F scores ranged between 86-89. Thus, something other than the training procedure appeared to be responsible for this latter decrement in performance. The subject did comment on day 24 that she was "tired," thus fatigue might have been partially responsible for these results.

Subject C02's transformed RCPM data. Similar to the data of subjects N01 and C01, subject C02's transformed RCPM data (see Figure 8) were initially well fit by the instruction-based regression line (adjusted r^2 was 87.72). About one third of the way into this phase, level of performance increased such that subject C02's transformed RCPM scores were now consistently well above the predicted level, and, this high level of performance continued over the next third of training. However, as discussed earlier, performance decreased somewhat during the latter shaping sessions, such that transformed RCPM scores began, once again, approaching the instruction-based regression line.

Return to instruction-based training. As with subject N02, subject C02 was briefly returned to instruction-based conditions, to assess what effects the removal of the shaping procedure might have on performance. Unlike subject N02, subject C02's RCPM score on the first session of this reversal phase was not meaningfully

Figure 8. Subject C02's per session, transformed (natural logarithmic) rate-correct-per-minute (RCPM) data plotted against the transformed (natural logarithmic) session number. Each phase is separated by a vertical line, and each data point represents fifteen 20-second trials of practice. A regression line was fit through the Phase A (instruction-based) data and projected (dotted line) through the Phase B (shaping) data.


Instructions
Shaping
Shaping + $$
4.9
4.8
4.7
4.6
4.5
4.4
4.3
4.2
(I)(S)
Ln(RCPM)
0
1
2
3
4
5
Ln(Session)

higher than previous scores, but rather, at 120.2 r/m, it was comparable to RCPM scores during latter shaping sessions. Over these remaining sessions of this brief reversal phase, RCPM increased slightly. However, RCPM never surpassed subject C02 personal best score, achieved on day 23, under shaping. When looking at the transformed RCPM data, subject C02 appeared, initially, to be performing at a level comparable to that found late in shaping. The transformed RCPM scores appeared to be on a declining trend during this phase. Considering the decline in performance found during the latter part of the shaping phase, it was unclear whether removal of shaping or some other factor was responsible for the decline. Again, because of the briefness of this training phase, the previous statement was offered as speculation only.

Return to shaping speeded responding. When shaping was reintroduced, RCPM increased (slightly) to its highest level to date, at 123.7 r/m. On the next shaping day, RCPM increased by almost 3 r/m, to 126.3 r/m, and then returned to 123.5 r/m by the end of this brief return-to-shaping phase. Looking at the transformed RCPM data, performance, here, was better than that during the previous phase, and that found during the latter few initial-shaping-phase sessions. Although performance appeared to be declining under return-to-Phase-A conditions, performance decreases were also found during the latter part of the initial shaping phase. As such, little could be concluded about the effects on performance of removing the shaping procedure. However, it was more difficult to discount the finding that reintroduction of shaping produced meaningful improvement over return-to-Phase-A performance, because the improvement was immediate and meaningful. This latter finding appeared to further support the notion that contingency-based procedures more effectively promote fluent responding relative to instruction-based procedures. Again, because of the brevity of the phase, conclusions had to be drawn with caution.

Phase C - Shaping-plus-monetary-reinforcement. When monetary reinforcement was introduced, RCPM increased to 125.9 r/m. Over the remaining three training days, RCPM continued to show slight though consistent improvement, finishing at 127.5 r/m on day 35. This continued improvement was also illustrated in the transformed RCPM data. The increase in RCPM was a result of a slight increase in accuracy (daily averages were between 96 and 97.5 r/m) in combination with decreasing latency.

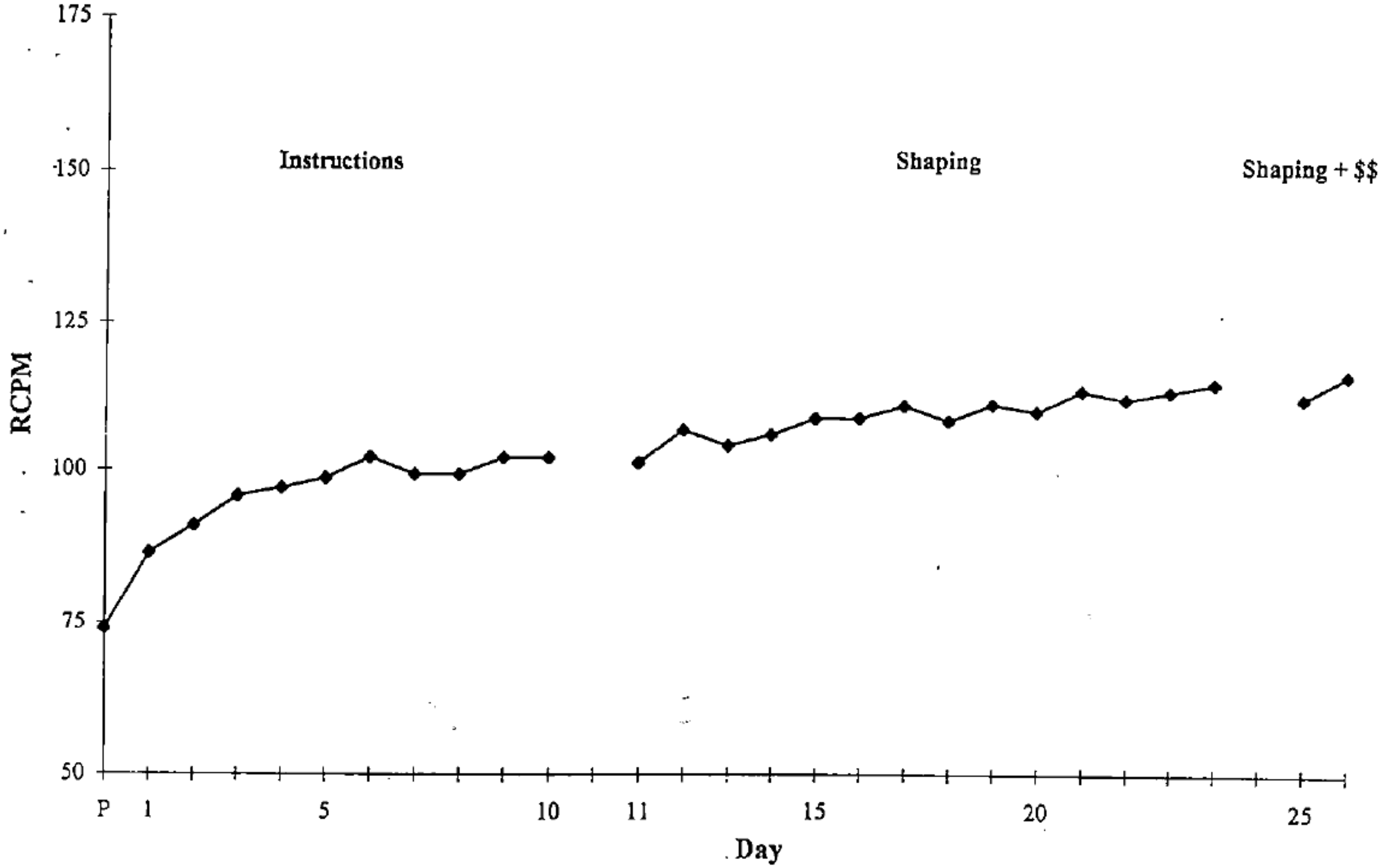
Overall summary of subject C02's training data. To summarize, as with all previous subjects, a large practice effect developed early, under instruction-based conditions. Similar to all but subject N01, progressively smaller improvements continued to be made throughout the remainder of the training phase. The shaping procedure had little initial effect on performance. However, about a third of the way into the shaping phase, performance was better than was predicted on the basis of projected instruction-based trends. On the latter two shaping days, performance fell off slightly, however, for the most part, subject C02's RCPM scores remained above the level predicted by the instruction-based regression line. It was uncertain whether the removal of shaping produced an early decrement in performance, because of the decrement in latter shaping performance. Clearly, though, the reintroduction of shaping produced an improvement in level of performance over that found during the second instruction-based training phase. When monetary reinforcement was added to the shaping procedure, both accuracy and latency improved, and subject C02 achieved her personal best RCPM score of 127.5 r/m.

Subject C03

Pretest. Similar to three of the four previous subjects, subject C03 scored 100% on two of the three trials, and 92.9% (one error) on the third. Therefore, it seemed appropriate to conclude that functional equivalence between training and generalization stimuli had been demonstrated.

Phase A - Instruction-based training. Subject C03's training RCPM data are located in Figure 9. During the initial, instruction-based training session (P), subject C03's RCPM was 74.1 r/m. Large improvements in performance occurred early in this phase, such that by the end of the first full day of training, subject C03's RCPM had increased by over 12 r/m, to 86.4 r/m. Over the next two training days, consistent though more modest performance gains were made, and by the end of the 4th day of training, RCPM was at 97 r/m. Subject C03's accuracy was fairly stable over first two thirds of instruction-based training (i.e., it was usually just above 90%), thus this improvement appeared to be a function of decreasing response latency. On day 6, subject C03's RCPM score reached a phase high of 102.1 r/m. Throughout the remainder of this initial training session,

Figure 9. Subject C03's daily rate-correct-per-minute (RCPM) data. Each data point (except the first of each phase) represents three 15-trial sessions.

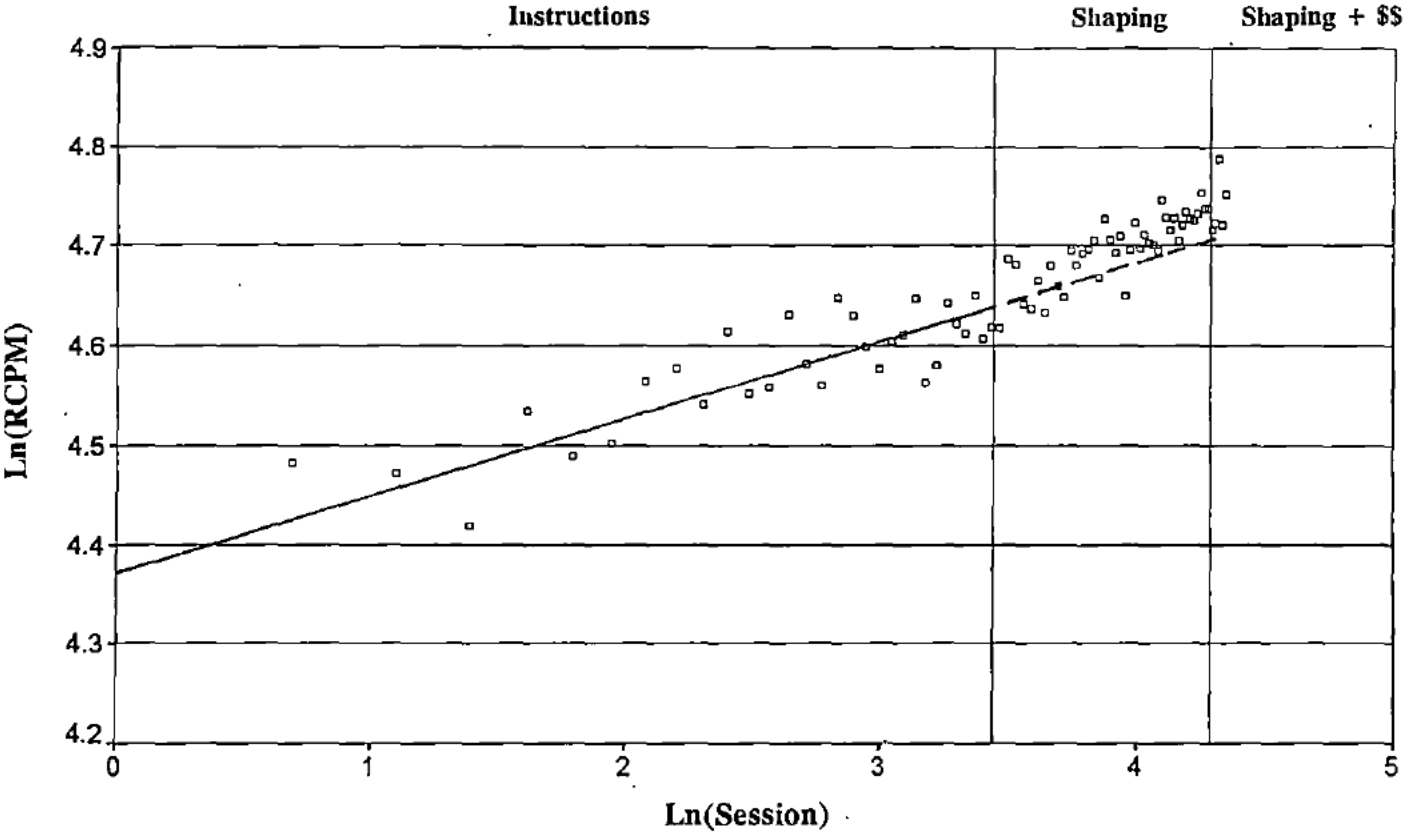

175
150
125
100
75
50
RCPM
Instructions
Shaping
Shaping + $$
P
1
5
10
11
15
20
25
Day

RCPM ranged between 99-102 r/m. Thus, like subject N01, subject C03 appeared to make no consistent additional performance gains during the latter third of instruction-based practice.

Phase B - Shaping speeded responding. When shaping was introduced, RCPM was 101.2 r/m, which was a slight decrease over the terminal Phase A value. Visual inspection of the raw RCPM data suggested continued, moderate improvement during the shaping phase. In fact, on the second shaping day, RCPM increased to 106.6 r/m. Thereafter, RCPM appeared to increase by approximately 2 r/m every few days, reaching a phase high of 114.6 on the final day of shaping. Improvement during this phase appeared to be a function of continued decreasing response latency in combination with an increase in accuracy (note the increase occurred on day 15), where accuracy was now consistently higher than earlier in shaping and throughout most of instruction-based training.

Subject C03's transformed RCPM data. Turning to the transformed RCPM data (Figure 10), the first thing to note is that the instruction-based data were only moderately well fit (i.e., the adjusted r^2 value was 77.72) by their regression line. When this regression line was projected through subject C03's shaping data, her performance was initially seemingly well described by the regression line. A few days into shaping, however, the subject was responding almost exclusively, and, most often, meaningfully above this predicted value. Note that this better-than-expected performance continued throughout the remainder of the shaping phase. Note, also, that this improvement occurred around the time that subject

Figure 10. Subject C03's per session, transformed (natural logarithmic) rate-correct-per-minute (RCPM) data plotted against the transformed (natural logarithmic) session number. Each phase is separated by a vertical line, and each data point represents fifteen 20-second trials of practice. A regression line was fit through the Phase A (instruction-based) data and projected (dotted line) through the Phase B (shaping) data.


Instructions
Shaping
Shaping + $$
4.9
4.8
4.7
4.6
4.5
4.4
4.3
4.2
Ln(RCPM)
0
1
2
3
4
5
Ln(Session)

C03's accuracy increased. It appeared, thus, that training under shaping resulted in better performance than was predicted on the basis of instruction-based trends.

Phase C - Shaping-plus-monetary-reinforcement. Subject C03 was unable to continue past the end of the 6th training week, making her exposure to shaping-plus-monetary-reinforcement limited. Initially, accuracy remained comparable to terminal Phase B values, while mean latency correct increased slightly. This suggested that the introduction of monetary reinforcement may have resulted in slightly slower responding. Subject C03's RCPM data appeared to support this notion as her initial, Phase C RCPM score of 112 r/m was lower than the terminal Phase B score. Similar to most of the previous subjects, subject C03's accuracy was higher on the second shaping-plus-monetary-reinforcement day. As a result, RCPM increased to a high (personal best) of 116 r/m. Thus, although not so initially, the addition of monetary reinforcement to the shaping procedure appeared to produce some improvement in performance. However, given the initial decrement in performance and the brief duration of this training phase, it was inappropriate to draw such a conclusion from the data. Therefore, the previous statement was offered for purposes of speculation only.

Overall summary of subject C03's training data. Similar to previous subjects, greatest performance gains occurred very early in training, followed by progressively smaller improvements as training continued. Performance was affected little when shaping was first introduced. However, a few days into the shaping phase, subject C03's performance was better than predicted under continued instruction-based practice. Although the data collected in the shaping-plus-monetary-reinforcement phase were limited, there was some suggestion that the introduction of monetary reinforcement resulted in slightly slower responding, followed by some improvement later in the session.

A summary of all subjects' training data. The following is a summary of findings across all subjects in this experiment. Contrary to the findings of Baron et al. (1983), relatively large practice effects developed after a very short period of instruction-based practice. Furthermore, three of five subjects demonstrated continued though progressively smaller improvement, even after extensive, instruction-based training. When shaping was introduced, only subject N01's performance was initially better than that predicted had she continued training under instruction-based conditions. The remaining subjects demonstrated better than predicted performance either fairly early on (subjects C01, C02, C03) or towards the end of the shaping phase (subject N02). Due to the short duration of the two reversal phases, it was considered inappropriate to draw definitive conclusions from the reversal data. However, the following observation could not be ignored. The present findings did not appear to support those of Baron et al. (1983), who reported that removal of the shaping procedure did not produce any decrement in performance. That is, there was some suggestion that removal of the shaping procedure, in the long term, might have resulted in an decrement in performance. Recall how, for subject N02, early, Return-to-Phase-A performance was strong, and, as the phase progressed, level of performance began decreasing.

Given further, instruction-based training, these data suggested that decrements in performance might have resulted. With subject C02, it was the improvement in performance that resulted when shaping was reintroduced, which suggested that contingency-based procedures most effectively promoted fluent responding on CRT tasks. When monetary reinforcement was introduced, accuracy and percent-correct-and-fast immediately increased for all but one subject (C01), while mean latency correct and RCPM scores appeared initially relatively unchanged.

Results and Discussion: Generalization Phase

Before proceeding with this final set of analyses, a number of points should be noted. First, unlike the training data, which were graphed on a daily basis, the generalization data were graphed on a trial-by-trial basis. This was done because of the short duration of the generalization phase. Second, performance on trial 1 of the generalization task is of potentially greatest interest as a pure measure of generalization, because it is only on this trial that subjects have no prior history of reinforcement for rapid responding to these stimuli. Admittedly, subjects had had prior exposure to the generalization stimuli during the pretest. However, at that time they were working under instructions which emphasized accuracy only, and had yet to experience speeded response training. Third, in the absence of complete generalization of fluent responding, subsequent data (i.e., trials 2-14) provide information about the rate-of-gain of speeded responding. These trials also provide information about stability of performance for training and generalization tasks alike.

Predictions can be made about the nature of the generalization data given either the absence or presence of generalization of fluent responding. First, in its absence, performance on the generalization task should be comparable to performance on the training task during early, instruction-based training. Data collected on day 1 of training have been included here to allow for such a comparison to be made. These data are referred to as "early-training performance" data throughout the remainder of this paper. Alternatively, given complete generalization of fluent responding, performance on the generalization task should be comparable to that on the training task during the generalization phase. Assuming the switch from training to the generalization phase has not disrupted performance, performance on the training and generalization tasks during the generalization phase should be comparable to performance on the training task on the last day of training. Data collected during the last training session were included for this comparison, and are referred to as "late-training performance" data throughout the rest of this paper. Finally, with an intermediate level of generalization, performance on the generalization task should be better than early-training performance, but poorer than late-training performance. The RCPM data was, again, chosen to be examined in detail. The reader can locate the graphs of each subject's accuracy and mean latency correct data, in the same order as the data are discussed in the next section, in Appendix B.

Numbers Task

Subject N01. Initial performance on the training task during the generalization phase (see Figure 11) was slightly better than late-training performance, suggesting that this subject was capable of making even further performance gains than those found during training. Subject N01's trial 1 RCPM on the generalization task was less than her trial 1 RCPM score on the training task, thus complete generalization of fluent responding was not displayed. However, subject N01's generalization task performance was clearly better than her initial training task performance. In fact, her generalization task performance was only moderately lower than her training task performance during the generalization phase. Overall, these data suggested that speeded response training resulted in a high degree of generalization of fluent responding to functionally equivalent stimuli.

Subject N02. First, subject N02's performance on the training task during the generalization phase (see Figure 12) was somewhat more variable than her late-training performance, suggesting that the move from training to the generalization phase resulted in a slight decrement in performance. Second, performance on the training task and on the generalization task during the generalization phase was initially only negligibly different. That is, RCPM scores on the training task were only negligibly higher than those on the generalization task. The following is a summary of performance on the remaining 12 trials of this generalization phase: on six of the trials, subject N01's training-task RCPM

Figure 11. Subject N01's generalization data. This graph contains her early and late training RCPM data, and her RCPM data on the training and generalization tasks during the generalization phase.

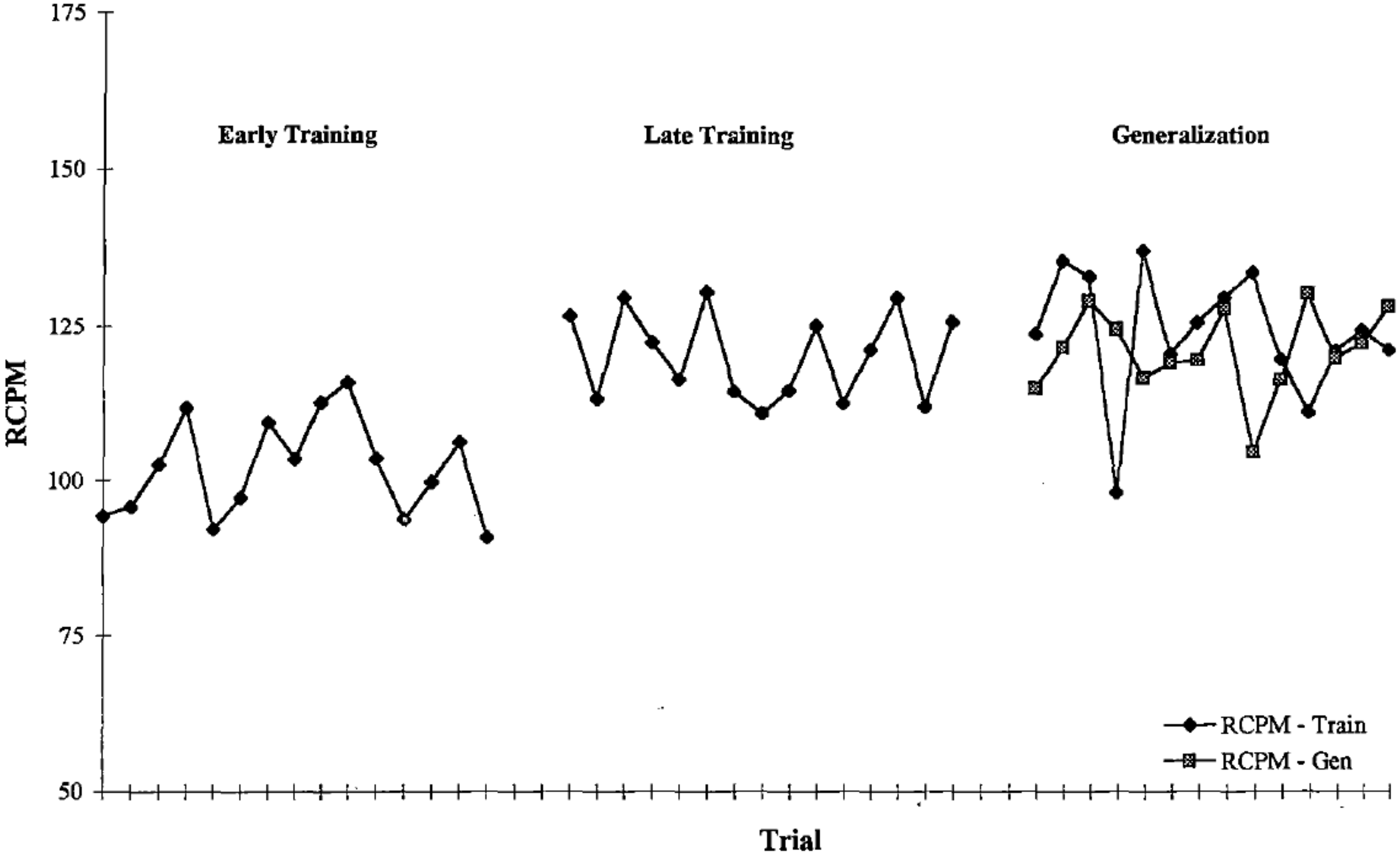
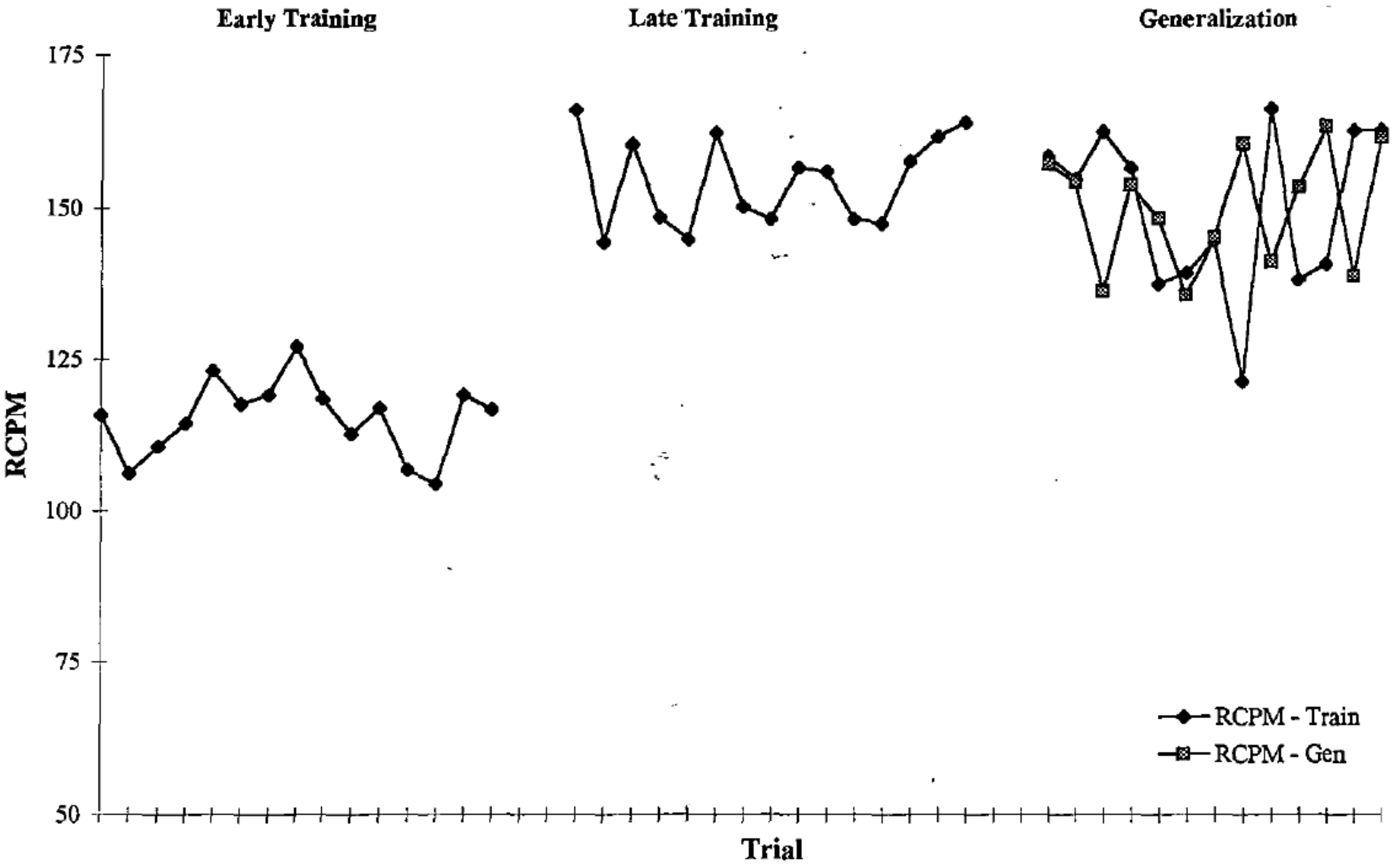

175
150
125
100
75
50
RCPM
Trial
Early Training
Late Training
Generalization
RCPM - Train
RCPM - Gen

Figure 12. Subject N02's generalization data. This graph contains her early and late training RCPM data, and her RCPM data on the training and generalization tasks during the generalization phase.

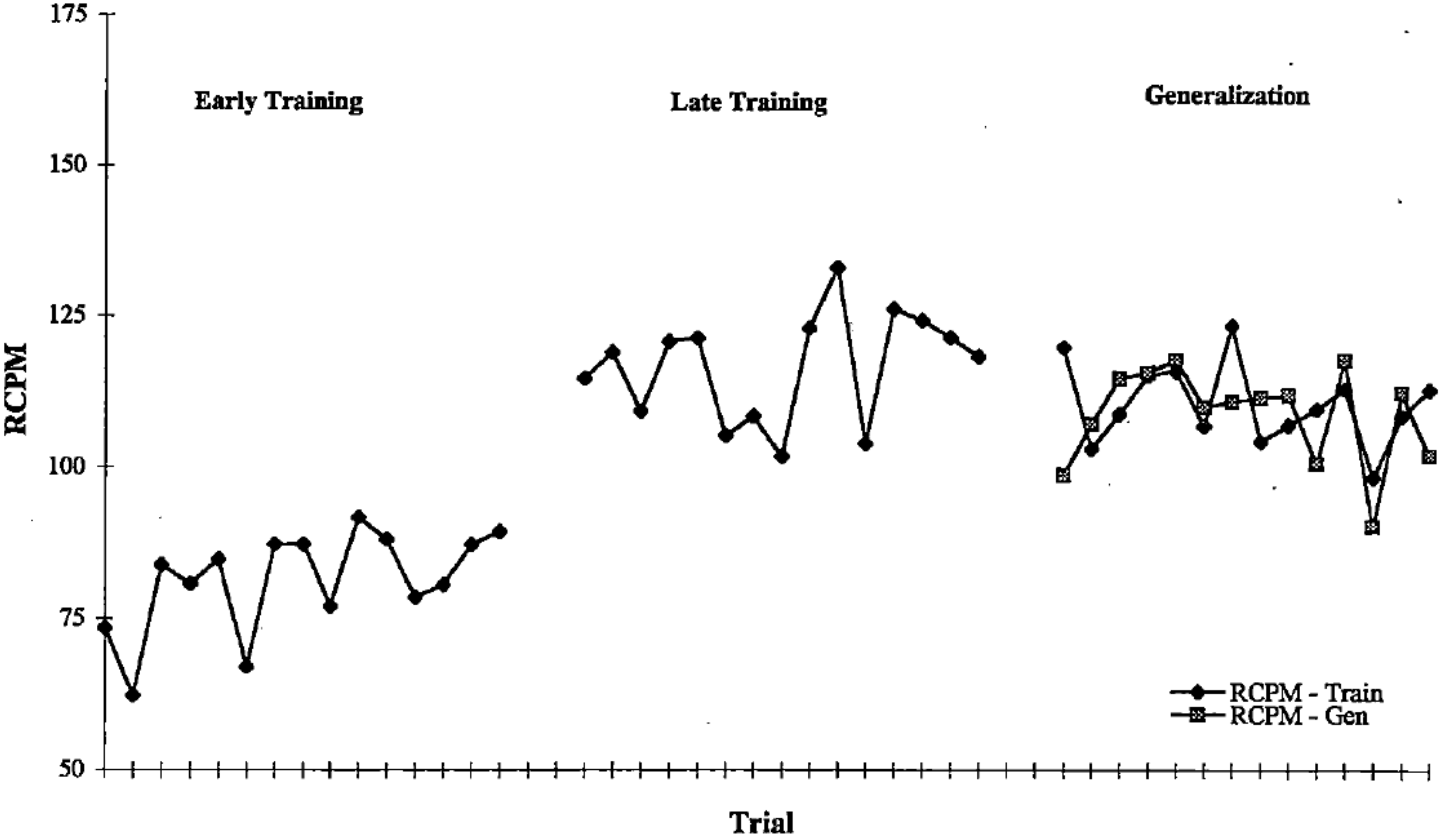

Early Training
Late Training
Generalization
175
150
125
100
75
50
RCPM
Trial
RCPM - Train
RCPM - Gen

scores were higher than her generalization-task RCPM scores; on four trials the opposite was the case; and on the remaining two trials, RCPM scores on both tasks were comparable. Overall, then, this subject demonstrated almost complete generalization of fluent responding to functionally equivalent stimuli upon completion of speeded response training.

Category Task

Subject C01. As illustrated in Figure 13, performance on the training task during the generalization phase was not as high as late-training performance. This suggested that switching from training to the generalization phase resulted in a slight decrement in performance. However, it should be noted also that performance during the generalization phase was generally less variable than late-training performance. The subject had a meaningfully higher initial RCPM score on the training task than she did on the generalization task, which suggested that subject C01 was initially more fluent on the training task than she was on the generalization task. However, on trials 2-6, subject C01 demonstrated greater fluency on the generalization task relative to the training task. These slightly higher RCPM score appeared to be a result of more slow and accurate responding to the generalization stimuli. Throughout the rest of generalization phase, it was difficult to assess to which task the subject was responding most fluently. On some trials her RCPM scores were higher on the training task, while on other trials, they were higher on the generalization task. Also, clearly, performance on the generalization task was superior to early training task

Figure 13. Subject C01's generalization data. This graph contains her early and late training RCPM data, and her RCPM data on the training and generalization tasks during the generalization phase.


Early Training
Late Training
Generalization
175
150
125
100
75
50
RCPM
Trial
RCPM - Train
RCPM - Gen

performance. Overall, these data suggested that speeded response training had resulted in much generalization of fluent responding to functionally equivalent stimuli.

Subject C02. Performance on the training task during the generalization phase (see Figure 14) was comparable to performance on the training task during latter training, thus there was no evidence in the RCPM data to suggest that switching from training to test affected performance on the training task. Like C01, this subject was initially more fluent on the training task (i.e., RCPM was 17.4 r/m higher) than on the generalization task. However, over the next three trials, unlike subject C01, subject C02 demonstrated greater fluency on the training task. Over the remainder of the generalization phase, subject C02 demonstrated, for the most part, greater fluent on the training task, however, the differences in RCPM across the two tasks was much reduced. Also, performance on the generalization task was far superior to early-training performance. Overall, then, much generalization of fluent responding to functionally equivalent stimuli resulted from speeded response training.

Subject C03. Subject C03's late-training performance (see Figure 15) and her performance on the training task during the generalization phase were both highly variable, making it difficult to establish whether moving from training to the generalization phase disrupted performance. Subject C03 was only slightly more fluent on the training task on trial 1 of the generalization phase than she was on the generalization task. On trial 2, performance on the training task was

Figure 14. Subject C02's generalization data. This graph contains her early and late training RCPM data, and her RCPM data on the training and generalization tasks during the generalization phase.

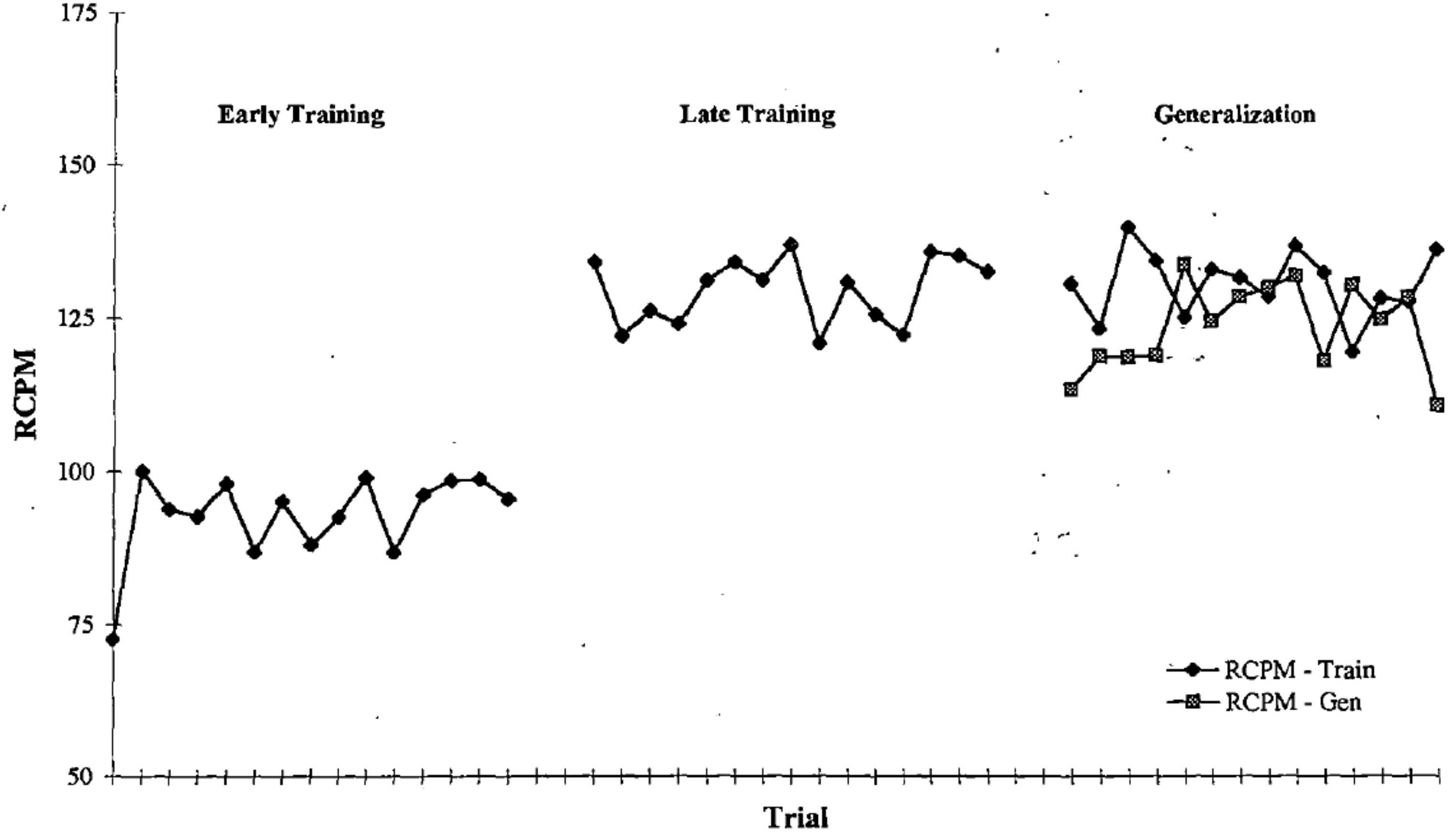
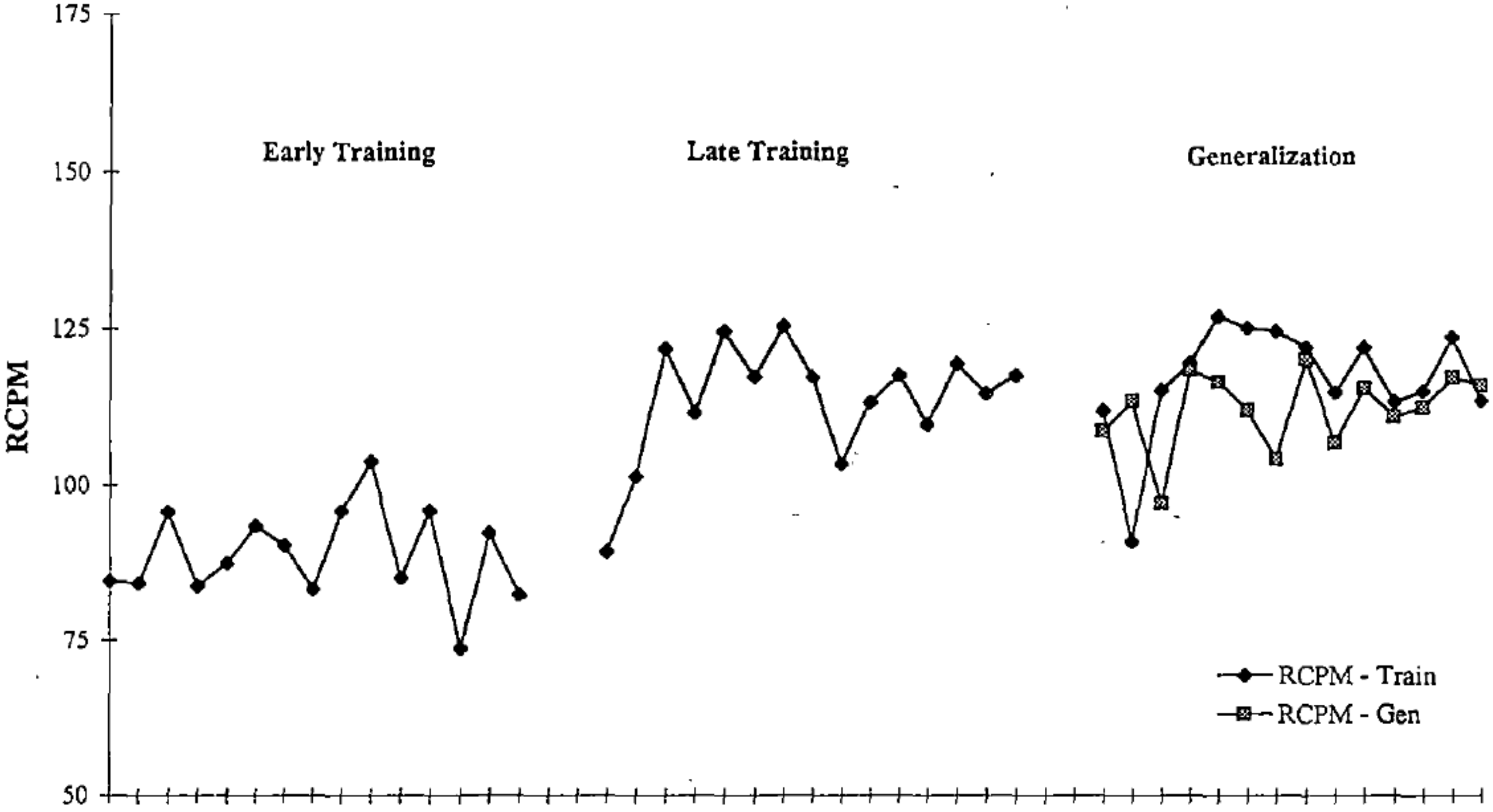

175
150
125
100
75
50
RCPM
Early Training
Late Training
Generalization
RCPM - Train
RCPM - Gen
Trial

Figure 15. Subject C03's generalization data. This graph contains her early and late training RCPM data, and her RCPM data on the training and generalization tasks during the generalization phase.


175
150
125
100
75
50
RCPM
Early Training
Late Training
Generalization
RCPM - Train
RCPM - Gen

atypically low, which resulted in greater fluency being demonstrated on the generalization task. Throughout the remainder of this generalization phase, subject C03's RCPM scores on the training task were almost consistently higher than those on the generalization task. When comparing subject C03's performance on the generalization task with her early-training performance, it was clear that although not complete, much generalization of fluent responding to functionally equivalent stimuli resulted from speeded response training.

A summary of the all subjects' generalization data. The following conclusions can be drawn from the generalization data collected in this experiment. First, switching from training to the generalization phase appeared to be only slightly disruptive to performance on the training task. Second, four of five subjects demonstrated superior performance on the training task on trial 1, whereas the remaining subject demonstrated comparable fluency on both tasks. Throughout the remainder of the generalization phase, subjects' RCPM scores on the training task were generally higher than those on the generalization task. However, when comparing performance on the generalization task with early-training performance for each subject, it was clear that much generalization of fluent responding to functionally equivalent stimuli resulted from speeded response training. In fact, often RCPM scores on the training task were only slightly higher than RCPM scores on the generalization task. Overall, then, these data provide at least some limited support for the claim made by precision teachers that training to fluency produces benefits beyond simple proficiency on the training task. Here, upon completion of speeded response training, subjects also demonstrated fluent performance on functionally equivalent, non-training stimuli.

General Discussion

In this experiment, instruction- and contingency-based procedures were studied as tools for promoting fluency on choice reaction time (CRT) tasks. The following conclusions can be drawn from the data. First, instruction-based practice on simple CRT tasks produces large performance gains in terms of decreasing response latency. Subjects demonstrated greatest improvement early in training, followed by continued, though more gradual improvement throughout the remainder of the instruction-based phase. In fact, three of five subjects demonstrated continued improvement over the entire training phase. Second, contingency-based procedures are superior to instruction-based procedures in facilitating the development of fluent responding to simple CRT tasks. When switched from instruction- to a contingency-based procedures, all subjects demonstrated greater improvement than was theoretically predicted to occur on the basis of projection of regression lines fitted to their log-log-transformed RCPM data. Third, the addition, late in training, of monetary reinforcement to the shaping procedure produces even further improvement. Here, four subjects demonstrated more accurate responding, while mean latency correct (and early RCPM scores) remained relatively unchanged. In addition, monetary reinforcement appears to increase subject motivation. Fourth, switching from training to the generalization phase produces only minor decrements in performance. Finally, speeded response training appears to promote considerable generalization of fluent responding to functionally equivalent non-training stimuli. The specific contribution the present experiment made to the various literatures and areas of research will now be discussed.

Traditional Instruction-based CRT Training. The present experiment made numerous contributions, several of which were methodological, to the CRT literature. First, the instruction-based data confirms the assertions made by CRT researchers that large performance gains develop over relatively short periods of training under simple, instruction-based conditions. Second, this experiment demonstrates that meaningful performance gains can continue to be made even after prolonged training. Long-term data collection is rarely found in the CRT literature, making most of our understanding of the development of practice effects limited to conclusions drawn from short-term data collection. From this standpoint, the present experiment made a unique and valuable contribution to the literature. Third, the typical CRT investigator rarely provides detailed data concerning accuracy and within-session variability; CRT researchers generally report that accuracy remained above some criterion value, but fail to discuss the extent to which accuracy varied within the specified range across training. Here, these measures are made available to the reader, to allow him or her to make an independent assessment of the magnitude of the practice effect found during any given period of training. Fourth, in the present experiment, the difficulty of assessing improvement across multiple performance measures (viz., accuracy _and_ latency) is addressed, and two additional performance measures (one which looked at latency while holding accuracy constant, the other [RCPM] which was sensitive to both accuracy and latency) are included to provide a more rigorous and interpretable assessment of performance.

Admittedly, the present instruction-based procedure differs from traditional instruction-based methodology in that subjects were continuously encouraged to be reducing response latency, and were given more extensive performance feedback (i.e., various response-by-response, end-of-trial and end-of session performance measures were presented to subjects) which informed them of their progress. Such is typically not found in the CRT research. This may, in part, explain why subjects appear to reach asymptote so early in the typical CRT study. That is, initial improvement occurs, and this improvement may be, to a large degree, a function of the training-appropriate discriminative stimuli acquiring general control over responding. However, as training continues, rate of improvement begins to progressively decline, and soon training is terminated.

Baron, Menich and Perone (1983). Although considered a replication of Baron et al. (1983), the present research stands apart from the former in many important ways. First, a multiple baseline (across subjects) design was adopted in the present experiment. Again, this was done to guard against the possibility that some unknown variables, correlated with time, rather than the manipulation of the independent variable, were responsible for changes in level of performance across training. Second, a latency adjustment procedure which conformed more closely to the definition of shaping was adopted in the present experiment. Here, the stringency of the latency criterion could be either increased or decreased depending upon the progress of the subject. As discussed earlier (recall subject N01's shaping data), this appeared to be an extremely important procedural modification. Most importantly, the confound suspected to be operating in the Baron et al. (1983) study was controlled for in the present study. Recall that to prevent reinforcement and instructional effects from being confounded, subjects received both accuracy- and latency-dependent end-of-trial feedback, were continually encouraged to be trying to reduce their median latency correct scores, and were not told in advance of the various phases of the experiment. In sum, their job throughout training was defined to be continually trying to drive their median latency correct scores down, while keeping accuracy high. A second confound, between treatment and duration of practice, was also present in the Baron et al. (1983) study, as it was in the present experiment. That is, as training progressed, the magnitude of the practice effect necessarily decreased. Baron et al. (1983) chose to transform latency scores to speeds, so they appeared to be aware of the confound. However, as discussed earlier, transforming the data by taking the reciprocal of latency did not adequately control for the confound. Despite the confound, Baron et al. (1983) reported that the introduction of shaping resulted in the development of a meaningfully larger practice effect. What is important to understand, here, is that had Baron et al. (1983) more effectively controlled for this confound, they might have collected even stronger evidence to support the notion that contingency-based procedures more effectively promote speeded responding than do instruction-based procedures. In other words, the effects of shaping were likely mitigated by the interaction between changes in treatment and duration of practice. In the present experiment, this confound was controlled for by taking natural logarithmic transformations of each subject's per-session RCPM data; plotting these scores against the natural logarithmic transformations of the respective session number (the log-log transformation making the data linear); fitting a regression line through each subject's instruction-based data; projecting this regression line through the contingency-based data; and then assessing how well the instruction-based regression line fit the contingency-based data. Prevention of the first confound, and control of the second, allowed the present experimenter to make a better comparison of traditional, instruction-based methodology, where both direct feedback for speed was not included, with contingency-based procedures, where direct feedback for speed was included, as a tool for promoting fluency on simple CRT tasks.

With the changes in procedure, came somewhat different results. Recall that in Baron et al's (1983) study, subjects demonstrated some improvement during baseline. However, this level of improvement was small relative to that which resulted when subjects were switched into the shaping procedure. In the present experiment, in contrast to Baron et al. (1983), greatest improvement occurred during the instruction-based training phase. However, similar to Baron et al. (1983), when shaping was introduced, all subjects made greater performance gains than would have been predicted under continued instruction-based training, even after long-term practice. These data suggest, then, that contingency-based procedures are superior to instruction-based procedures in promoting the development of fluent responding, at least with these types of CRT-tasks, because, again, even after several weeks of training, when shaping was introduced, better-than-predicted performance resulted. Finally, Baron et al. (1983) reported that removal of the shaping procedure did not result in a decrement in performance, and when shaping was reintroduced, no further improvement was found. In the present experiment, two subjects experienced reversal conditions. Although caution must be used when examining these data, they appear to question Baron et al's (1983) findings, in that there is some suggestion, in both reversal subjects' data, that removal of the shaping procedure would have resulted in performance decrements over the long term. It is interesting to note that Baron et al's (1983) subjects trained for a much shorter duration under reversal conditions. Perhaps Baron et al. (1983) would have found performance decrements had their subjects trained longer under return-to-baseline conditions. After all, the present reversal subjects did not show immediate decrement in performance. Rather, performance decrements became evident a couple of sessions into the reversal phase. Furthermore, although improvement was small during the latter part of the initial shaping phase, there was no evidence to suggest that subjects had reached their fastest responding, even after a total of eight weeks of training. In sum, despite smaller and more subtle benefits of shaping and monetary reinforcement than evident in Baron et al. (1983), the present findings more than suggest that for optimal CRT training, both shaping and monetary reinforcement are worthwhile.

A Closer Consideration of the Shaping Result and Procedure. Though there was evidence that contingency-based procedures more effectively promoted fluent responding in the present data, certain other findings must not be ignored. Most notably, initial performance under shaping for four of five subjects was not meaningfully different from that predicted from the instruction-based data. That is, although better-than-expected improvement was found during the shaping phase, immediate improvement was found in only one subject's data. As greater-than-predicted improvement was eventually found with all subjects, this lack of early improvement might have been, in part, a reflection of some inadequacy in the shaping procedure chosen by the experimenter. Was the 90th percentile of the distribution of latencies for correct responses an appropriate starting latency, given that reaction time data is positively skewed? After all, the more skewed the initial latency distributions were, which would necessarily drive up the latency at the 90th percentile, the more liberal the initial latency criterion would have been. In addition, the difference between the 40th and 50th percentiles was often very small (i.e., between 8-15 ms). Thus, given an overly liberal latency criterion and arguably small adjustments to the stringency of the latency criterion when subjects' % C&F scores surpassed 85, it may have taken many sessions before an effective (i.e., one where the upper limit of acceptable response latencies is not overly liberal) latency criterion was, in fact, in place. On the other hand, shaping by its very nature is a slow process, so expecting to see immediate change was, perhaps, unreasonable.

Mainstream educators, precision teachers and micromolar behavioral researchers. As discussed earlier, mainstream educators, precision teachers and micromolar behavior theorists disagree on the emphasis that should be placed on accuracy and response speed during training. The present experiment, in support of precision teachers, confirms experimentally that at least with these types of CRT tasks, emphasizing accuracy and appropriate response speed from the onset of training is an appropriate training strategy.

Admittedly, the training task chosen for this experiment was somewhat artificial, at least relative to, for example, a study of the development of fluency on writing answers to simple math operations or speed reading. As such, some might argue that the results of the present research have little real-world significance. However, it is often necessary to begin by eliminating as much external variation as possible to establish how variables covary at the most fundamental level. Subsequent research, which includes some or all of the previously excluded variables, can then be done to identify and describe the more complex relations between multiple variables, like those typically found in the real world. By proceeding in this fashion, a better understanding of the relationship among variables will result. Interestingly, fluency research appears to have proceeded in the opposite direction. That is, most of the early research was conducted by precision teachers and other behaviorally oriented educational researchers in the classroom, while more recently, some of the research has moved into the laboratory.

In response to the assertions made by micromolar behavior researchers that some more complex tasks require initial accuracy-intensive practice, a precision teacher would likely recommend dividing the complex behavior into component parts, and programming fluency on each part before having the student attempt the complex task. Often, when a complex task is broken into component parts and those parts are then each trained to fluency, the complex behavior emerges with no additional training. The term contingency adduction has been used to describe this phenomenon (Andronis, in Binder, 1995; Johnson & Layng, 1992; K. Johnson, personal communication, September 28, 1995). Thus, where complex tasks can be reduced into component parts, the concern expressed by micromolar behavioral theorists, that initial accuracy-intensive practice is essential, appears to have been answered by precision teachers.

Functional equivalence-based generalization. The generalization test was, admittedly, limited in nature, but none-the-less, the resultant data make a valuable contribution to the equivalence literature. As discussed earlier, behavior analytic theory has often been criticized for not being able to adequately explain certain characteristics (e.g., the emergence of novel utterances) of language development. However, the present findings, that fluent responding generalized to non-training stimuli which were functionally equivalent to the training stimuli, may offer some insight into this issue. Perhaps one of the factors which determines if novel utterances emerge is whether or not the stimulus portion of the trained (S-R) relation is a member of some predefined, functionally equivalent set. If such is the case, and an S-R relation is trained between one member of the set and some response, it seem reasonable to predict, appealing to stimulus generalization, that a response trained to one member of a functionally equivalent set of stimuli will also be emitted in the presence of an untrained member of the set. This prediction, of course, requires direct and thorough examination. Similar proposals have been made by equivalence researchers, who are becoming increasingly more able to describe the environmental factors which affect language acquisition (Hall & Chase, 1991). In any case, the present data suggest that response latency is a potentially informative variable, and should, thus, continue to be studied in the equivalence paradigm. Moreover, finding evidence of generalization of fluent responding to functionally equivalent stimuli may offer some empirical support for the assertion made by precision teachers that training to fluency produces benefits in terms of application to transfer of training.

The implications of accommodating special populations with more time for responding. Behavioral fluency is important for many reasons. (Recall the benefits [greater retention; longer endurance; application] said to accompany achieving fluency on a given task.) As such it is worthwhile to speculate on the implications of the practice of accommodating certain populations (i.e., learning disabled students; persons with physical disabilities; elderly citizens) with more time for responding, rather than helping them to achieve the greatest possible level of fluency on important tasks.

As discussed earlier, teachers of learning disabled students have traditionally favoured training their students on non-challenging tasks (i.e., where the learner can easily respond with close to perfect accuracy) (Mercer & Mercer, 1993; Scott et al., 1990). Thus little or no emphasis is typically placed on training appropriate response speed. In fact, rather than receiving appropriate response speed training, learning disabled students are often simply given more time to complete a given task (J. Parsons, personal communication, January 10, 1996). Recall, however, that Scott et al. (1990) reported that learning disabled students made larger educational gains when emphasis was shifted away from maintaining high accuracy on relatively easy tasks, to training on more challenging tasks, where rate-pacing was also incorporated into the procedure. Recall, also, that Johnson and Layng (1992) demonstrated that students diagnosed with attention-deficit disorder (ADD), "when given extensive endurance training on a variety of tasks were able to greatly increase their attentions span" (starting from 20-second timings and building towards the target duration) (p. 1481). Furthermore, Binder et al. (1990) suggested that learners "typically lack the ability to maintain steady performance levels for extended periods of time [prior to] attaining certain minimum levels of speed and accuracy on individual curriculum tasks" (p. 25).

Finally, fluency on basic skills (e.g., basis arithmetic) is often a prerequisites for successful completion of a more complex task (e.g., higher-level math problems) (Binder, 1993; Gagné, 1983; Haughton, in Johnson & Layng, 1991). It follows that teachers of learning disabled students, when accommodating their students by giving them more time to complete basic tasks (i.e., tool skills), are not only preventing their students from enjoying the additional benefits that accompany training to fluency, but they may also be inadvertently imposing a ceiling on learning more complex tasks.

Training to fluency also offers benefits to persons with physical disabilities. For example, Sue Imbriglio (1992), a physical therapist, incorporates training to fluency of some motor skills in her treatment of patients with Huntington's Disease. Huntington's disease (HD) is a "progressive and degenerative genetic condition of the central nervous system," producing cognitive deficit, motor impairment, and eventually death (Imbriglio, 1992, p. 62). To assist persons with HD in maintaining maximum independence for as long as possible, "both ability and fluency of movement [are measured] with the goal of increasing the rate at which a patient is able to perform a task" (Imbriglio, 1992, p. 66). Thus where training (or retraining) fluent performance is possible and beneficial to the patient, providing that patient with more time to perform a given task, may, in the end, lead to the patient's premature loss of independence.

Consider, next, the case of someone who suffers a mild brain injury such that he or she has not suffered complete behavioral deficits, but, rather, can no longer perform certain tasks quickly and efficiently. Here, retraining fluent performance, rather than accommodating that person with more time to complete those tasks, would maximally benefit that person, because it would allow her or him to regain the maximum level of proficiency on the given tasks, and also enjoy the additional benefits produced by training to fluency.

Lastly, as discussed earlier, Welford reported that as one ages, his or her reaction time becomes progressively slower. However, Baron et al. (1983) found that although older (i.e., senior citizens) subjects responded initially more slowly than their younger (undergraduates) counterparts, both younger and older subjects made comparably large speed gains as a result of training. That is, the initial difference between the older and younger subjects was maintained, however both groups of subjects made large speed improvements (Baron et al., p. 282). This suggests that <u>some</u> of the slowing that accompanies aging may be the result of reduced opportunities to practice rapid responding, rather than irreversible changes in the central nervous system. Thus where older citizens are capable of responding more fluently, but are, instead, given more time to respond, the slowing that accompanies changes in the central nervous system may inadvertently be magnified.

Suggestions for future research

Laboratory corroboration of findings reported by precision teachers and others is essential for the progression of the science, therefore the following suggestions for future research are offered. First, it seems worthwhile to investigate further whether the removal of the shaping procedure results in a decrement in performance as training continues. Recall that Baron et al. (1983) reported that performance was not affected by the removal of the latency criterion. In contrast, in the present study, performance decrements were apparent either immediately upon removal of the shaping procedure (subject N02) or upon reintroduction of shaping (subject C02). If the findings in the present study are validated, than further support for the superiority of contingency-based procedures in both programming and maintaining fluent responding will have been collected.

Although shorter response latencies were progressively shaped in the present study, the procedure was complicated somewhat by the fact that accuracy, although it was not being shaped, had to be maintained at a certain level. An alternative approach that could have been taken, and one that should be considered by future CRT researchers, is to shape increasing RCPM scores, a statistic which approximates that used in precision teaching. With both accuracy and latency being reflected in the computation of the statistic, study of this performance measure appears to offer an advantage over the present methodology. It should be recognized, however, that in choosing to study RCPM, an experimenter may, in some cases, be trading-off some precision-of-measurement and feedback. Recall that an increase in RCPM does not necessarily reflect decreasing response latencies. In fact, an increase in RCPM could be the result of any one of the following occurrences: increased accuracy with no change in latency; increased accuracy with an accompanying decrease in latency; no change in accuracy with a decrease in latency; decreased accuracy with a proportionally greater decrease in latency. If latency is the variable of interest, then, clearly, studying the RCPM data alone, rather than the latency data, or some combination of the two, does not appear to be the most appropriate procedure. However, if both accuracy and speed are being investigated, depending upon the emphasis placed on each, the opposite might be true. On the other hand, in terms of the precision of feedback, since RCPM feedback can necessarily be provided only on a cumulative basis (i.e., end-of-trial is the first time this statistic can be computed), its effectiveness as a training procedure might be less than one based on response-by-response feedback of latency. Alternatively, if the goal is to train rapid responding over time, (i.e., across a series of tasks), then a cumulative measure is, arguably, perfectly appropriate. Ultimately, of course, this is an empirical question. Further, as is often the case, there may not be one (unequivocally) ideal approach. In the end, the experimenter must clearly identify his or her variable of interest and choose what appears to be the most appropriate procedure, while, at the same time, acknowledging the potential limitations which accompany that choice of procedure.

In retrospect, the addition of monetary reinforcement, contingent upon % C&F, may have failed to promote faster responding, due to the fact, overlooked by the experimenter, that increasing % C&F scores do not necessarily reflect decreasing response latencies. That is, as % C&F increases, more responses are meeting the latency criterion. However, the latencies of those extra few responses might be only a few milliseconds shorter than the criterion. Note, also, that given the relatively brief duration of a training trial, a seemingly large increase in % C&F (e.g., with some subjects a 5% C&F increase was equivalent to one additional correct-and-fast response) may have resulted from only one or two additional correct responses meeting the latency criterion. This might explain why RCPM scores did not appear to increase when monetary reinforcement was introduced. Also, where additional responses were now correct and just meeting the latency criterion, mean latency correct could easily have increased slightly, suggesting a decrement in performance. It would be interesting, then, to compare these results with those from a study where monetary reinforcement was made contingent upon decreasing median latency correct scores, provided accuracy remained above some criterion level.

It would be worthwhile, also, to investigate the extent to which speeded responding generalizes to other non-training tasks. The generalization data collected in the present experiment suggested that, at least with CRT-type tasks, much generalization of fluent responding to functionally equivalent stimuli results from training. Research of this nature must continue because, after all, the more generalization that can be programmed into a training procedure, the more effective and valuable the training technique. Further research on the benefits said to accompany training to fluency would also greatly benefit the precision teaching literature. Looking back, it would have also been interesting in the present study to have run follow-up sessions to assess level of performance ("retention") after long periods of no practice. After all, "true mastery" of a task requires, among other things, that the learner be able to demonstrate fluent performance after a significant period of non-practice (Johnson & Layng, 1992).

Perhaps the most important recommendation that can be made is to urge other educational researchers and practitioners in the behavior analytic community to continue studying variables suspected to affect the development of fluency, and to continue promoting those procedures which have been shown to facilitate fluent performance on both academic and non-academic tasks. Furthermore, greater emphasis must be placed upon publishing the data from such studies (in both academic and non-academic journals), if precision teachers and other behavioral educational researchers are to be successful in their attempt to convince mainstream educators and trainers that training to fluency is necessary and important. Although much work lies ahead, these and other important challenges that have yet to be faced, like "error-filled" learning, should be viewed as opportunities for further learning and for further advancement of the science of human behavior.

References

Ackerman, P.L. (1988). Determinants of individual differences during skill acquisition: cognitive abilities and information processing. Journal of Experimental Psychology, 117(3), 288-318.

Baron, A., Menich, S.R., & Perone, M. (1983). Reaction times of younger and older men and temporal contingencies of reinforcement. Journal of the Experimental Analysis of Behavior, 40(3), 275-287.

Battig, W.F., & Montigue, W.E. (1969). Category norms for verbal items in 56 categories: a replication and extension of the Connecticut category norms. Journal of Experimental Psychology Monograph, 80(3, part 2), 1-45.

Binder, C. (1987). Fluency building. Precision Teaching and Management Systems. Inc., Nonantum, Ma., 1-12.

Binder, C. (1988). Precision Teaching: measuring and attaining exemplary academic achievement. Youth Policy, 10(7), 12-15.

Binder, C. (1993, October). Behavioral fluency: a new paradigm. Educational Technology, 8-14.

Binder, C. (1995). Behavioral fluency: evolution of a new paradigm. Manuscript submitted for publication.

Binder, C., Haughton, E., & Van Eyk, D. (1990, Spring). Increasing endurance by building fluency: Precision teaching attention span. Teaching Exceptional Children, 24-27.

Catania, A.C. (1984). Learning (3rd ed.). Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Catania, A.C., Horne, P., & Lowe, C.F., (1989). Transfer of function across members of an equivalence class. Analysis of Verbal Behavior, 7, 99-110.

Ellis, H.C., & Hunt, R.R. (1993). Fundamentals of cognitive psychology (5th ed.). Indianapolis, Indiana: WCB Brown & Benchmark.

Gaddes, W.H., Edgell, D. (1994). Learning disabilities and brain function: A neuropsychological approach (3rd ed.). New York: Springer-Verlag New York Inc.

Gagné, R.M. (1982). Some issues in the psychology of mathematical instruction. Journal of Research in Mathematics Education, 14(1), 7-18.

Hale, D.J. (1969). Speed-error tradeoff in a three-choice serial reaction task. Journal of Experimental Psychology, 81(3), 428-435.

Hall, G.A., & Chase, P.N. (1991). The relationship between stimulus equivalence and verbal behavior. The Analysis of Verbal Behavior, 9, 107-119.

Hicks, L.H., Birren, J.E., (1970). Aging, brain damage and psychomotor slowing. Psychological Bulletin, 74, 377-396.

Imbriglio, S. (1992). Huntington's disease at mid-stage. Clinical Management, 12(5), 62-72.

Johnson, K.R. & Layng, T.V.J. (1991). Breaking the structuralist barrier: Literacy and numeracy with fluency. American Psychologist, 47(11), 1475-1490.

Judd, W.A., & Glaser, R. (1969). Response latency as a function of training method, information level, acquisition, and overlearning. Journal of Educational Psychology, 60(4), 1-30.

Keel, S.W. (1973). Attention and Human Performance. Pacific Palisades, California: Goodyear Publishing Company, Inc.

Koorland, M.A., Keel, M.C., & Ueberhorst, P. (1990). Setting aims for precision learning. Teaching Exceptional Children; 22(3), 64-66.

Kristofferson, M.H. (1977). The effects of practice with one positive set in a memory scanning task can be completely transferred to a different positive set. Memory & Cognition, 5(2), 177-186.

Kučera, H. & Francis, W. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.

Leonard, J.A., & Newman, R.C. (1965). On the acquisition and maintenance of high speed and high accuracy in a keyboard task. Ergonomics, 8, 281-304.

Lezak, M.D. (1983). Neuropsychological assessments (2nd. ed.). New York: Oxford University Press.

Lindsley, O.R., (1990a). Our aims, discoveries, failures and problems. Journal of Precision Teaching, 7, 7-17.

Lindsley, O.R. (1990b, Spring). Precision Teaching: By teachers for children. Teaching Exceptional Children, 10-15.

Logan, F.A. (1956). A micromolar approach to behavior theory. Psychological Review, 63(1). 63-73.

Logan, F.A. (1963). Micromolar behavior theory and performance speed in education. Harvard Educational Review, 33(2), 178-185.

Martin, G., & Pear, J. (1988). Behavior modification: what it is and how to do it (3rd ed.). New Jersey: Prentice Hall.

Mercer, C.D., & Mercer, A.R. (1989). Teaching students with learning problems (4th ed.). Don Mills, Ontario: Maxwell Macmillan Canada Inc.

Newell, A., & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson's (Ed.) Cognitive skills and their acquisition (pp 1-55). Hillsdale, New Jersey: Lawrence Earlbaum Associates.

Pashler, H., & Baylis, G. (1991). Procedural learning: 1. Locus of practice effects in speeded choice tasks. Journal of Experimental Psychology: Learning, Memory,and Cognition, 17,(1), 20-32.

Potts, L., Eshleman, J.W., & Cooper, J.O. (1993). Ogden R. Lindsley and the historical development of precision teaching. The Behavior Analyst, 16,(2), 177-189.

Rabbitt, P., & Banerji, N. (1989). How does very prolonged practice improve decision speed? Journal of Experimental Psychology: General, 118,(4), 338-345.

Rogers J.G. (1968). Input and output speed components of learning to learn. Journal of Experimental Psychology, 78(2), 233-241.

Schneider, W., & Fisk, A.D. (1984). Automatic category search and its transfer. Journal of Experimental Psychology: Learning, Memory, and Cognition 10,(1), 1-15.

Scott, J., Stoutimore, J., Wolking, B., Harris, C. (Spring, 1990). Challenging reading for students with mild handicaps. Teaching Exceptional Children, 32-35.

Shiffrin, R.W., & Logan, F.A. (1965). Performance speed as a function of practice speed. Journal of Verbal Learning and Verbal Behavior, 4(4), 335-338.

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M.D. Zeiler (Eds.), Analysis and integration of behavioral units (pp. 231-245). Hillsdale, NJ: Erlbaum.

Spence, K.W. (1956). Behavior theory and conditioning. New Haven: Yale University Press.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction time experiments. American Scientist, 57, 421-457.

Stevens, J.C., & Savin H.B. (1962). On the form of learning curves. Journal of the Experimental Analysis of Behavior, 5(1), 15-18.

Tabachnick, B.G., Fidell, L.S. (1989). Using multivariate statistics. (2nd ed.). New York: Harper Collins Publishers, Inc.

Vaughan, W. (1988). Formation of equivalence sets in pigeons. Journal of Experimental Psychology: Animal Behavior Processes, 14(1), 36-42.

Vernon, P.A. (1990). An overview of chronometric measures of intelligence. School Psychology Review, 19(4), 399-410.

Welford, A.T. (1980). Relationships between reaction time and fatigue, stress, age and sex. In A.T. Welford (Ed.), Reaction times (pp. 321-354). Toronto: Academic Press.

Woodworth, R.S., & Schlosberg, H. (1965). Experimental psychology (rev. ed.). Toronto: Holt, Rinehart, & Winston.

Appendix A

The following series of graphs (two per subject) illustrate the daily accuracy (%C), mean latency correct (MLC) and standard deviation of latency correct (SDLC) data, and the daily "latency on 100% accuracy trials data" for each subject.

<u>Figure 16.</u> Subject N01's accuracy (%C), mean latency correct (MLC), and standard deviation of latency correct (SDLC) training data. Note, the scale in parentheses on the primary y-axis describes the standard deviation data.

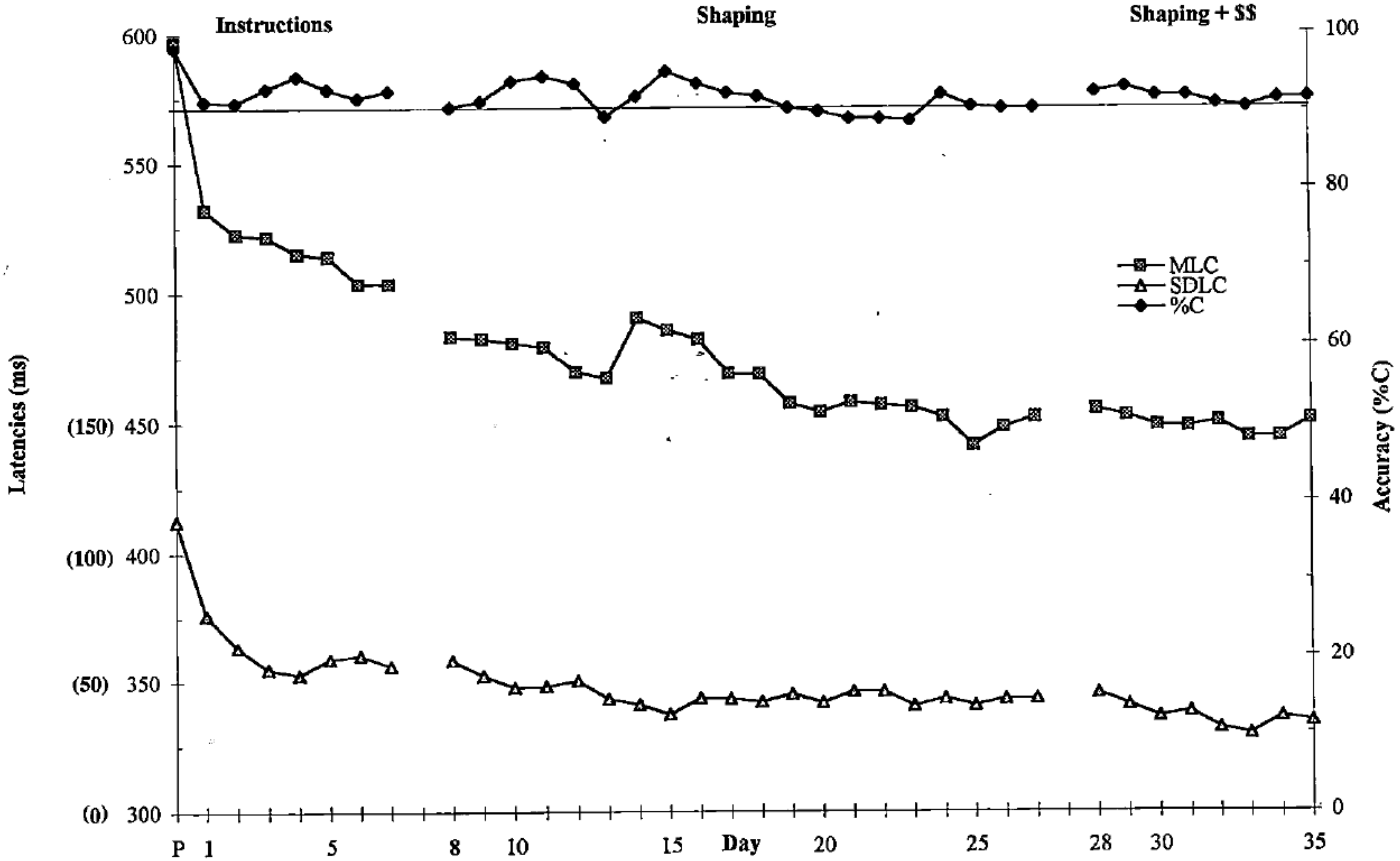
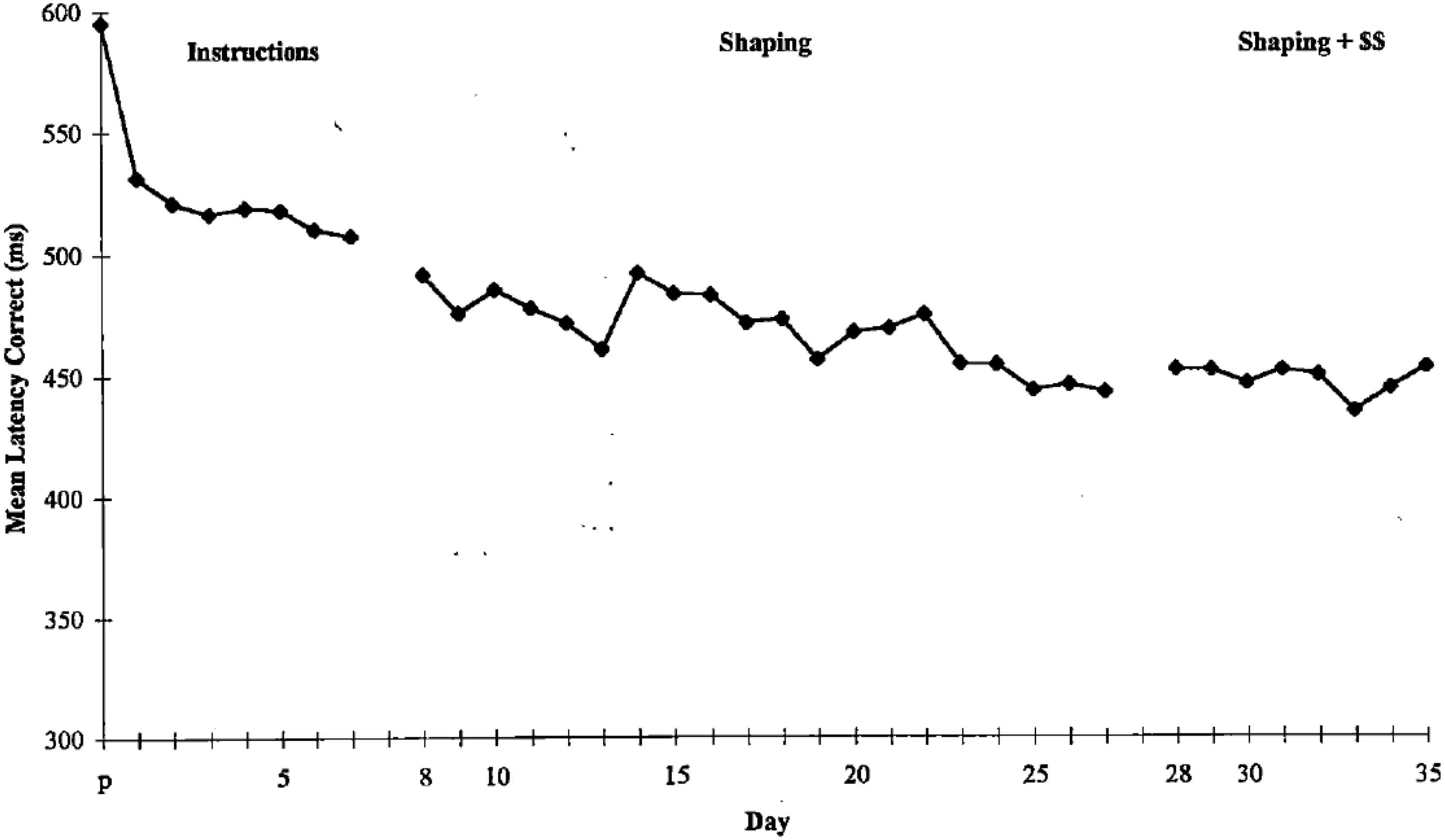

Instructions
Shaping
Shaping + $$
Latencies (ms)
600
550
500
(150) 450
(100) 400
(50) 350
(0) 300
Accuracy (%C)
100
80
60
40
20
0
MLC
SDLC
%C
P
1
5
8
10
15
Day
20
25
28
30
35

Figure 17. Subject N01's latency on 100% accuracy trials data.



Instructions
Shaping
Shaping + $$
Mean Latency Correct (ms)
600
550
500
450
400
350
300
p
5
8
10
15
20
25
28
30
35
Day

Figure 18. Subject N02's accuracy (%C), mean latency correct (MLC), and standard deviation of latency correct (SDLC) training data. Note, the scale in parentheses on the primary y-axis describes the standard deviation data.

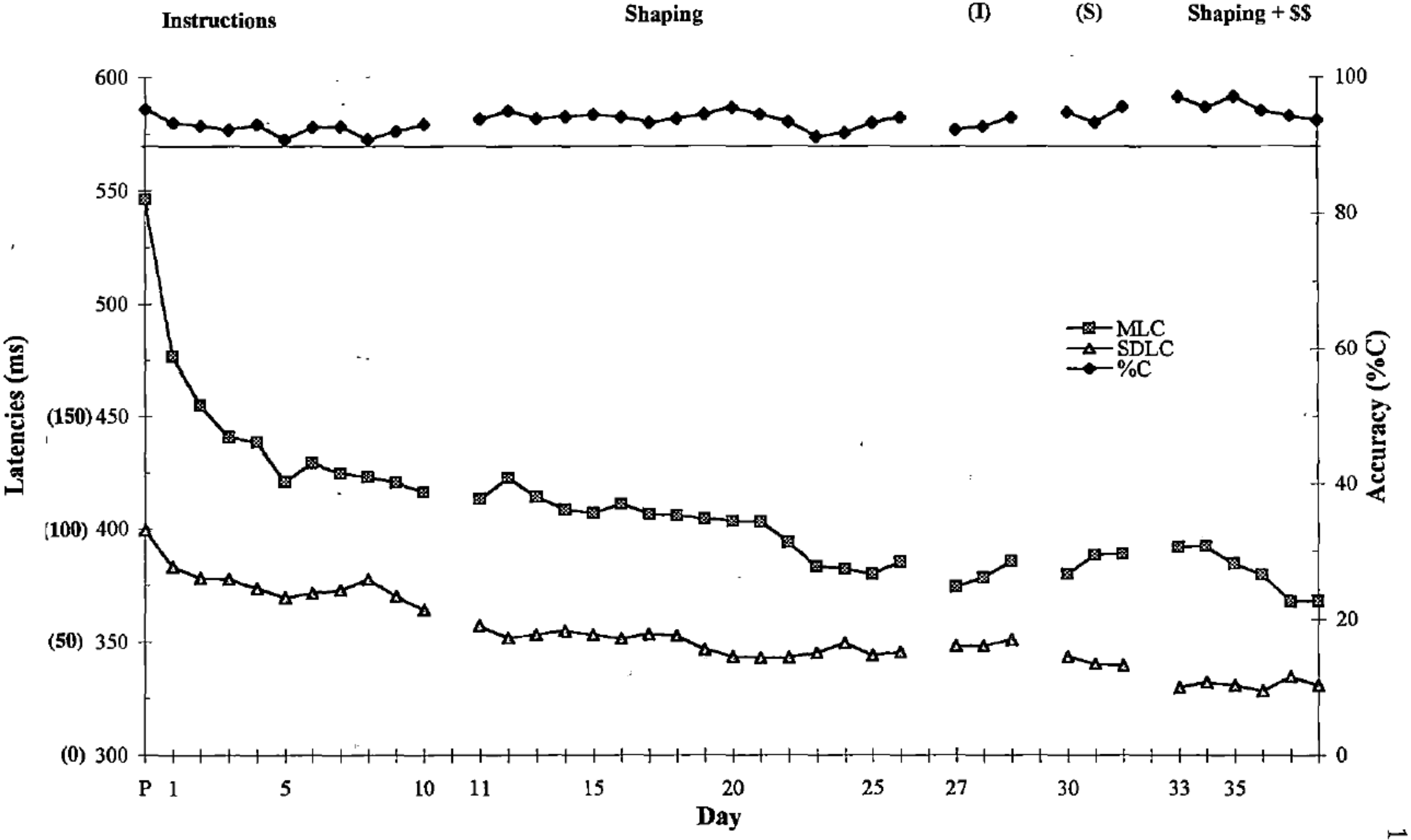

Instructions
Shaping
(I)
(S)
Shaping + $$
Latencies (ms)
600
550
500
(150) 450
(100) 400
(50) 350
(0) 300
Accuracy (%C)
100
80
60
40
20
0
MLC
SDLC
%C
P
1
5
10
11
15
20
25
27
30
33
35
Day

Figure 19. Subject N02's latency on 100% accuracy trials data.

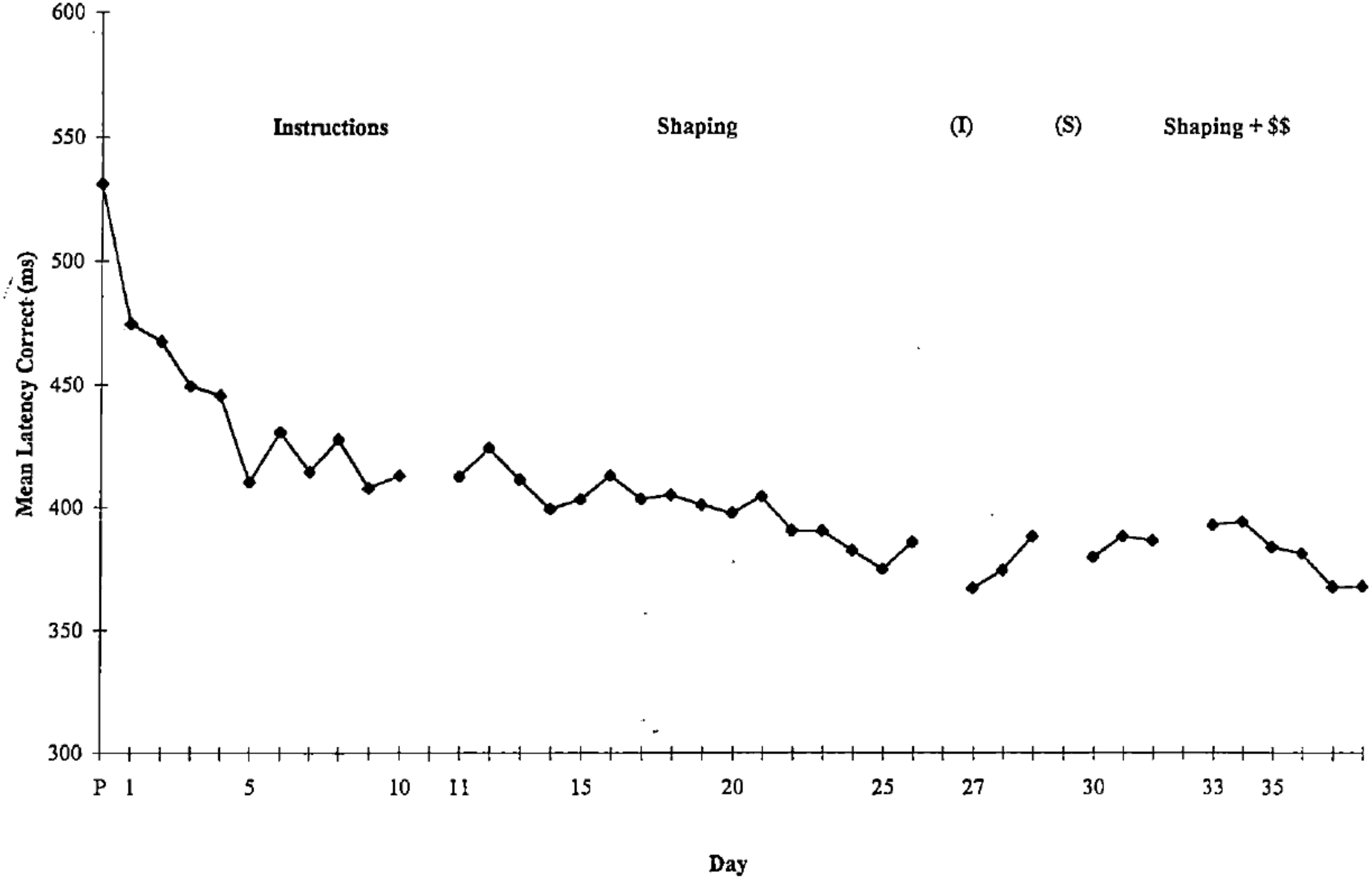

Instructions
Shaping
(I)
(S)
Shaping + $$
Mean Latency Correct (ms)
600
550
500
450
400
350
300
P
1
5
10
11
15
20
25
27
30
33
35
Day

Figure 20. Subject C01's accuracy (%C), mean latency correct (MLC), and standard deviation of latency correct (SDLC) training data. Note, the scale in parentheses on the primary y-axis describes the standard deviation data.

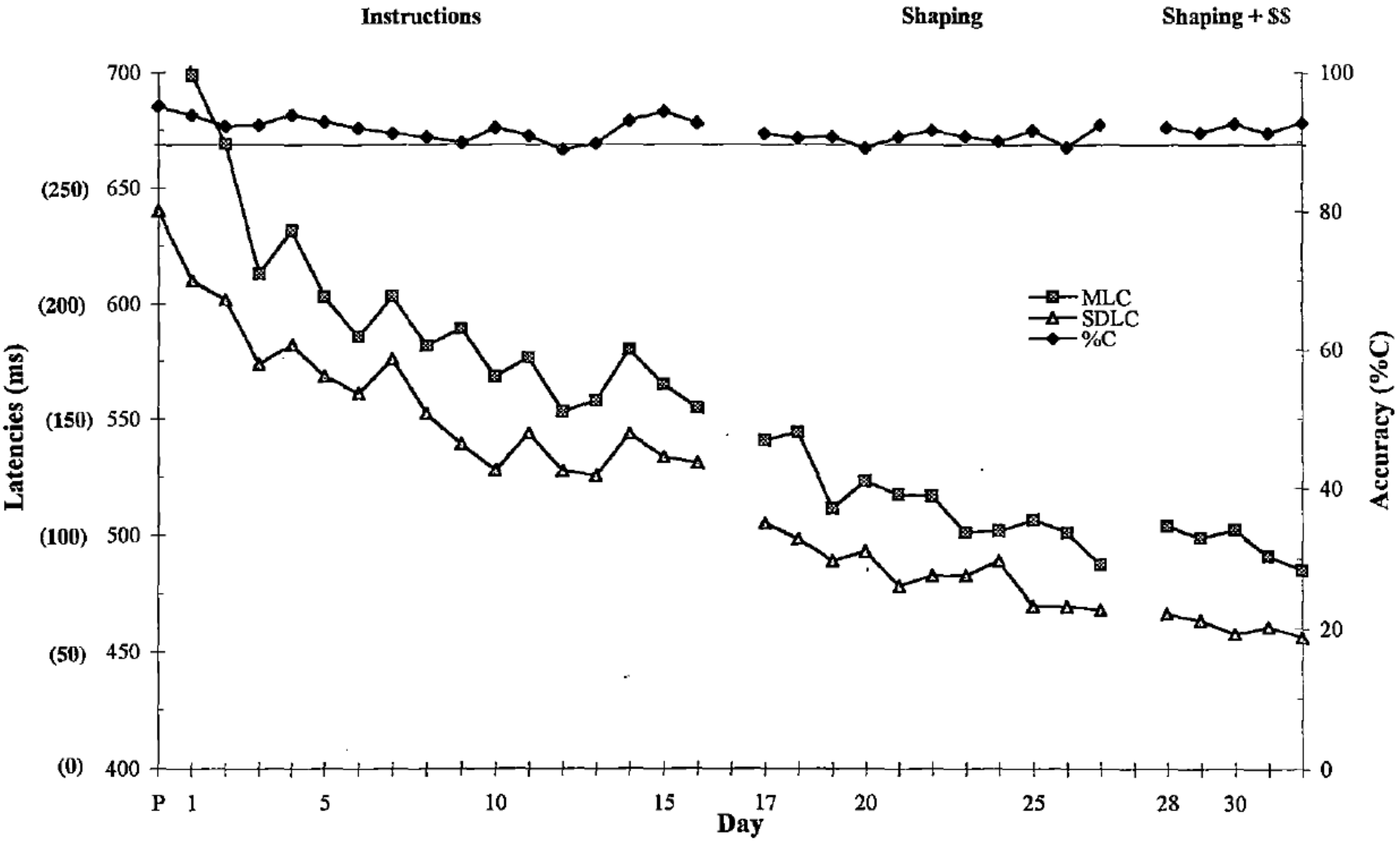

Instructions
Shaping
Shaping + $$
Latencies (ms)
Accuracy (%C)
700
(250) 650
(200) 600
(150) 550
(100) 500
(50) 450
(0) 400
100
80
60
40
20
0
MLC
SDLC
%C
P
1
5
10
15
17
20
25
28
30
Day

Figure 21. Subject C01's latency on 100% accuracy trials data.

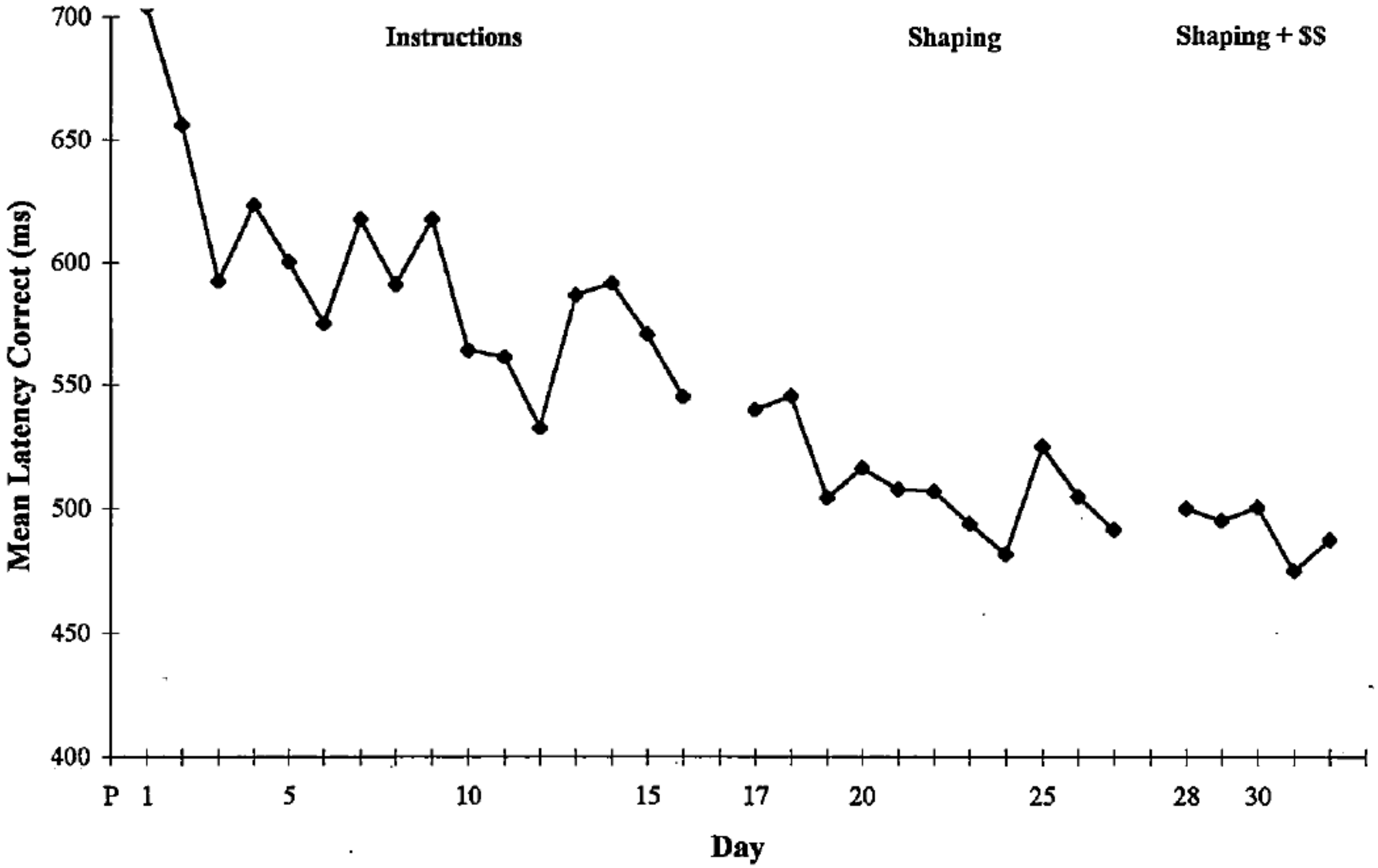

Instructions
Shaping
Shaping + $S
Mean Latency Correct (ms)
700
650
600
550
500
450
400
P
1
5
10
15
17
20
25
28
30
Day

Figure 22. Subject C02's accuracy (%C), mean latency correct (MLC), and standard deviation of latency correct (SDLC) training data. Note, the scale in parentheses on the primary y-axis describes the standard deviation data.

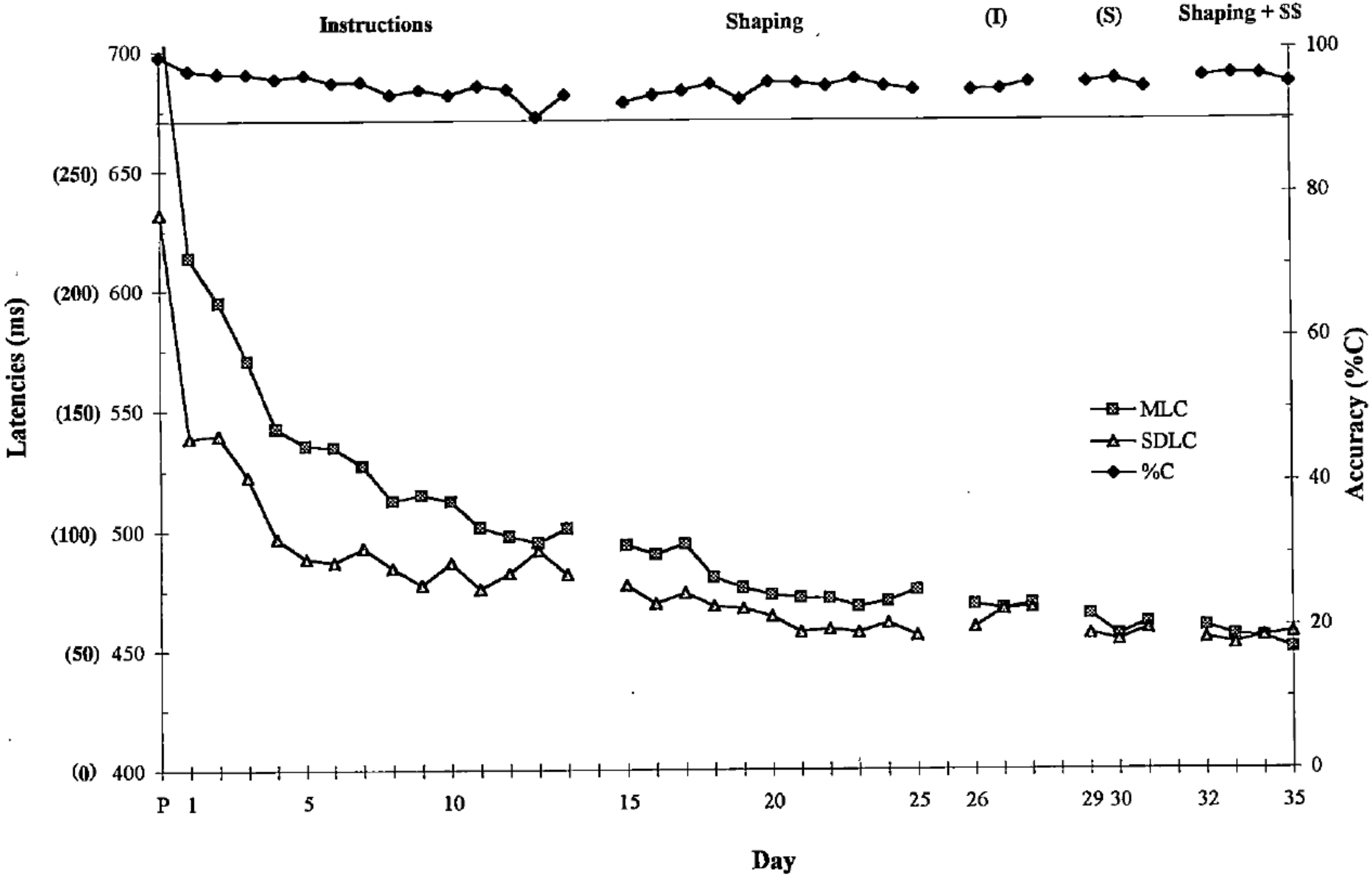
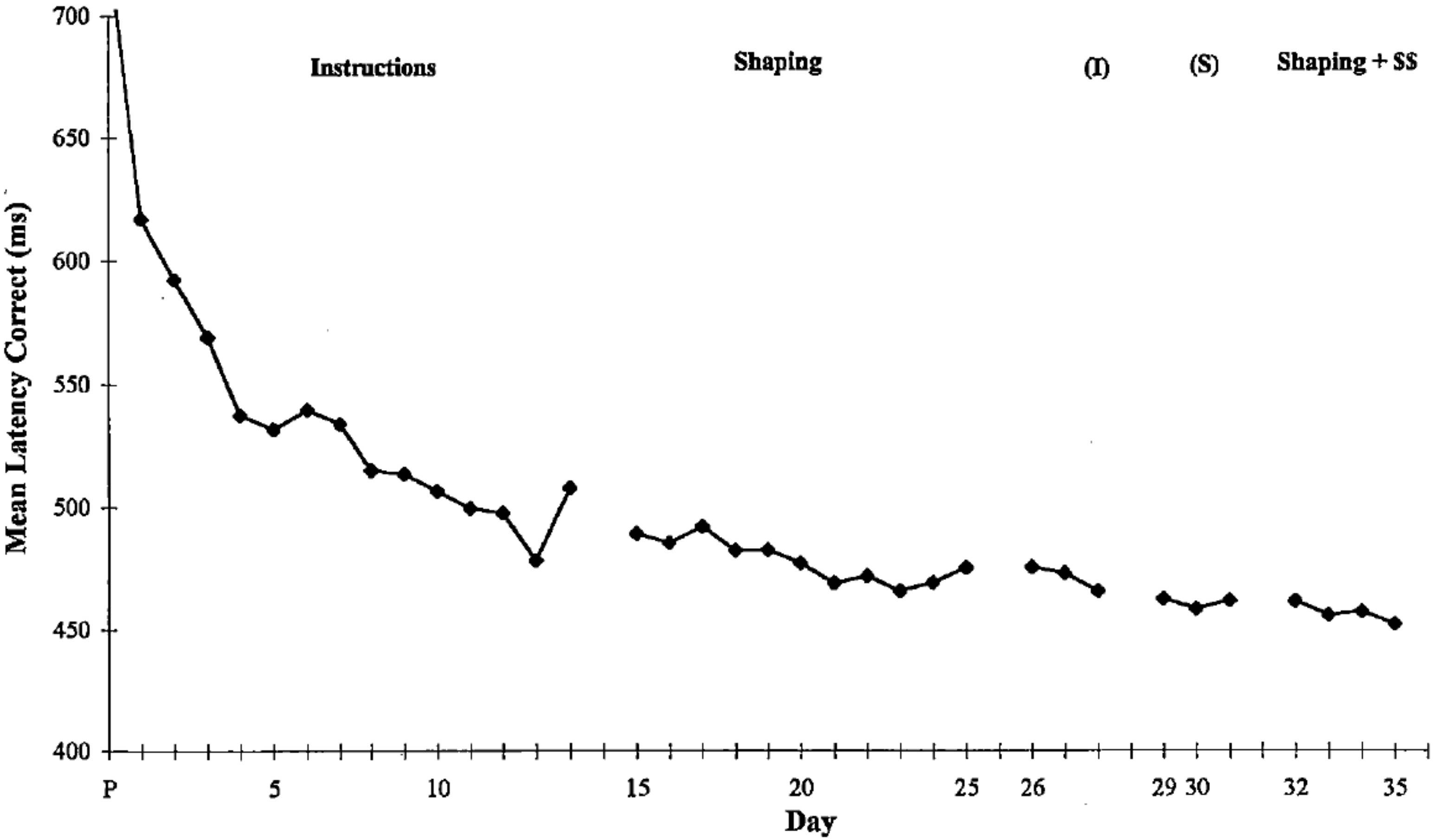

Instructions
Shaping
(I)
(S)
Shaping + SS
Latencies (ms)
700
(250) 650
(200) 600
(150) 550
(100) 500
(50) 450
(0) 400
Accuracy (%C)
100
80
60
40
20
0
MLC
SDLC
%C
P
1
5
10
15
20
25
26
29 30
32
35
Day

Figure 23. Subject C02's latency on 100% accuracy trials data.



Instructions
Shaping
(I)
(S)
Shaping + $$
Mean Latency Correct (ms)
700
650
600
550
500
450
400
P
5
10
15
20
25
26
29 30
32
35
Day

Figure 24. Subject C03's accuracy (%C), mean latency correct (MLC), and standard deviation of latency correct (SDLC) training data. Note, the scale in parentheses on the primary y-axis describes the standard deviation data.

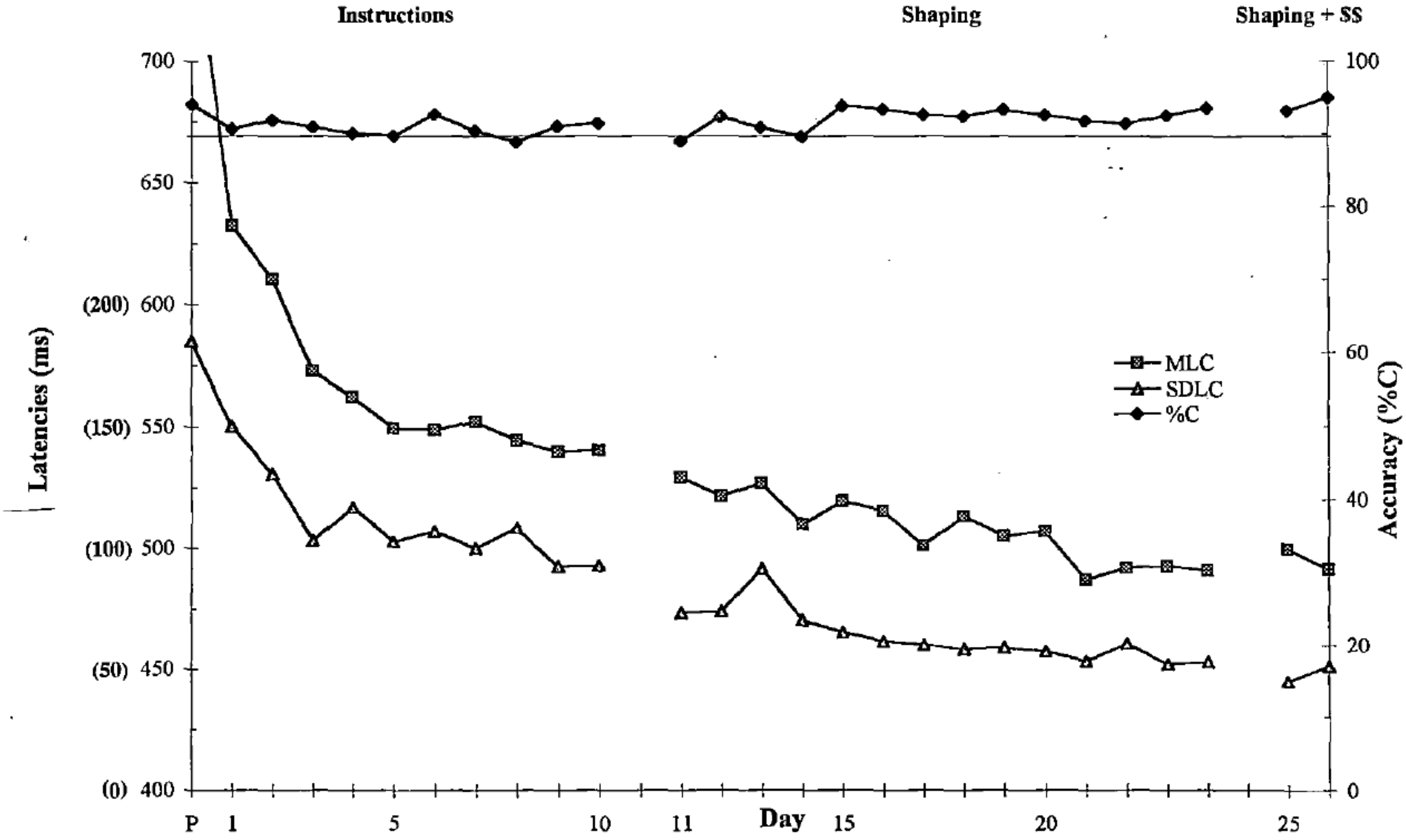
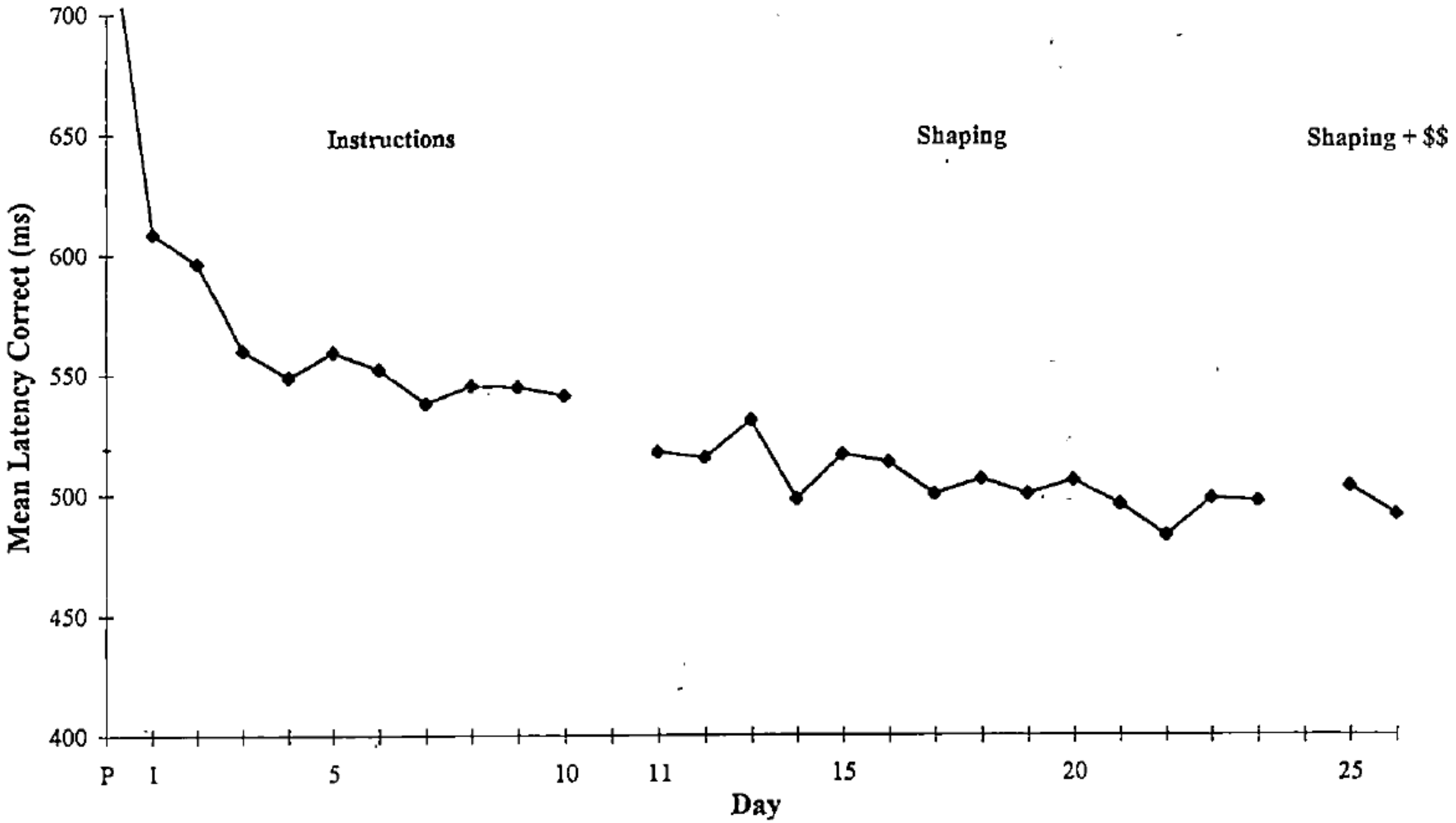

Instructions
Shaping
Shaping + SS
Latencies (ms)
700
650
(200) 600
(150) 550
(100) 500
(50) 450
(0) 400
Accuracy (%C)
100
80
60
40
20
0
MLC
SDLC
%C
P
1
5
10
11
Day
15
20
25

Figure 25. Subject C03's latency on 100% accuracy trials data.


700
650
600
550
500
450
400
Mean Latency Correct (ms)
Instructions
Shaping
Shaping + $$
P
1
5
10
11
15
20
25
Day

Appendix B

The following series of graphs (one per subject) illustrate the accuracy (% C) and mean latency correct (MLC) generalization data for each subject.

<u>Figure 26.</u> Subject N01's generalization data. This graph contains her early and late % C and MLC data, and her % C and MLC data on the training and generalization tasks during the generalization phase.

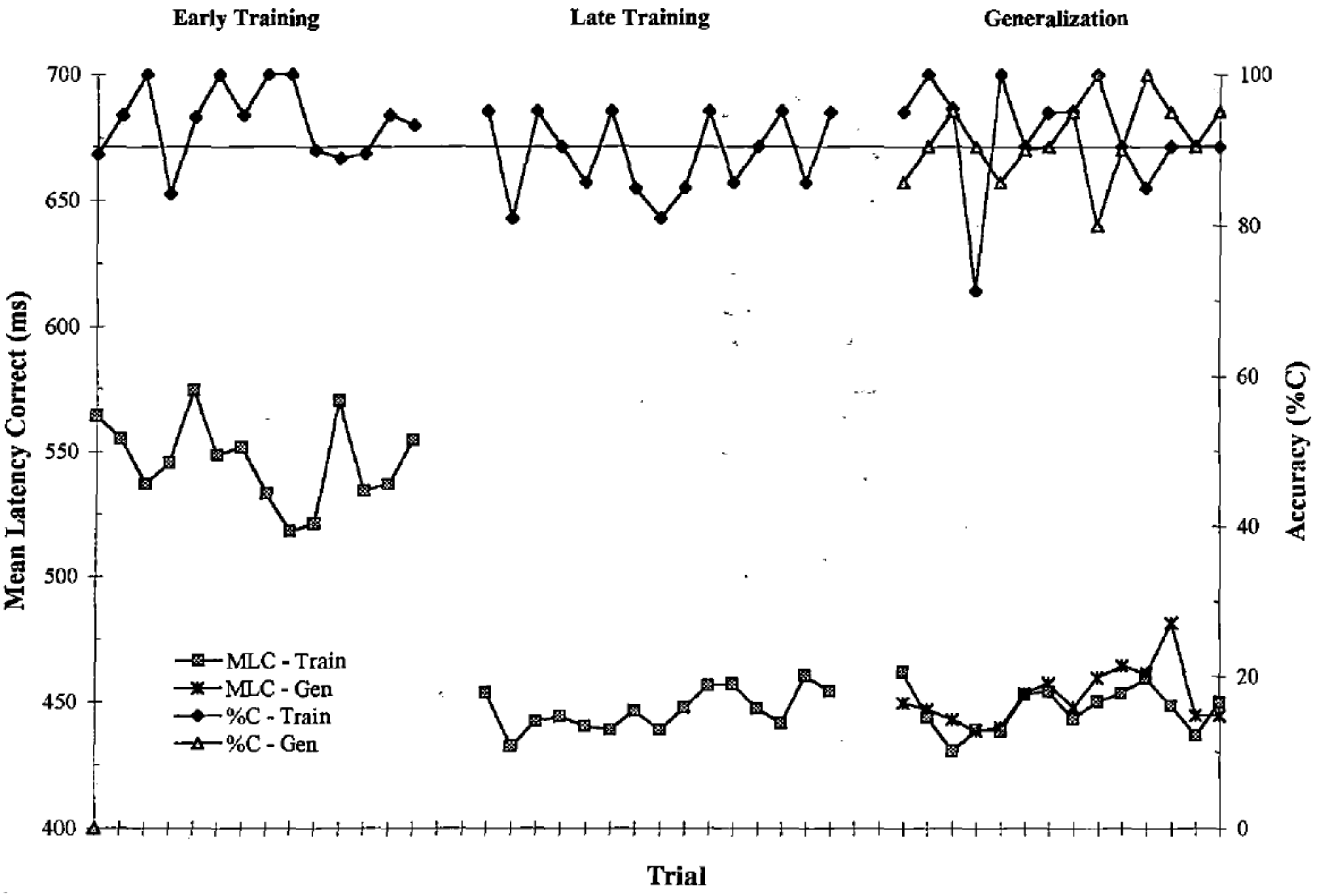

Early Training
Late Training
Generalization
Mean Latency Correct (ms)
700
650
600
550
500
450
400
Accuracy (%C)
100
80
60
40
20
0
Trial
MLC - Train
MLC - Gen
%C - Train
%C - Gen

Figure 27. Subject N02's generalization data. This graph contains her early and late % C and MLC data, and her % C and MLC data on the training and generalization tasks during the generalization phase.

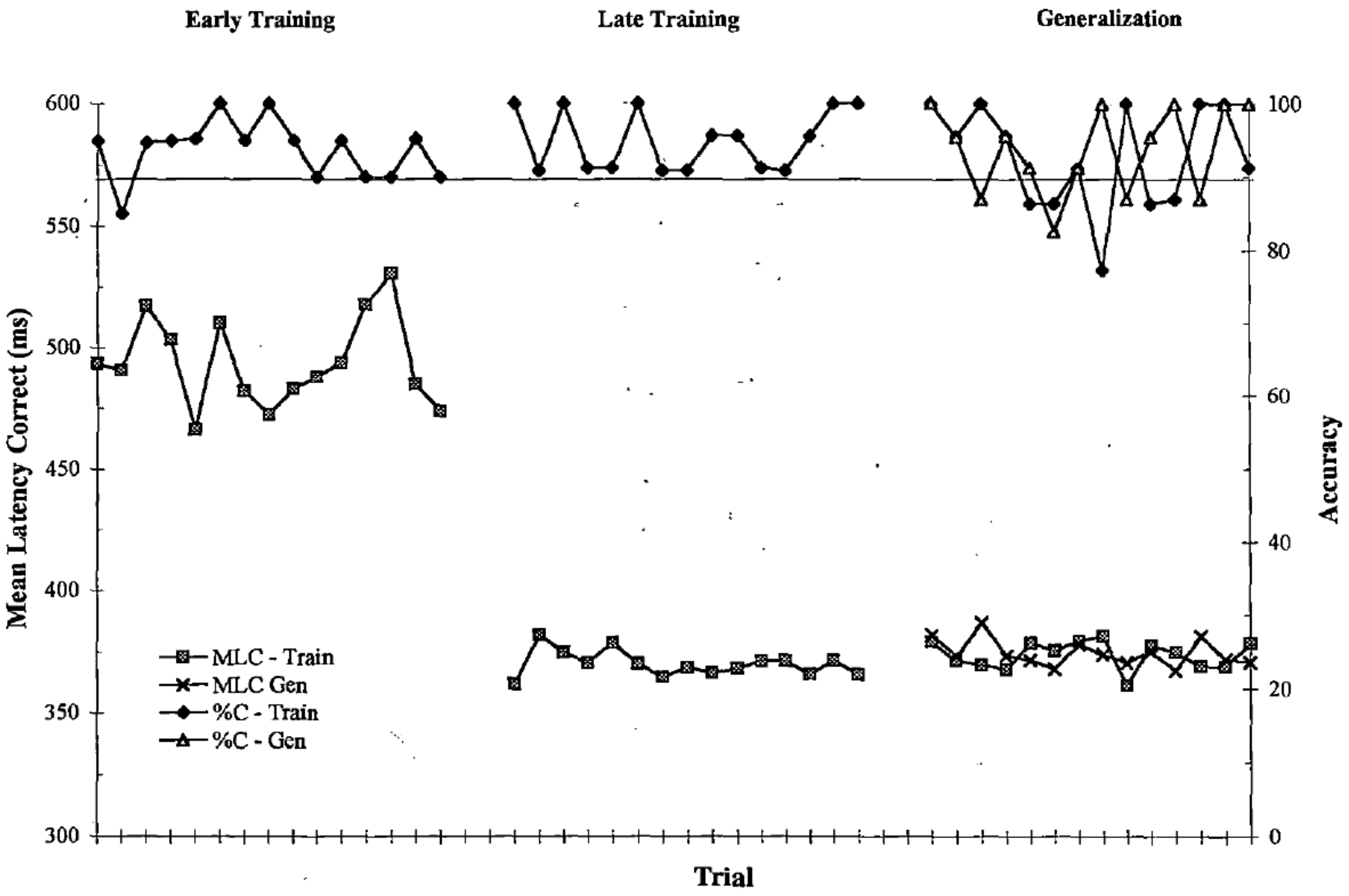
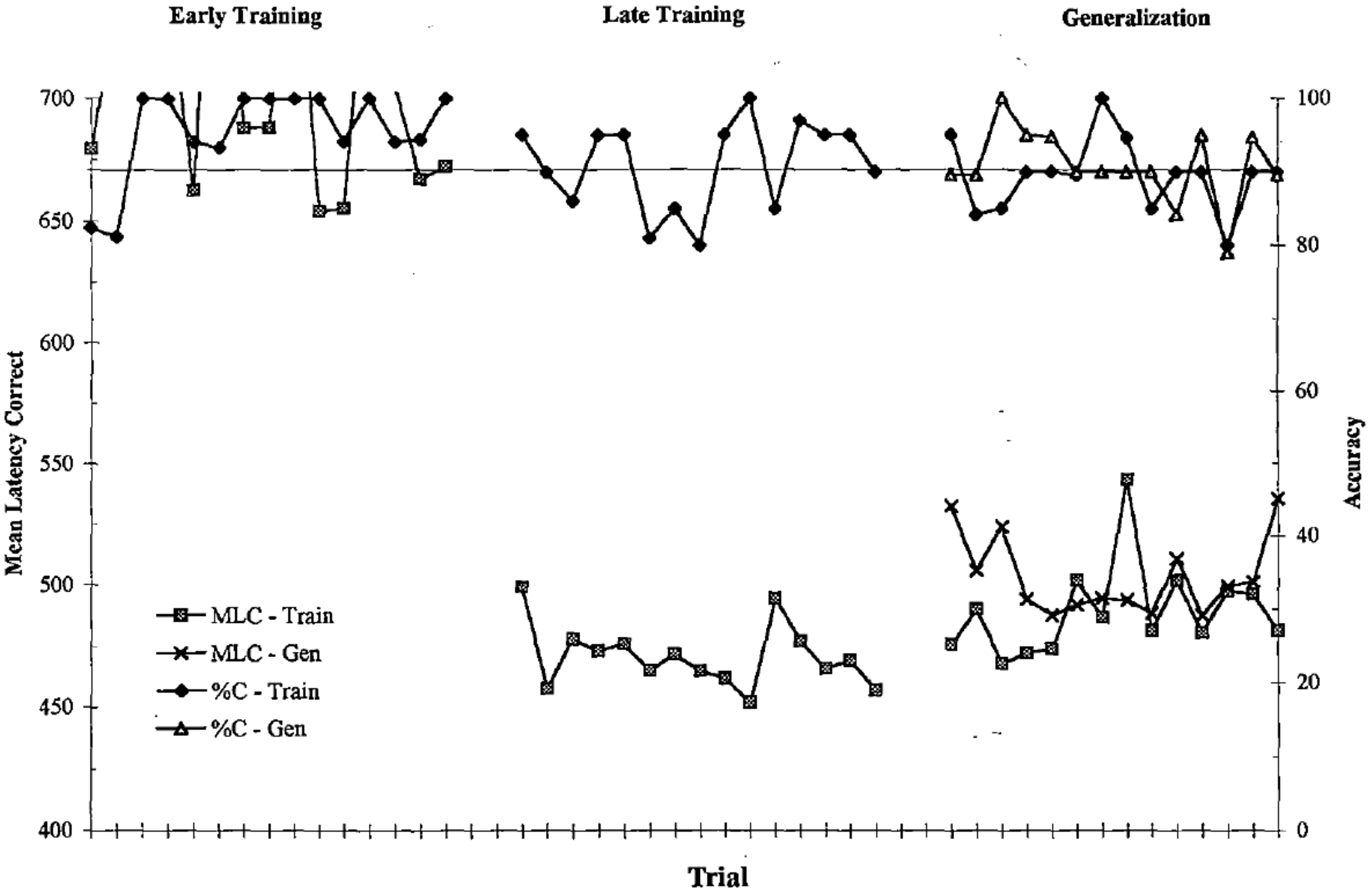

Early Training
Late Training
Generalization
600
550
500
450
400
350
300
Mean Latency Correct (ms)
100
80
60
40
20
0
Accuracy
Trial
MLC - Train
MLC Gen
%C - Train
%C - Gen

Figure 28. Subject C01's generalization data. This graph contains her early and late % C and MLC data, and her % C and MLC data on the training and generalization tasks during the generalization phase.


Early Training
Late Training
Generalization
Mean Latency Correct
700
650
600
550
500
450
400
Accuracy
100
80
60
40
20
0
Trial
MLC - Train
MLC - Gen
%C - Train
%C - Gen

<u>Figure 29.</u> Subject C02's generalization data. This graph contains her early and late % C and MLC data, and her % C and MLC data on the training and generalization tasks during the generalization phase.

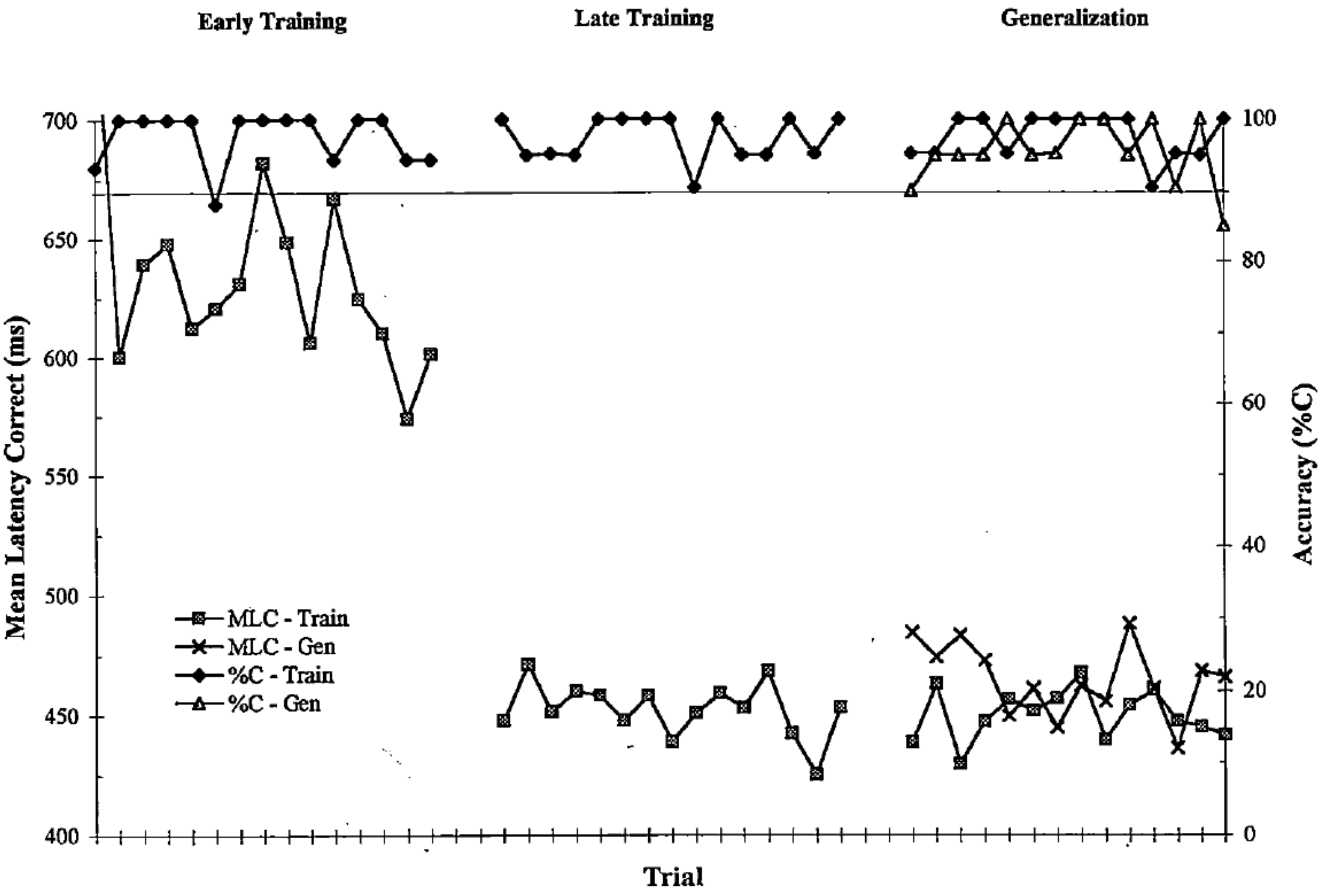
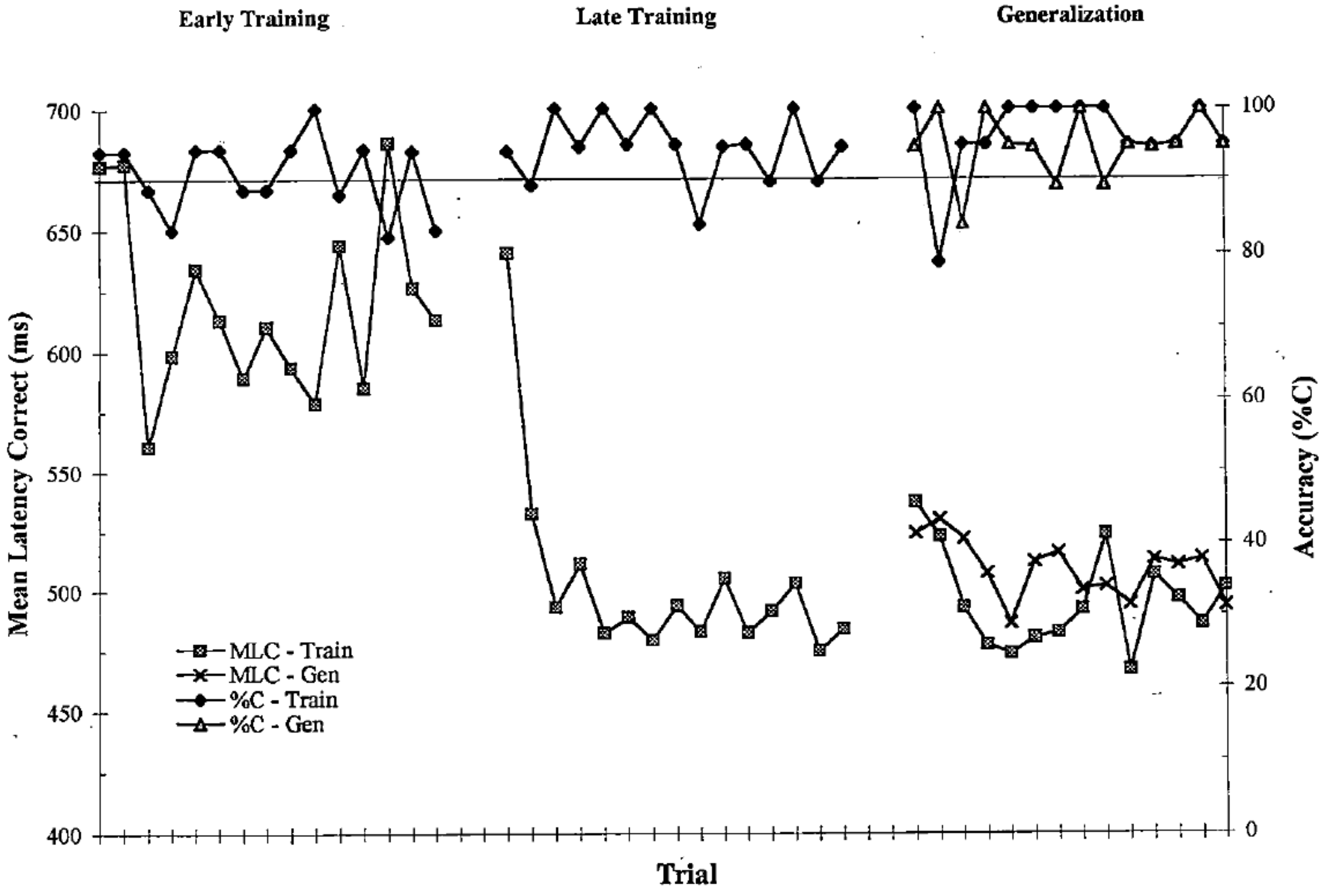

Early Training
Late Training
Generalization
Mean Latency Correct (ms)
700
650
600
550
500
450
400
Accuracy (%C)
100
80
60
40
20
0
MLC - Train
MLC - Gen
%C - Train
%C - Gen
Trial

Figure 30. Subject C03's generalization data. This graph contains her early and late % C and MLC data, and her % C and MLC data on the training and generalization tasks during the generalization phase.


Early Training
Late Training
Generalization
Mean Latency Correct (ms)
700
650
600
550
500
450
400
Accuracy (%C)
100
80
60
40
20
0
MLC - Train
MLC - Gen
%C - Train
%C - Gen
Trial

VITA

Surname: Grabavac Given names: Diana Marie

Place of Birth: Victoria, British Columbia Date of Birth: January 11, 1970

Educational Institutions Attended:

University of Victoria	1987-1995
Université du Québec à Chicoutimi	Spring, 1988

Degrees Awarded:

B.A.	University of Victoria	1987-1991

Honours and Awards:

The Ian McPherson Award, 1995;
NSERC PGS A 1993-1995;
The President's Research Scholarship, 1994-1995;
The President's Research Scholarship, 1993-1994;
The President's Scholarship for Part-Time Undergraduates, 1993;
The President's Scholarship, 1991;
Provincial Government Bursary to Study French in Quebec 1988;
The Rotary Club of Sidney Academic Scholarship, 1987

Publications and Presentations:

Goldwater, B.C., Grabavac, D.M., & Acker, L.E., A "User-Friendly" Recognition Quiz for Encouraging Frequent studying. Submitted for publication.

Grabavac, D.M., Goldwater, B.C., & Acker, L.E., (1995, May). Studying fluency in the lab using choice reaction time tasks. Poster presented a the 21st Annual Association for Behavior Analysis Conference in Washington, DC.

Tan, K., Bredin, J., Polson, D.A.D., Grabavac, D.M., & Parsons, J.A., (1995, May).

The effects of topography- and selection-based training and testing on the emergence of symmetry. Poster presented at the 21st Annual Association for Behavior Analysis Conference in Washington, DC.

Bredin, J., Tan, K., Grabavac, D.M., Parsons, J.A., & Polson, D.A.D., (1994, October). Looking for topography-based symmetry. Poster presented at the 6th Annual Northwest Association for Behavior Analysis Conference in Seattle, Washington.

Goldwater, B.C., Grabavac, D.M., & Acker, L.E., (1992, May). Effects of frequent testing on performance. Poster presented at the 18th Annual Association for Behavior Analysis Conference in San Francisco, California.

Grabavac, D.M., Parsons, J.A., & Polson, D.A.D., (1992, May). The influence of stimulus versus response familiarity on bidirectional learning. Poster presented at the 18th Annual Association for Behavior Analysis Conference in San Francisco, California. (Manuscript in preparation.)

Wong, W.K., Polson, D.A.D., Parsons, J.A., & Grabavac, D.M., (1991, May). Using *Think Fast* to examine variables affecting the efficacy of flash card instruction. Poster presented at the 17th Annual Association for Behavior Analysis Conference in Atlanta, Georgia.

Title of Thesis:

Training Fluency on Choice Reaction Time (CRT) Tasks: A Comparison of Instruction-Based and Contingency-based Procedures

Author [redacted]

Diana Marie Grabavac
January 30, 1996