12-1999

A Comparison of the Effects of Temporal Requirement and Response Form on Retention, Equivalence, and Endurance

Amy Lorraine McCarty
Western Michigan University

A COMPARISON OF THE EFFECTS OF TEMPORAL REQUIREMENT AND RESPONSE FORM ON RETENTION, EQUIVALENCE AND ENDURANCE

by

Amy Lorraine McCarty

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Psychology

Western Michigan University
Kalamazoo, Michigan
December 1999

A COMPARISON OF THE EFFECTS OF TEMPORAL REQUIREMENT AND RESPONSE FORM ON RETENTION, EQUIVALENCE AND ENDURANCE

Amy Lorraine McCarty, Ph.D.

Western Michigan University, 1999

Many researchers and practitioners have reported dramatic learning gains that they attribute to the fluency procedure, often citing rate of response as a critical feature. Others have been unsuccessful at demonstrating these claims empirically. A variety of differences between methods and procedures that may account for the different outcomes are discussed in this paper. In particular is the difference in outcomes when topography-based as opposed to selection-based responses were targeted for training. Fluency practitioners, and the researchers who have obtained similar positive outcomes, have for the most part used topography-based forms of response during the drill and practice sessions, or sprints, that constitute the bulk of fluency training. Researchers who have targeted selection-based response forms have not produced data consistent with the positive findings of the fluency practitioners. This study examines the relative effects of training involving speed versus practice, and topography-based versus selection-based response forms, and the behavioral processes that may account for differences in observed outcomes.

UMI Number: 9954044

ACKNOWLEDGEMENTS

My heartfelt appreciation goes to my advisor and mentor, Dr. Jack Michael, for the encouragement, guidance, and gentle persuasion he provided continually throughout my college education. If there were an award for Teacher of the Millennium, he would be a leading candidate.

I am deeply indebted to Dr. Alyce Dickinson for putting me through the paces of numerous practical exercises in Performance Management, and providing much valuable feedback. Likewise, I extend my thanks to Dr. Dale Brethower whose never-ending supply of creative exercises drove home the meaning of Performance and Instruction.

A tip of my hat goes to Mike Makepeace, a dear friend and source of much needed inspiration.

Finally, I wish to thank my parents, Leon and Joan McCarty, my sisters, June Schutter, Joy Balasko, Lori Maleski, and Sally Rynberg, and the rest of my extended family. Without their loving support and encouragement, this work would not have been accomplished. For them, I am eternally grateful.

Amy Lorraine McCarty

TABLE OF CONTENTS

CHAPTER

CHAPTER

Table of Contents—Continued

CHAPTER

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Fluency and the Fluency Procedure

The importance of fluent responding for effective function, retention and resistance to disruption is increasingly emphasized in the behavioral education literature. Lindsley (1956) described fluency as a high level of accurate performance that is maintained over time without practice, and one in which the resulting skills are applied in new ways. Johnson and Layng (1992) define it as "the rate of performance that makes skills not only useful in everyday affairs but also remembered even after a significant period of no practice" (p. 1476). Binder (1996) describes behavioral fluency as "doing the right thing without hesitation" (p. 164), "the fluid combination of accuracy plus speed that characterizes competent performance" (p. 164), and "that combination of accuracy plus speed of responding that enables competent individuals to function efficiently and effectively in their natural environments" (p. 163). Johnson and Layng (1992) emphasize that "the definition of fluency requires the skill to be available to the selecting environment as a behavior that can be readily linked or combined with other behaviors, thereby allowing students to perform complex tasks and solve complex problems" (p. 1476). With such goals in mind, fluency-based educators have designed and developed effective methods of instruction for teaching students at all levels.

Johnson and Layng (1992), Binder (1996), Lindsley (1995), and Haughton (1980) discuss several criteria for an effective fluency procedure. During training, they are concerned with accuracy, speed, and endurance. After initial training, they are interested in retention and performance during applications similar to those found in natural settings. If, after a period of no practice, a subject's performance doesn't meet predetermined measures of accuracy, speed, and endurance, they recommend increasing the stringency of the criterion, administering more training and/or practice, and further evaluating the performance before moving on. The fluency procedure is a key component in Precision Teaching and the Morningside Model, and reports of these instructional methods describe unprecedented learning gains in a variety of populations (Beck & Clement, 1991; Binder, 1996; Binder & Bloom, 1989; Haughton, 1971; Johnson & Layng, 1992, 1994; Lindsley, 1990; McDade & Goggans, 1993).

Comparison of Traditional Instruction and The Fluency Procedure

What are the main differences between the instructional methods used by fluency practitioners and those used in conventional instruction? And what is it about the fluency procedure that gets better results than conventional instructional methods? Only a careful and thorough empirical investigation can provide definitive answers to these questions. Although such a project is beyond the scope of the current research, it is reasonable to predict that if conducted, it might suggest some of the following explanations.

Accuracy and Speed Criteria

Fluency practitioners attribute much of that procedure's effectiveness to the combination of accuracy and speed criteria imposed during spaced practice sessions. In these sessions, students perform sprints consisting of timed question/answer drills until the combined speed and accuracy criteria are met.

The value of the accuracy criterion is obvious—it strengthens correct responding while eliminating incorrect responses from the learner's repertoire. The fluency procedure's high accuracy criterion (often 100%) also results in more practice than does conventional instruction. In conventional educational settings, 60% or 70% is often accepted, and students move on to new, often more advanced, topics before mastering the current material. Intuitively, one would expect more practice to result in better retention, application, etc.

A rate criterion also generates more practice. In conventional instruction, practice stops at a pre-determined accuracy level. With the fluency procedure, practice continues beyond 100% accuracy until a rate criterion is also met.

Generative Instruction

The fluency procedure is sometimes combined with generative instruction (Johnson & Layng, 1992) in which skills are broken down into their components, and then further broken down into their underlying tool skills. Component skills are taught only after their tool skills are mastered, and composite skills are taught only after mastery of their components. When these two instructional procedures are combined, the student receives a great deal of practice in all components of the target skill, ensuring that the skill in its entirety is firmly established in the student's behavioral repertoire. Generative instruction in combination with a fluency criterion is not commonly used in conventional education.

Direct Instruction

Direct Instruction (Engelmann & Carnine, 1982) is often incorporated into the fluency procedure and probably accounts in part for the beneficial outcomes. The occurrence or non-occurrence of a response consequence in conventional instruction can be somewhat arbitrary (generally left up to the teacher), whereas in Direct Instruction, a consequence is specifically programmed to occur after each response. Additionally, in Direct Instruction, consequences are specifically designed to strengthen or weaken the preceding response. In conventional instruction, a consequence might or might not be academic-related. It might be intended to instruct, to lift a student's spirit, to increase self-esteem, to reduce or increase class participation, or any number of possibilities. As a result, changes in the frequency of accurate responding could be accidental rather than by design. Although some responses will be serendipitously strengthened or weakened, this will be achieved more effectively and efficiently with Direct Instruction.

In Direct Instruction, consequences are programmed to follow responses immediately; whereas the length of response consequence intervals in conventional instruction is usually arbitrary (assignments might be graded and returned to students hours or days after completion). The immediacy of consequences in Direct Instruction increases its power to strengthen accurate responding and weaken inaccurate responding.

Efficacy of Remote Consequences

Features of remote consequences associated with the fluency procedure might also account for some of its beneficial effects. Compared to remote consequences typical of traditional instruction, those incorporated into the fluency procedure involve a shorter interval between student performance and remote consequence, and are more certain, more personal, more public, and more intense.

Delayed feedback in the form of scores plotted on a graph is usually provided after each sprint. Although delayed relative to the time of individual responses, this consequence occurs far more quickly than teachers return graded assignments in conventional educational settings (seconds or minutes as opposed to hours or days).

An integral part of the fluency procedure, this feedback is provided with a greater degree of certainty than in conventional settings, where teachers have the option of withholding feedback, and students are free to ignore it.

After fluency sprints, students are encouraged to plot their own scores, making the delayed consequence more personal than when scores are recorded by the teacher. Sometimes, fluency practitioners post graphs of their students' performance. When not posted publicly, these graphs are usually readily available for viewing by the student, increasing the number of opportunities for the student to be affected (e.g., motivated) by a record of his/her performance. In conventional settings, record books are generally kept in locked desks.

Graphs are a more intense form of stimuli than are scores in record books in that they provide more information (i.e., graphic depictions of performance change). Learners might grasp them more readily they would a table of numbers. This too may increase the probability that the record of performance will affect the student's future behavior.

Praise, corrective feedback, and/or encouragement, often follow practice sprints and accompany the plotting of data. Although these consequences are not necessarily programmed features of the fluency procedure, they might occur more frequently than in conventional settings due to the smaller student/teacher ratio and the greater degree of student/teacher interaction. They might also be more personalized with respect to individual students' performances.

It is not surprising that educators who use the fluency procedure observe greater gains in learning than do those who use conventional methods of instruction. Large bodies of research have demonstrated the value of Generative Instruction (Alessi, 1987; Johnson & Layng, 1994), Direct Instruction (Engelmann & Carnine, 1982), programmed versus nonprogrammed consequences (Barbetta, Heward, Bradley, & Miller, 1994; Comunidad Los Horcones, 1987), immediate versus delayed consequences (Van Houten, Hill, & Parsons, 1975; Van Houten, Morrison, Jarvis, & McDonald, 1974), public versus private feedback (Van Houten et al., 1974; Van Houten et al., 1975), and many of the other distinctive practices associated with the fluency procedure (Binder, 1996; Johnson & Layng, 1992).

CHAPTER II

REVIEW OF RELATED LITERATURE

Research Examining the Effects of High Versus Low Rate

High rate is a critical attribute included in all definitions of behavioral fluency. Following is a review of some of the published reports in which the effects of high and low rate are examined.

In a study by Haughton (1972, as cited in Binder, 1993 and 1996), elementary students' correct and incorrect response rates were recorded as they wrote digits and read (orally) random digits. Later, the same students practiced writing answers to simple arithmetic problems. Only students who performed the earlier tasks at rates around 100 per minute reached rates as high as 50 to 60 per minute on the arithmetic task. And only students who performed the simple arithmetic task at rates as high as 40 to 50 were able progress smoothly through more advanced math curricula.

Orgel (cited in Binder, 1993) conducted a study in which college students used flash cards to practice saying calculus formulas and rules. Students who achieved rates of 50+ accurate responses per minute during practice scored nearly twice as high as did those with slower rates when retention tests were conducted six weeks later.

In an unpublished pilot study by Binder (as cited in Binder, 1996), two adult subjects learned the relationships between arbitrarily paired numbers and Hebrew characters, and then practiced answering simple addition problems written in the Hebrew characters. One subject completed more practice and attained a higher rate than the other subject. In a test session, the subjects wore headphones while attempting to add pairs of Hebrew characters. At certain times, a recording of a voice saying random numbers was played through the headphones. Responses during both conditions were plotted with a cumulative recorder.

Disruption associated with the distracting stimulus was measured by determining the average response rate in the absence of that stimulus, and dividing it by the response rate in its presence. In the cumulative record, visual evidence of disruption took the form of dips in the record between ticks indicating the onset and offset of the distracting stimulus. The performance of the subject who attained higher rates during training showed less disruption during testing. This suggests that high rates of accurate responding during training may be associated greater endurance.

In a similar unpublished study (Binder, 1987, as cited in Johnson and Layng, 1992), numbers were arbitrarily paired with Hebrew characters, and children learned to say the numbers in the presence of their corresponding Hebrew characters. They continued to practice long after reaching 100% accuracy, but without a speed criterion. During a test session, the children attempted to add pairs of Hebrew characters. A recording of a voice saying random numbers was played at certain times through headphones worn by the children. The distracting stimulus "completely disrupted their adding performance" (p. 1481). After additional practice in which combined speed and accuracy criteria were imposed, the students were able to add the pairs of Hebrew characters at a consistent pace, even when distracting conditions were reintroduced.

White (1986) described an intervention used with a third-grade learning disabled student named Patsy in which pronouncing short vowels was the target response. Patsy practiced the sounds for one minute each day with a goal of 40 sounds per minute (the fluency of her nonhandicapped peers). The program was terminated after she reached her goal for one day, but within a few weeks she began to mispronounce short vowels as she had done before. After a reinstating the same program, Patsy reached her speed goal in 10 days, with no errors in the last nine. The program was again terminated, and she maintained her performance for a few days, after which the problem recurred. In her third attempt to correct the problem, Patsy's teacher doubled her fluency aim from 40 to 80 sounds per minute. The child met the goal in eight sessions at which time the program was stopped, and at the time White's (1986) report was published, Patsy continued to say the vowels correctly during her daily activities.

Research Comparing the Effects of Rate to the Effects of Practice

The studies described in the preceding pages have demonstrated a correlation between high rate and high levels of performance. However, as rate increases, practice also increases. Following are descriptions of some of the published reports in which the effects of rate and practice are compared.

In a four-phase, between-group study, Berquam (1981) compared the effects of a combined practice/speed criterion to those of practice-alone condition with measures of retention as the dependent variable. Thirty-four third graders served as subjects, learning the relationships between 10 three-letter trigrams (e.g., "LAH") and corresponding one-digit numbers (0 through 9). The target response was writing a one-digit number. Phase 1 was divided into two stages. In the first stage, all students practiced 100 problems, matching (by writing) the numbers to their trigrams on a practice sheet. The practice sheet consisted of 100 trigrams printed in 10 columns of 10 on a page (each trigram appeared 10 times, randomly placed on the page). During this stage, a legend (a correct depiction of each pair, i.e., GAC 1, MUN 2, NUR 3...LAH 0), was printed along the top of the practice sheet. An accuracy check was performed after each practice session in which students wrote the numbers corresponding to each of the 10 trigrams without the legend.

The second stage of Phase 1 began after the median accuracy for both groups reached 100%. During this stage, subjects continued to use 100-problem practice sheets, but no legend was provided. The experimental group engaged in five daily 1-

min practice sessions, during which they were encouraged to increase their response rate. (The manner in which rate was emphasized was not described in the report.) The control group completed one practice sheet per day for four days with no time constraints. On the fifth day, the control group performed a 1-min session exactly as did the experimental group. Although the number of practice sessions for each group was equivalent, the number of problems within the sessions varied. The average number of problems for control subjects during this stage was 450, whereas the average number for experimental subjects was 279. The average total time on task was 8.4 minutes for control subjects and 5 minutes for experimental subjects.

Phase 2 consisted of a two-week retention period at the end of which another 1-min frequency check was performed. During Phase 3, both groups engaged in one 1-min practice session per day for five days during which frequency of response was emphasized. Phase 4 involved another two-week retention period ending with a 1-min frequency check.

Although subjects in the experimental group received less practice during the initial learning phase than did those in the control, they achieved significantly higher rates of correct responses on both retention tests (44.3 and 56.1 responses per minute, compared to 38.7 and 41.9). After the relearning session, in which both groups received fluency practice, the control group's scores on subsequent retention tests increased from 38.7 per minute to 41.9.

In a study by Ivarie (1986), fourth-grade students were instructed in the relationships between Arabic and Roman numerals, with writing a Roman numeral (e.g., "IV") in the presence of its corresponding Arabic numeral (e.g., "4") as the targeted response. After initial instruction, students engaged in alternating practice sessions and 1-min test sessions, in which they were given a sheet with Arabic numbers and were instructed to write the corresponding Roman numerals. A legend that enabled them to check their answers for accuracy was printed on practice sheets but not on the test sheets.

After three consecutive 1-min tests, the students' scores were separated into two groups: Group One consisted of scores of students who achieved and maintained proficiency rates of 70 responses per minute with 7 or fewer errors on three tests; Group Two consisted of scores of students who achieved and maintained proficiency rates of 35 responses per minute with 4 or fewer errors on the three tests. Mean scores for students in the high rate group were significantly higher (72.12%) than those in low rate group (54.62%) after 3 months of follow-up tests. Within each of the two groups, scores were divided into three subgroups (above average, average, and below average) based on the students' performance on the Iowa Test of Basic Skills.

When analyzed for an interaction effect, data indicated that the performance of students categorized as average or below-average accounted for the superior performance of the high rate group, suggesting that high proficiency rates were particularly beneficial for students in average and below-average achievement categories.

In Experiment I of a three-part study, Kelly (1996) compared mastery learning with a fluency criterion and mastery learning without a fluency criterion using four sets of Dolch[1] words. Saying individual words was the targeted response. One mildly retarded, 5-year-old boy served as a subject, and an ABACABAC design was employed. Under each A (baseline probe) condition, three new Dolch reading words were presented and subject responses were recorded. No reinforcement or correction followed responses during this condition.

Under B conditions, the Dolch words introduced in the preceding A condition were taught in "learn units" in which (a) the teacher modeled the target behavior and prompted the subject to read the Dolch word presented, (b) the subject responded, and (c) the teacher provided feedback (praise, or corrective feedback followed by a repeat of the trial). When the subject completed a ten learn unit session in 13 seconds or less for two consecutive sessions, treatment was terminated.

During C conditions, the Dolch words introduced under the immediately preceding A condition were taught. Treatment proceeded as under B conditions, but with an accuracy-alone criterion. When the subject achieved 90% accuracy or higher

---

[1]E. W. Dolch, a school psychologist, compiled a list of service words that are difficult to depict in pictures and are generally taught as sight words. The list includes pronouns, adjectives, adverbs, prepositions, conjunctions, and verbs such as *these*, *warm*, *soon*, *away*, *but*, and *bring* (Buckingham and Dolch, 1936).

for two consecutive sessions, treatment was terminated. Maintenance probes were conducted one week after the completion of each treatment (B and C condition) phase. These probes were conducted in exactly the same way as the baseline probes (condition A). Maintenance scores were grouped according to their associated treatment conditions and then averaged. The subject's mean accuracy score for words learned under condition B (fluency) was 99%, whereas his mean accuracy score for words learned under condition C (accuracy-alone) was 29.3%.

Kelly (1996) replicated these results in Experiment II of her study with two subjects of different functional levels from each other (one learning disabled and the other speech and language impaired) and from the subject in Experiment I. She used the same procedure, altering only the order of treatment phases. Although results were very similar to those obtained in Experiment I, Kelly noted that the procedure resulted in more instructional time during fluency conditions than during accuracy-alone conditions. This observation led her to conduct a third study, which is discussed below.

To test whether fluency criteria or increased instructional time accounts for gains obtained with the fluency procedure, in Experiment III, Kelly (1996) examined the relationship between mastery learning with fluency criteria and the maintenance of sight words, while controlling for the number of instructional units presented. Two 6-year-old boys served as subjects, one identified as speech and language impaired and ADHD, the other both learning and emotionally disabled.

The emission of spoken words from four sets of Dolch words were targeted for study, and an ABADABAD[2] design was employed. Under each condition A (baseline), three new Dolch reading words were presented and subject responses were recorded. Under condition B, the Dolch words introduced in the preceding condition A were taught in learn units as in her previous experiments. When a subject completed two consecutive ten-learn-unit sessions with 10 seconds per session, treatment was terminated. Under the D condition, the Dolch words introduced in the preceding condition A were taught. Treatment proceeded as in the previous B condition but a yoking procedure was used to ensure an equal number of trials for both treatment conditions. For example, if a subject completed 100 trials during a condition B phase, the subsequent condition D phase would terminate upon the completion of 100 trials.

To control response rate (that is, to ensure a lower response rate), interresponse times were lengthened such that each condition D phase was twice as long as the preceding condition B phase. Baseline and maintenance probes were conducted exactly as in Experiments I and II. Mean accuracy scores for one subject were 99% for words learned under condition B (fluency) and 0% for words learned under condition D (yoked and IRT-controlled). Mean accuracy scores for the other

---

[2]Kelly (1996) uses *D* to refer to the third condition in Experiment III, probably to distinguish it from condition *C* of the first two experiments in the same article.

subject were 99% for words learned under condition B and 30% for words learned under condition D.

In a four-phase, within-subject study, Shirley and Pennypacker (1994) compared the effects of combined accuracy/practice criteria to those of combined speed/accuracy/practice criteria. Two eighth-grade boys diagnosed with learning disabilities in reading and spelling served as subjects. The targeted response was the writing of spelling words. In Phase 1 (Baseline 1) students participated in three sessions a week for four weeks, with a new set of 10 spelling words each week. On Mondays, they read each word aloud, and wrote it on paper. On Wednesdays, they practiced writing each word five times and in a sentence. On Fridays, they took a timed test that was followed by corrective feedback. (Neither the timing procedure nor the method of feedback was described in the report.) Phase 2 (Baseline 2) was identical to Phase 1 except students were given two lists of 10 words each week instead of one.

Two lists of 10 words each were given to students at the beginning of Phase 3 (Treatment 1). Students copied the words and wrote each five times and in a sentence as in the previous phases. Each practice session was followed by a test session the next day. This continued until the student performed with 100% accuracy on either of the two lists. In this way, students were subject to a 100% accuracy criterion for one list, while the amount of practice for the other was yoked to that of the first list. Two new lists were presented, and the procedure was repeated. Phase 4 (Treatment 2) was conducted in exactly the same way as Phase 3 with two exceptions: One hundred percent accuracy was required for both lists, and an additional speed criteria was imposed for one of the lists. (The speed criterion for each subject was set at the rate at which he wrote his own name, measured earlier.)

Follow-up tests conducted approximately 10 days after the end of training were conducted in exactly the same way as the performance tests during training. During these tests, the researchers measured accuracy rates (correct responses per minute) and accuracy ratios (number correct divided by number incorrect). Although the actual scores were not reported in their article, the researchers reported that when comparing between-phase data, no performance differences were apparent. That is, when follow-up data were grouped according to the phase in which the letters were learned, it appeared that subjects performed at about the same rate and level of accuracy regardless of learning conditions. However, when comparing within-phase data, accuracy-alone and fluency conditions appeared to produce better follow-up performance. When comparing follow-up data for letters learned during Phase 3, both students responded at a slightly higher rate of accuracy on letters learned under a 100% accuracy criterion than for practice alone. One student's accuracy ratio was higher for letters learned under the 100% accuracy condition in that phase. When comparing follow-up data for letters learned during Phase 4, both students obtained higher accuracy ratios for letters learned under the fluency condition than under the accuracy-alone condition. Accuracy rates for letters learned under the fluency condition were slightly higher for those learned under the accuracy-alone condition.

Grindle (1997) studied the effects of three levels of fluency on retention over a six-month period. Thirty-three college students were assigned to three groups. Each had an accuracy goal of 95%, but differing speed goals (34-39, 26-31, or 18–23 responses per minute). Using a computer-based training program, subjects answered true/false, multiple-choice, and short-answer questions during training sprints of 41 trials each. Subjects were tested at various times over the next six months. No significant differences between the three groups were observed.

Using college-students as subjects, Makepeace (1998) conducted three experiments to assess the relative effects of accuracy, practice and speed on generalization and co-adduction. In a complex differential matching-to-sample procedure, subjects responded by selecting and clicking squares on a computer screen. Makepeace manipulated criteria to produce varying response rates in the three experiments, and equated the number of responses in two of them. He administered three generalization/adduction tests to each subject upon completion of training. His results indicated that increased practice always led to better performance whether trained to a high rate or a low rate. When the amount of practice was kept equal, he observed no significant performance differences between behaviors taught to high rates and those taught to low rates.

Numerous reports provide compelling evidence that a high rate of accuracy leads to increased acquisition, retention, endurance, and application (Binder, 1993; Binder, 1976, 1979, 1984, as cited in Binder, 1996; Binder & Bloom 1989; Binder, Haughton, & Van Eyk, 1990; Cohen, Gentry, Hulten, & Martin, 1972; Haughton, 1972; Johnson & Layng, 1992, 1994; Kelly, 1996; Orgel, 1984). However, although many published reports emphasize combined accuracy and rate criteria as critical features of the fluency procedure, the added practice generated by the rate criterion might account more directly for the observed gains than the rate criterion itself. In that case, it would be more accurate to emphasize accuracy and practice as the critical components of the fluency procedure, and to discuss rate in terms of practicality rather than criticality. Studies that examined rate independent of practice produced mixed results. When practice was held constant, Berquam (1981), Ivarie (1986), Kelly (1996), and Shirley and Pennypacker (1994) got better test performance with behavior trained to high rates than with behavior trained to low rates. However, their results are not supported by the findings of Grindle (1997) and Makepeace (1998).

The majority of studies do suggest that the imposition of combined rate and accuracy criteria produces greater learning gains than an accuracy-alone criterion (all else being equal). If future empirical research bears this out, a new question arises: What is it about the rate criterion that produces greater gains in learning?

CHAPTER III

ANALYSIS OF THE BEHAVIOR

Direct S-R Relationship Versus Conditional Discrimination

A variety of direct S-R events often occurs between an initial stimulus presentation and a terminal response. In the following example, a person with normal hearing is learning the manual alphabet. After initial instruction of the relationship between the Roman letter *L* and signing the handshape for "L", a sequence of events resulting in an accurate response might proceed as illustrated in Figure 1. In this example, the relationship between (a) private stimuli associated with the sight of *L* and (b) the response of signing "L" is a conditional one—it requires mnemonic behavior and a history of reinforcement for signing "L" in the presence of *L*.

Eventually, a more direct S-R relationship emerges. Private stimuli associated with the sight of *L* begin to evoke signing "L" directly, without any intervening events. The new sequence is depicted in Figure 2.

In some situations, the development of a direct S-R relationship is impossible. A matching-to-sample procedure is an example of such a situation and is illustrated in Figure 3. When presented with a letter, a subject is to select the corresponding handshape from an array of letters from the manual alphabet. Mnemonic behavior can

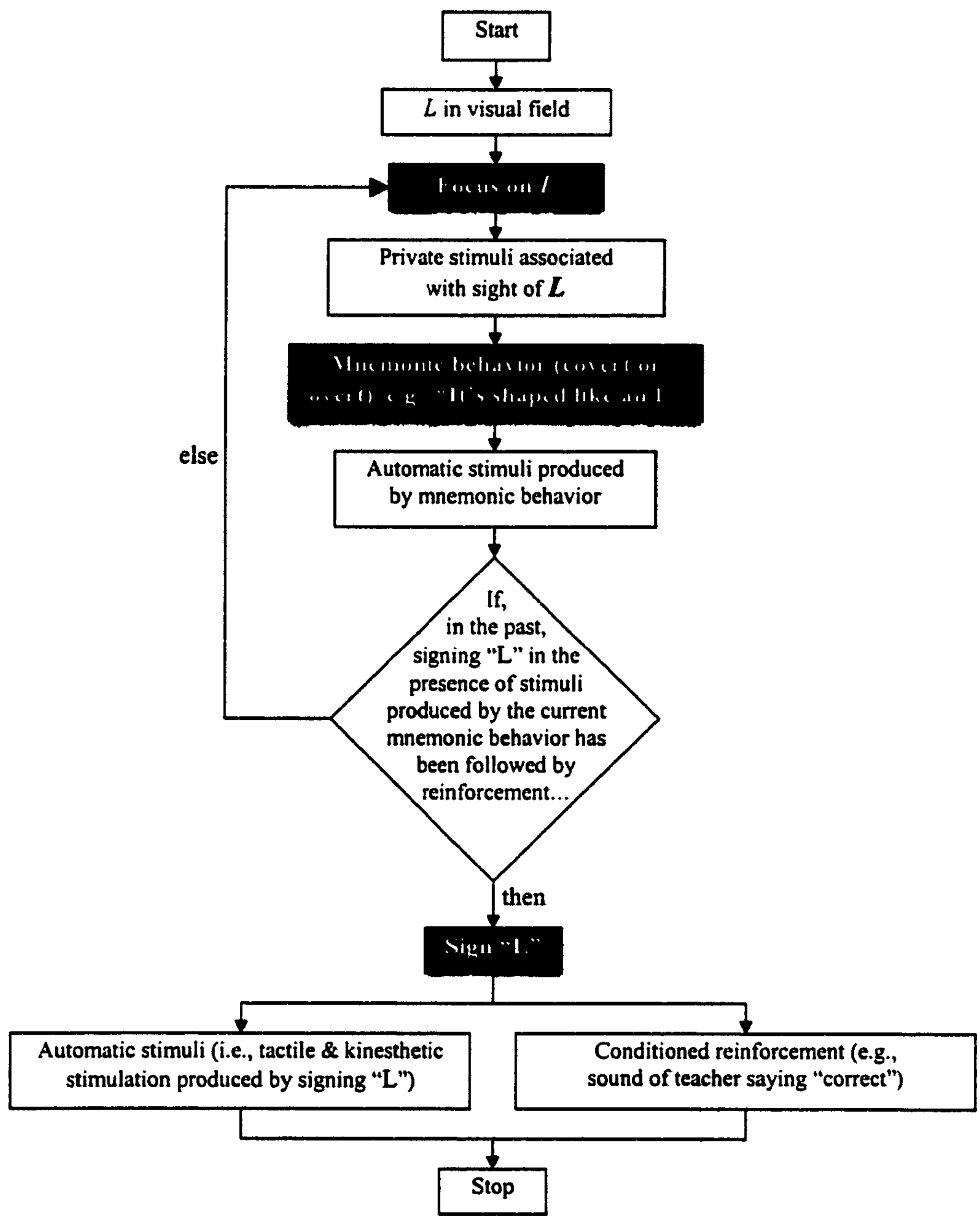


Figure 1. Sequence of Events Involving a Conditional Discrimination, Terminating With an Accurate Topography-Based Response.

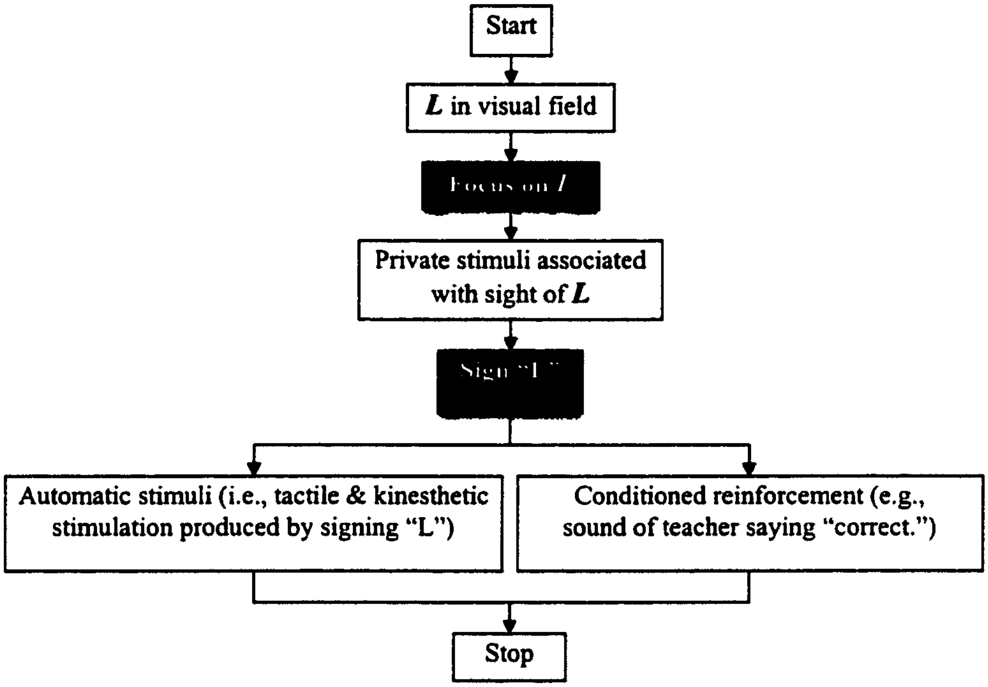


Figure 2. Sequence of Events Involving Only Direct S-R Relationships, Terminating With an Accurate Topography-Based Response.

be phased out as in the previous example. However, in this example, the conditional discrimination cannot be eliminated from the sequence because the target response cannot be emitted without first scanning the array and then focusing on one or more choices.

Imposition of Combined Accuracy/Rate Criteria

When combined accuracy/rate criteria are imposed, such as in the fluency procedure, reinforcement is contingent on fast, accurate responding. As response duration and the duration of intervening events involved in a conditional discrimination occur more quickly, overall response rate increases. Stimuli associated

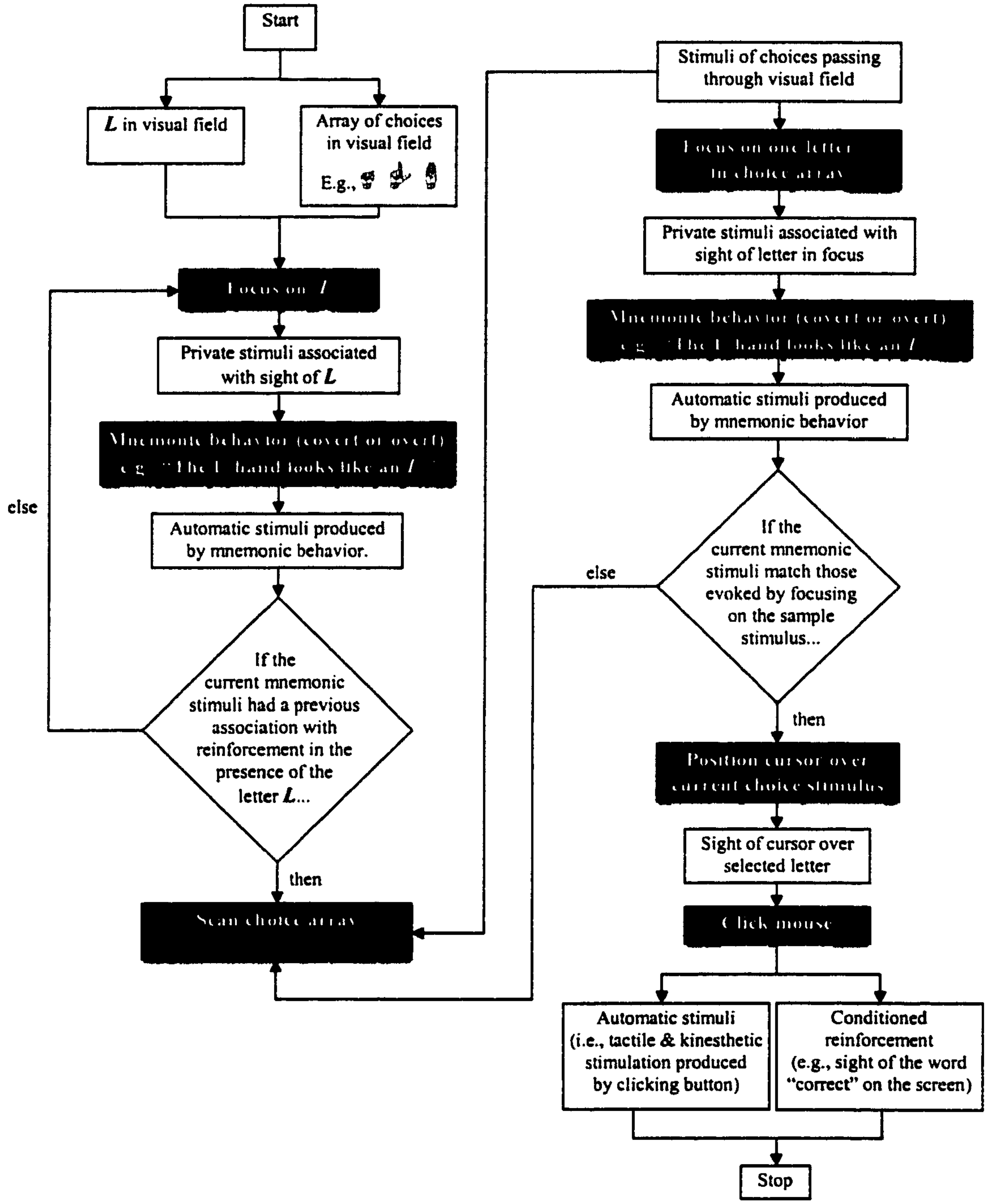


Figure 3. Sequence of Events Involving Two Conditional Discriminations, Terminating With an Accurate Selection-Based Response.

with increasing rate are paired with conditioned reinforcers (e.g., the sound of the word "correct" and the sight of an upward slope on a celeration graph), and the response rate continues to increase until the limits of the individual responder are reached in a ceiling effect.

A conditional discrimination involves a variety of stimuli and a variety of responses. Eventually, if not conditional upon interaction with the external environment, responses that previously occurred later in the sequence events (e.g., those that immediately preceded reinforcement) are evoked by the initial stimuli, resulting in the elimination of some events from the chain. Response rates can begin increasing again at this point due to the decreased number of intervening events. Response rates continue to increase until another ceiling effect occurs. Eventually, S-R relationships that are not conditional upon interaction with the external environment become direct (as in Figure 1), rather than conditional (as in Figure 2). All else being equal, behavioral relations from which conditional discrimination can be eliminated should always produce higher response rates than those for which conditional discrimination is a necessary component.

Retention, Adduction, and Endurance

A direct stimulus-response relationship might be more resistant to extinction than an indirect one. If so, it could account for the increased retention reported in the fluency literature. Intuitively, it seems that repertoires involving direct stimulus-response relationships would also be more readily available to the selecting environment than those involving conditional discriminations. This circumstance, if it exists, might account for the adduction (recombination of component behaviors in new ways without additional training) and the increased endurance (accurate performance under distracting conditions) that is said to occur with fluent repertoires.

Topography-Based Versus Selection-Based Behavior

Michael (1985) distinguishes two types of behavioral relations as topography-based and stimulus-selection-based. Topography-based relations, he explains, are those that involve "*units of behavior* [that] can be described as an increased strength of a distinguishable topography given some specific controlling variable" (p.1). Examples include speaking, writing, signing. Stimulus-selection-based relations (hereon referred to as selection-based) are those that do not involve distinguishable topographies, such as pointing at or touching a stimulus.

Sundberg (1990) and Wraikat (1990, 1991) compared topography- and selection-based relations in terms of ease of acquisition with respect to trials to an accuracy criterion. In all three of these studies, topography-based relations were easier to learn than selection-based ones. McCarty (1997) made a similar comparison, but imposed a fluency criterion (combined speed/accuracy criteria). Her results supported those of the previous researchers, and further suggested that greater fluency may be attained with topography-based repertoires than with selection-based ones.

In a between-group study by Stratton (1992) subjects learned faster under selection-based conditions than topography-based. However, procedural differences might account for his contrary results. Under both conditions, a sample stimulus (a Japanese symbol) was presented on a computer screen. Under selection-based conditions, subjects responded by clicking on a corresponding choice stimulus (one from an array of English words) on the computer screen. Under topography-based conditions, subjects responded by saying the corresponding English word. However, during topography-based phases only, the researcher stayed in very close proximity (within inches) to the subject so he could evaluate responses and provide feedback (the computer performed this function under selection-based conditions). In addition, during topography-based phases, half of the computer screen was obstructed from view of the subject because it displayed current correct answers for use by the researcher in providing feedback. Any combination of variables related to these procedural differences might have influenced the outcome of Stratton's experiment.

Cresson (1994) distinguished subjects (college students) in terms of slow and fast learners, and observed that the slow-learners learned better under topography-based conditions, whereas the fast learners learned better under selection-based conditions. An even greater dichotomy existed between subjects in Sundberg's (1990) and Wraikat's (1990, 1991) studies and those in Stratton's (1992). Developmentally delayed children participated as subjects the first two, and college students participated in the third. The results of these studies support the findings of Cresson (1994) when considered in terms of an inferred capacity for fast or slow learning.

Like Stratton (1992), McCarty (1997) used normal-functioning adults as subjects, but designed the experiment to avoid the procedural differences between conditions discussed earlier. She demonstrated that when combined accuracy and speed criteria are imposed, topography-based performance requires fewer trials to criteria than does selection-based. Subjects learned arbitrary relationships between Kanji symbols and letters of the manual alphabet. A new set of relationships was introduced in each of several sessions using a repeated-acquisition procedure. Under selection-based conditions, when a Kanji symbol appeared on the computer screen, a correct response consisted of selecting and clicking the corresponding handshape from an array of handshapes on the screen. Under topography-based conditions, when a Kanji symbol appeared on the computer screen, a subject responded by making the corresponding sign with her own hand. Each session continued until the subject completed 24 consecutive correct responses with individual response latencies of 1-sec or less, or until 200 trials were completed. Subjects met the combined accuracy/speed criteria in 92% of the topography-based sessions, whereas they met those same criteria in only 19% of the selection-based sessions.

In that study, subjects reported looking for images resembling real objects in the Kanji symbols that they could relate to the corresponding handshape (e.g., one reported seeing the head of a bull in a Kanji symbol which related to the 

handshape). All three subjects reported using mnemonics throughout all selection-based sessions. Subject 3 said she stopped using them under topography-based conditions because they became unnecessary. Subject 2 said she "[thought] of the mnemonic quicker" (p. 29) when making the handshapes than when clicking on them. Subject 1 continued to use mnemonics during all topography-based sessions.

As a function of practice, the speed criterion, a combination of the two variables, or some other variable(s), intervening events (such as those involved in mnemonic behavior and/or those inherent in the matching-to-sample procedure) dropped from two of the subjects' topography-related response chains, leaving fewer events between initial stimulus and terminal response. This enabled a more direct stimulus-response relationship (see Kanji symbol→make handshape).

With practice, mnemonic behavior during selection-based conditions might have decreased. However, an interesting question is "why was there no apparent decrease in mnemonic behavior under selection-based conditions?" One possibility is that events associated with selection-based behavior (e.g., eye movements, changing visual stimuli, hand movements, etc.) might have interfered with processes that would otherwise have brought about this decrease. However, even if the subjects had eliminated mnemonic behavior from their selection-based response chains, that type of response still required a conditional discrimination (see Kanji symbol→scan screen→see picture of handshape→click picture) that would have prevented them from achieving speeds comparable to those attained with topography-based responses.

Although McCarty (1997) did not measure retention in that study, she did measure latency. Those latencies illustrate an interesting phenomenon that is particularly relevant to the notion of fluency. On several of the graphs for topography-based sessions (p. 37-39), it can be seen that latencies decreased steadily and remained at very low levels compared to those of selection-based sprints. Three examples are depicted in Figure 4. In the selection-based data, latencies decreased early in each session but showed much more variability as the session progressed than did topography-based latencies.

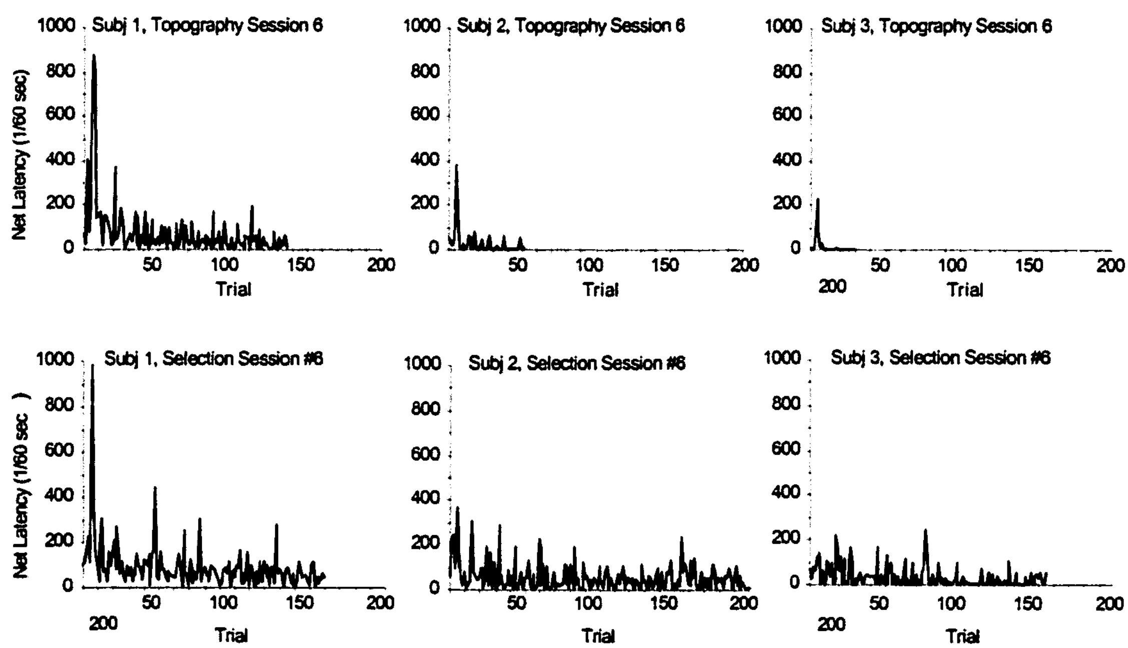


Figure 4. Net Latencies of Individual Trials for Repeated Acquisitions of Topography- and Selection-Based Tasks.

Source: From *Comparing selection- and topography-based fluency using repeated acquisition* (p. 37-39) by McCarty, A., 1997. Unpublished master's thesis, Western Michigan University, Kalamazoo. Adapted with permission.

As discussed above, some events associated with mnemonic behavior eventually dropped out of the behavioral chain. This would have enabled direct S-R relationships to develop, the occurrence of which seems to be indicated by the steady and persistent decrease in latencies. No corresponding steady and persistent decrease occurred under selection-based conditions, possibly because the nature of that type of relation prohibits the complete elimination of intervening events associated with conditional discriminations from the S-R sequence.

Implications of Response Form in Fluency Research

Increases in rate (or more precisely, decreases in latency) seem to be indicators that the relationship between S and R is becoming more direct, involving fewer intervening events. Perhaps the target S-R relationship becomes stronger and more resistant to extinction as a result of the closer relationship. Empirical research comparing the effects of rate, independent of practice, has produced inconclusive results. A key issue might be topography- versus selection-based behavior. If a rate criterion facilitates learning by forcing the rapid elimination of intervening variables, it would do so most effectively if the target behavior were topography-based.

Although they were not studying topography- versus selection-based behavior directly, McDade, Austin, and Olander (1985) produced results that tend to support this idea. They compared the use of SAFMEDS to computer-based fluency drills, using 33 college students as subjects. With SAFMEDS (an acronym for "Say All Fast

A Minute Each Day Shuffled") subjects answered (vocally) the questions on the cards as quickly as possible while a tester timed them and provided accuracy feedback. Under computer-based conditions, subjects responded to multiple-choice questions presented on the computer. A mastery-criterion was imposed under both conditions. Several units of two college courses were tested in this way, and students were allowed to choose which test condition would be in effect for each unit, as long as the units were evenly distributed between conditions. Students were also permitted to stop upon reaching the predetermined mastery criterion of 15 correct responses per minute, or to continue to a higher level of mastery.

With respect to preference, results indicated that students preferred the computer-based condition. When given a choice, they usually chose the computer-based condition, getting the SAFMEDS condition for a later unit by default. Almost twice as many students opted to practice beyond mastery in the computer-based condition than in the SAFMEDS condition. On the other hand, whereas 31 of the 33 students performed more trials under the computer-based condition than under SAFMEDS, 29 of them "showed higher best performances on SAFMEDS" (p. 50). Performances under both conditions were very high, usually double the minimum mastery criterion.

The researchers in the above study did not distinguish the target behavior in the two conditions in terms of selection- versus topography-based, nor did they attribute the better performance under SAFMEDS to response form. They did note that "when higher fluencies, especially verbal ones, are desired, SAFMEDS would be preferable to computer-generated tests" (p. 50). It is conceivable that performances during the SAFMEDS condition were better because topography-based "saying" behavior enabled direct S-R relationships to develop between the initial stimuli (the questions on the cards) and the terminal behavior (the spoken answers). The target behavior under computer-based condition was selection-based, requiring a conditional discrimination. The poorer performance under this condition might be more appropriately attributed to the form of response rather than to the fact that it was computer-based. If voice activation software had been available at the time, topography-based behavior might have been possible under both conditions, and results might have favored computer-based conditions.

In the studies reviewed in Chapter II, where the effects of rate were compared to those of practice, only two targeted selection-based behaviors during fluency sprints. During the fluency sprints in Makepeace's study, only selection-based responses (clicking an image on a computer screen) were targeted. Thirty of the 41 target responses in Grindle's (1997) study were purely selection-based (clicking true/false or multiple choice selections on a computer screen). The remaining eleven responses targeted by Grindle required three keystrokes each. For experienced touch-typists, each three-keystroke combination might have constituted one topography-based response. For less-skilled typists, this combination of keystrokes might have constituted three selection-based responses. In neither of these two studies were greater learning gains observed with high- versus low-rate behaviors.

The remainder of the studies reviewed targeted topography-based behaviors during fluency sprints. Berquam (1981), Ivarie (1986), and Shirley and Pennypacker (1994) targeted writing. Binder (1979), Kelly (1986) and White (1986) targeted saying. In all of these studies, greater learning gains were observed with high-rate behaviors than with low-rate ones.

In a comprehensive review of the fluency procedure, Binder (1996) described several studies in which the fluency procedure produced learning gains, all of which employed topography-based responses such as saying, writing, assembling, dressing, and typing bursts on a keyboard. During an internship at Morningside academy, I observed much saying and writing during fluency sprints, but never any selection-based behavior.

Purpose of the Present Study

Results of studies in which learning gains are achieved using the fluency procedure are compelling. However, the underlying processes by which these gains are achieved have not been thoroughly examined. The extent to which the fluency procedure stimulates learning might depend heavily on whether the form of the targeted response is topography- or selection-based. The purpose of the current study was to identify and measure the relative effects, if any, of topography- versus selection-based response forms and fast versus slow speed goals on retention, endurance, equivalence, and strength of response form.

CHAPTER IV

METHOD

Subjects

Five adult women served as subjects. All were already proficient users of a computer mouse, and none had experience with Morse Code or Kanji symbols. Before being accepted for participation in this study, subjects were required to complete two prerequisite skill exercises. These exercises are described later in this paper. Each subject was paid $5.00 for each session she attended.

Protection of Human Subjects

The Human Subjects Institutional Review Board of Western Michigan University Human reviewed and approved subject participation (see Appendix A).

Informed Consent

Each subject read an anonymous consent form (Appendix B) before participating in the study.

Setting

The study was conducted in a 10' x 10' office, furnished with a worktable, a filing cabinet, a bookcase, two chairs, and a desk with a computer.

Apparatus and Software

With the exception of a study sheet and the occasional administration of a paper/pencil test, a microcomputer controlled all experimental events and recording of data. The computer program was designed and developed by the researcher using Toolbook II Instructor (version 5.0, [c]Asymetrix Corp., 1990-1996). Stimuli were presented via the computer monitor, and subjects responded using the computer mouse.

Independent Variables

Each subject was exposed to each of four conditions consisting of different combinations of two levels of two independent variables. During each condition, subjects learned six relationships between Kanji and Morse Code symbols, for a total of 24 relationships. The first independent variable was response form, with the two levels being topography- and selection-based. The second independent variable was performance goal, with the two levels being high-speed accuracy and low-speed accuracy.

Under condition one, a selection-based matching-to-sample procedure (see Kanji symbol→select Morse code symbol) was used, and a 100% accuracy goal was emphasized. Subjects were encouraged to maintain response rates under 15 responses per minute, and to focus on accuracy rather than speed.

In condition two, a topography-based matching-to-sample procedure (see Kanji symbol→tap out Morse code) was used, and a 100% accuracy goal was emphasized. Subjects were again encouraged to maintain response rates under 15 responses per minute, and to focus on accuracy rather than speed.

Under condition three, a selection-based matching-to-sample procedure was used (same as condition one), and a 100% accuracy goal was emphasized. Subjects were encouraged to respond quickly and accurately, as fast as they could manage. On performance feedback graphs, a speed goal of 30 responses per minute was displayed.

In condition four, a topography-based matching-to-sample procedure was used (same as condition two), and an accuracy/speed goal (same as condition three) was emphasized.

Under all four conditions, at the end of each training drill, a feedback graph appeared displaying the subject's performance relative to the current condition's training goal.

Dependent Variables

The accuracy, gross latency, net latency, and response duration of every single trial in each phase of the experiment was measured. With respect to accuracy, responses were counted as either correct or incorrect. Gross latency was determined by measuring the elapsed time between onset of a sample stimulus and start of a terminal response. Net response latencies were determined by subtracting the mean reflexive S-R gross latency (described later) for a particular relationship from each of its corresponding paired associate gross latencies. For selection-based responses, response duration was determined by measuring the time between the press and release of the mouse button. For topography-based responses, response duration was determined by measuring the time between the first button-press of the response and the last button-release of the response. All times were measured in 1/1000ths of a second.

Nine tests examining retention, equivalence, and endurance were scored in terms of percent correct. Additional data were collected and analyzed separately with respect to the type of response form emitted when both selection- and topography-based response forms were possible. These were expressed in terms of percent of responses emitted in the original response form trained.

Procedure for Establishing Kanji/Morse Pairs

In an earlier attempt at this experiment, Kanji symbols were arbitrarily paired with Morse Code characters. However, all of the subjects dropped out the experiment before the first test. The main reason given for leaving was the difficulty of the task—some felt they would never learn the relationships, and others stated it would take more time to learn than they were willing to spend. To decrease the difficulty (and thereby increase the probability that future subjects would complete the experiment), Kanji symbols were paired with Morse Code characters with which they had some features that could be related. For example, in the 名:–•–• pairing, the two curved lines in the Kanji symbol could correspond to the two dashes in the Morse Code, and the two boxes to the two dots.

Assignment of Kanji/Morse Pairs to Experimental Conditions

Two main constraints were considered when assigning the Kanji/Morse pairs to experimental conditions. First, the degree of similarity between a given Kanji symbol and its corresponding Morse character could influence the ease with which subjects learned a relationship. Second, Morse Code characters differed as to the time required to tap them out, which could influence response rate. To minimize the influence of these constraints, the researcher attempted to assign Kanji/Morse pairs to treatment conditions such that (a) each group had an equal number of dots and dashes, and (b) the pairs were equally distributed in terms of subjective levels of similarity within pairs. Although both constraints could not be overcome completely, the pairs were distributed as evenly as possible.

Design and Procedure

This experiment employed a within-subject alternating treatment design. A prerequisite skill training session was conducted on the day before the first experimental session. Only candidates who met predetermined fluency criteria during prerequisite skill training were accepted for participation as research subjects. Three training sessions were conducted on three consecutive days, during which subjects were exposed to each of the four conditions. A test session was conducted the day after completion of the third training session. A second test session was conducted one week later, and a third one week after that. Table 1 summarizes the session schedule.

Table 1

Session Schedule

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
Week 1	Prerequisite Skill Session	Experimental Session #1	Experimental Session #2	Experimental Session #3	Test Session #1
Week 2					Test Session #2
Week 3					Test Session #3

Prerequisite Skills Training

Before beginning the first training session, subjects completed four matching-to-sample fluency exercises, two selection-based and two topography-based. Subjects were advised that they must achieve predetermined fluency criteria during this training to participate further in the study. Subjects were given a fluency goal of 20 responses per minute at 95 percent accuracy for all exercises in this session (see Figure 5). The stated goal was slightly higher than the actual criteria for participation. Subjects who achieved performances of at least 18 responses per minute at 90 percent accuracy were invited to be in the study. Of the seven people who attempted this test, six met criteria and were invited to participate. Of those six, five elected to do so.

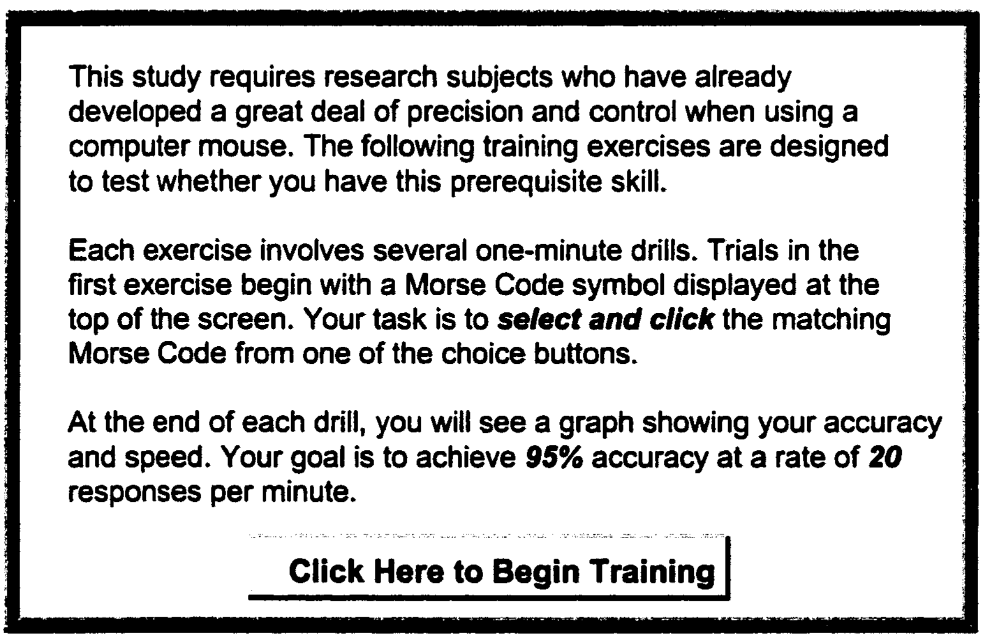


Figure 5. Appearance of Computer Screen Displaying Instructions for Selection-Based Prerequisite Skill Exercise.

During the two selection-based exercises a subject saw a sample stimulus (a Morse Code character) on the computer screen, and an array of choice stimuli (all Morse Code characters) one of which was the same as the sample (see Figure 6). When the subject positioned the cursor over any one of the choice buttons, that button darkened. When the cursor was moved away from a choice, the button lightened. A press and release of the mouse button while the cursor was over one of the choice stimuli constituted one response.

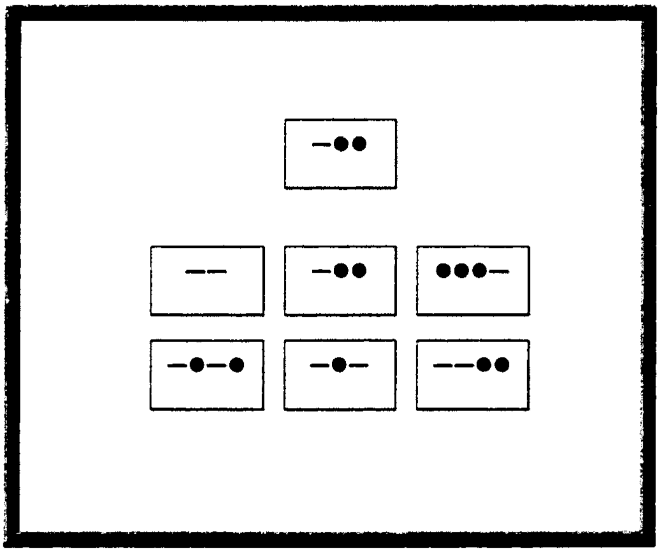


Figure 6. Appearance of Computer Screen Displaying Sample and Choice Buttons for Selection-Based Prerequisite Skill Exercise.

Incorrect responses were followed by the brief appearance of the word "incorrect" on the computer screen, accompanied by a computer-generated "beep" sound, an illustration of the correct relationship, and a repeat of that trial. Correct responses were followed by the brief appearance of the word "correct," and then a new trial with different sample and choice stimuli.

This training proceeded in 1-min drills. At the end of each drill, the subject saw a graph depicting her performance relative to the training goal (see Figure 7). A selection-based exercise ended after the fluency goal was met, or when nine drills were completed, whichever occurred first, and was followed by a short break. Two selection-based exercises were conducted in this manner. If a subject met the fluency goal on any of the drills in either of the selection-based exercises, she then continued to topography-based exercises.

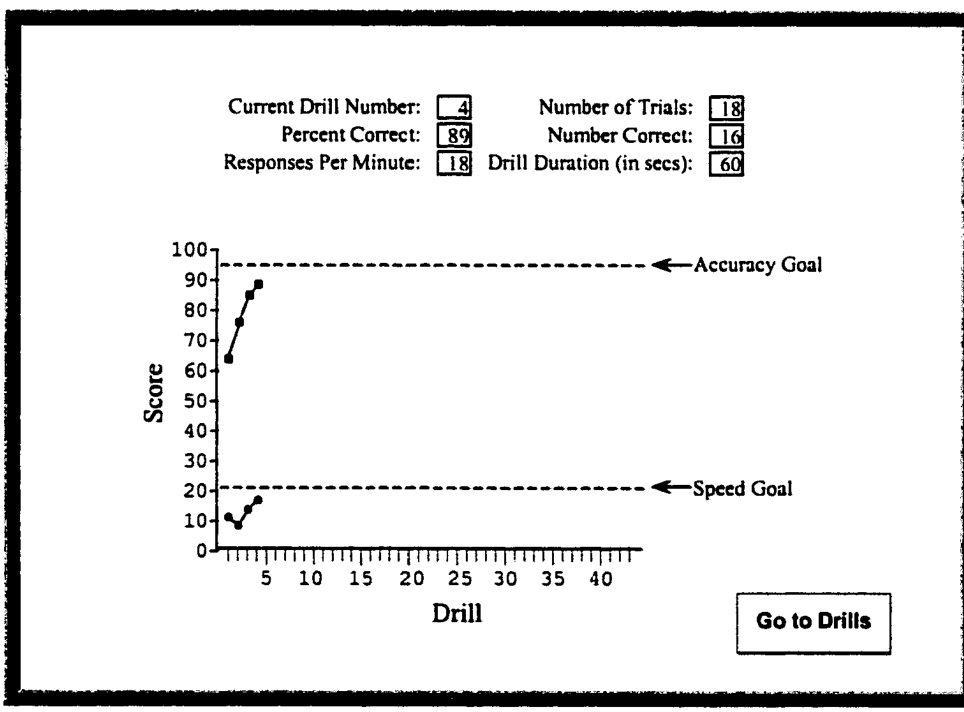


Figure 7. Appearance of Computer Screen With Graph Depicting Subject Performance During Prerequisite Training.

During topography-based prerequisite exercises, subjects were instructed to tap out Morse Code on the computer mouse, and again to strive toward 95% accuracy and 20 responses per minute (see Figure 8).

When a sample stimulus (a Morse Code symbol) appeared on the computer screen, the subject tapped out the code on the computer mouse, and dots and dashes corresponding to the taps appeared beneath the sample (see Figure 9).

The computer program treated a pause of .5-sec as an indication that the response was completed. (This .5-sec was not included in the response duration measurement.) Incorrect responses were followed by the brief appearance of the word "incorrect" on the computer screen, accompanied by a computer-generated “beep” sound, an illustration of the correct relationship, and a repeat of that trial. Correct responses were followed by the brief appearance of the word "correct," and a new trial with a different sample stimulus.

This training proceeded in 1-min drills, each followed by the appearance of a graph depicting performance relative to the training goal for that drill. Topography-based exercises ended after the fluency goal had been met, or when nine drills had been completed, whichever occurred first, and were followed by a short rest break. Two topography-based exercises were conducted in this manner. If the subject met fluency criteria in either of the topography-based exercises, she was invited to participate in the rest of the study.

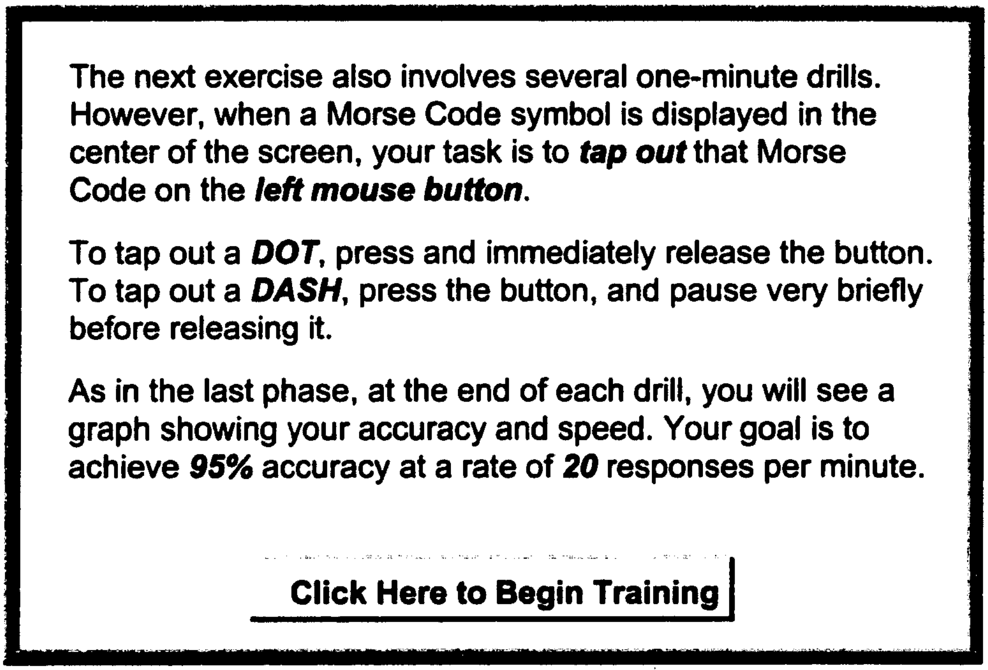


Figure 8. Appearance of Computer Screen Displaying Instructions for Topography-Based Prerequisite Skill Exercise.

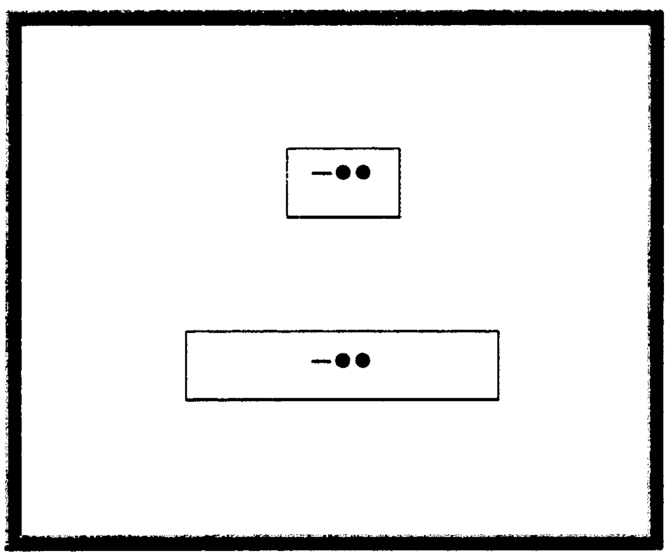


Figure 9. Computer Screen as It Appeared During Topography-Based Prerequisite Skill Exercises.

Experimental Sessions / Acquisition Phase

During each experimental session, subjects practiced four sets of relationships between Kanji symbols and Morse Code characters. Each session was divided into four subsessions, with rest breaks in between. Each subsession consisted of eight drills of 18 trials. A different treatment condition was in effect during each subsession, the order of which varied from one day to another.

During the first experimental session only, each subsession began with an introduction to the relationships to be learned and practiced in that subsession. A one-page legend was provided (see Appendix C). Subjects were encouraged to devise mnemonic methods for remembering the relationships, and the researcher demonstrated some mnemonic behavior. The subject was then asked to study the legend for five minutes.

At the end of each experimental session, subjects did a couple of reflexivity exercises (e.g., A=A). These were identical to the prerequisite skill training exercises, except they were limited to two drills for each response form. The computer captured the net S-R latency for each trial and a mean net latency was determined for each Morse Code character. Each of these mean net latencies represented an estimate of the time required to respond to a stimulus when conditional discrimination involvement was at a minimum. With this information, the researcher could estimate the duration of intervening events involved in conditional discriminations during that session's paired associate training.

Condition One: Selection-Based Training With Slow-Speed Practice Goal

At the beginning of each subsession of this type, an instruction field (see Figure 10) appeared on the screen informing the subject of the selection-based response form and the slow-speed performance goal (100% accuracy at rates under 15 responses per minute).

Select Code

In this exercise, the focus will be on ***practice***.

At the end of each drill, you will see a graph showing your accuracy and speed. Your goal is to maintain ***100% accuracy*** ***without exceeding 15 responses per minute***.

Work **slowly** and **carefully**—do not sacrifice accuracy for speed.

Click Here to Begin Training

Figure 10. Appearance of Computer Screen Displaying Instructions for Selection-Based Training with Slow-Speed Practice Goal (Condition One).

During drills, the subject saw a sample stimulus (a Kanji symbol) on the computer screen, and an array of six choice stimuli (all Morse Code characters) one of which corresponded to the sample (see Figure 11). She responded by clicking on one of the choice stimuli. Incorrect responses were followed by the brief appearance of the word "incorrect" on the computer screen, accompanied by a computer-generated "beep" sound, an illustration of the correct relationship, and a repeat of that trial. The brief appearance of the word "correct," and a new trial with different sample and choice stimuli followed correct responses. This training proceeded in drills of 18 trials each. At the end of each drill, the subject saw a graph depicting her performance relative to the training goal (as in Figure 7). The subsession ended when the subject had completed 8 drills of 18 trials.

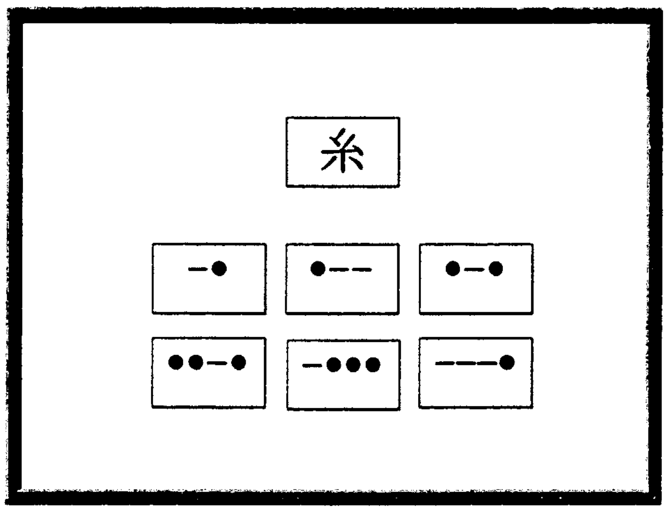


Figure 11. Appearance of Computer Screen With Sample and Choice Stimuli During Selection-Based Training With Slow-Speed Practice Goal (Condition One).

Condition Two: Topography-Based Training With Slow-Speed Practice Goal

At the beginning of each subsession of this type, an instruction field (see Figure 12) appeared on the screen informing the subject of the *topography*-based response form and the same *slow*-speed performance goal as in Condition One.

During drills, the subject saw a sample stimulus (a Kanji symbol) on the computer screen, and responded by tapping out the corresponding Morse Code character on the computer mouse (see Figure 13). In all other respects, subsessions proceeded as in Condition One, above.

Condition Three: Selection-Based Training With Fluency Goal

At the beginning of each subsession of this type, an instruction field (see Figure 14) appeared on the screen informing the subject of the *selection*-based response form and the *high*-speed performance goal (100% accuracy at rates as fast as possible). In every other respect, subsessions proceeded as in Condition One, above (see Figure 11).

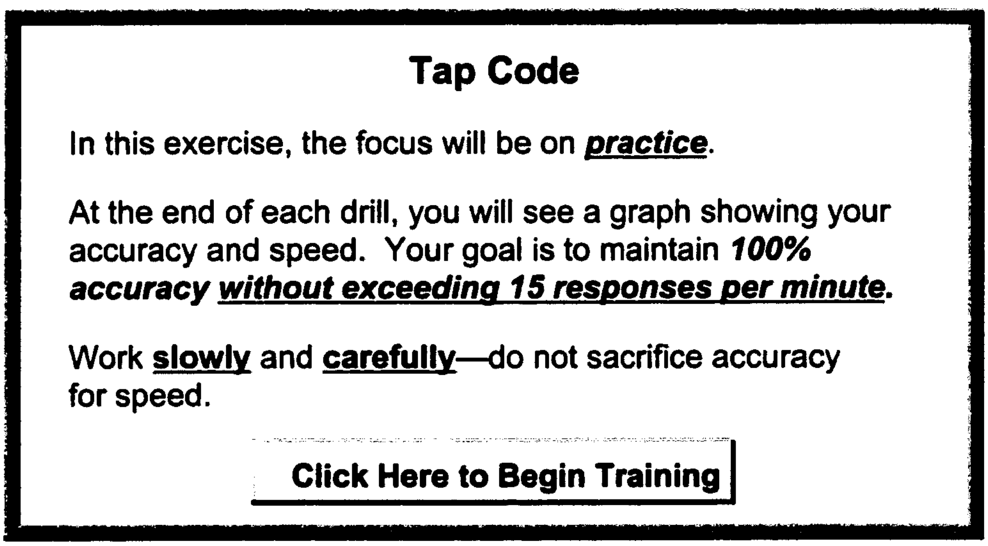


Figure 12. Appearance of Computer Screen Displaying Instructions for Topography-Based Training With Slow-Speed Practice Goal (Condition Two).

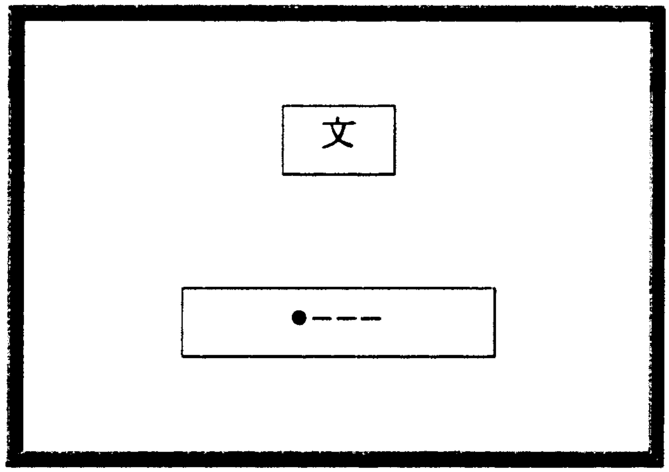


Figure 13. Appearance of Computer Screen During Topography-Based Training With Slow-Speed Practice Goal (Condition Two).

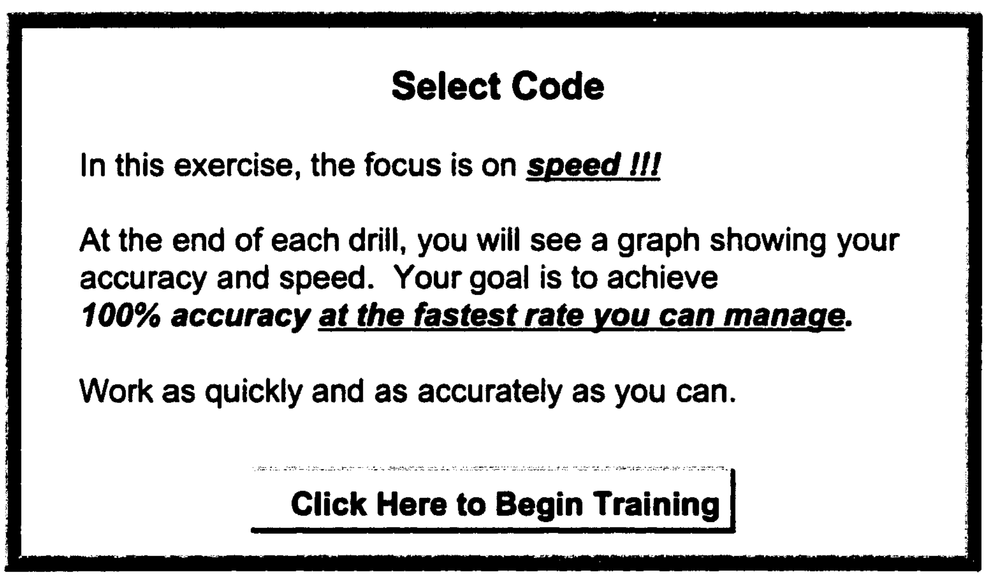


Figure 14. Appearance of Computer Screen Displaying Instructions for Selection-Based Training With Fluency Goal (Condition Three).

Condition Four: Topography-Based Training With Fluency Goal

At the beginning of each subsession of this type, an instruction field (see Figure 15) appeared on the screen informing the subject of the *topography*-based response form and the same *high*-speed performance goal as in Condition Three. In all other ways, subsessions proceeded as in Condition Two (see Figure 13).

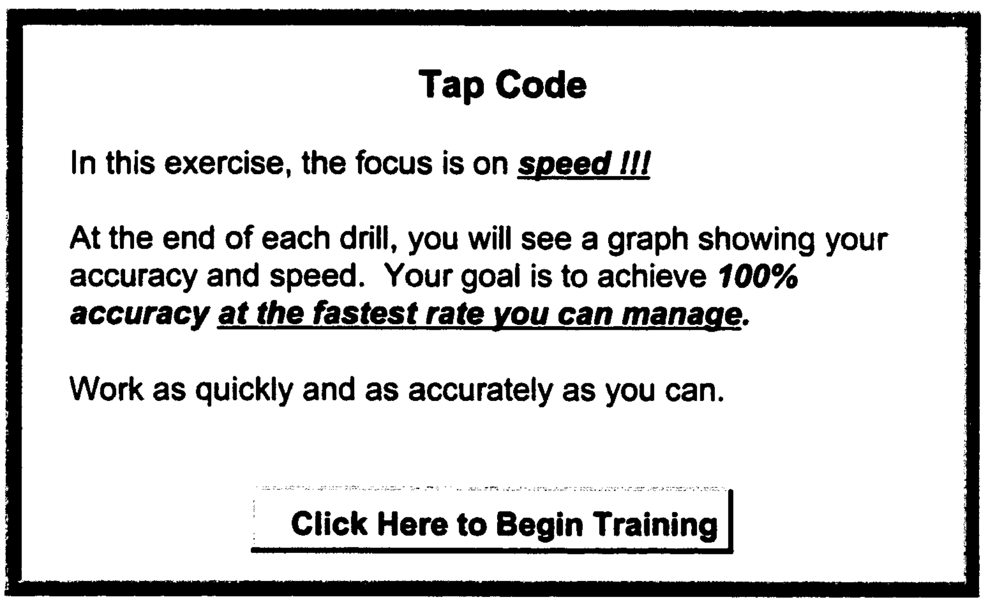


Figure 15. Appearance of Computer Screen Displaying Instructions for Topography-Based Training With Fluency Goal (Condition Four).

Testing Sessions

A series of nine tests were administered during each testing session, with a short break (a few minutes) after each test. All 24 relationships were tested in each test, regardless of the condition under which respective relationships were taught.

Tests 1 through 7 each comprised 48 items, two for each Kanji/Morse relationship. Tests 8 and 9 each comprised 24 items, one for each of the relationships.

Test 1: Selection-Based Retention

On each item in Test 1, subjects saw a Kanji symbol and were instructed to click on the corresponding Morse Code character in an array of choices.

Test 2: Topography-Based Retention

On each item in Test 2, subjects saw a Kanji symbol and were instructed to tap out on the corresponding Morse Code character on the computer mouse.

Test 3: Symmetric Equivalence

On each item in Test 3, subjects saw a Morse Code character and were instructed to click on the corresponding Kanji symbol in an array of choices.

Test 4: Transitive Equivalence 1

On each item in Test 4, subjects saw a Morse Code character flashed out on the computer screen, and were instructed to click on the corresponding Kanji symbol in an array of choices.

Test 5: Transitive Equivalence 2

On each item in Test 5, subjects heard a Morse Code character played through the computer speakers, and were instructed to click on the corresponding Kanji Code character in an array of choices.

Test 6: Response Form Evocative Strength

On each item in Test 6, a Kanji symbol appeared on the screen along with an array of choice stimuli and a screen object onto which subjects could tap out Morse Code characters. Subjects were instructed before the test began that for each trial, they would have their choice of response forms: they could use the same response form throughout the test, or they could select and click some answers, and tap out others. Both response forms were possible throughout the test.

Test 7: Endurance

This test was conducted in exactly the same way as Test 6, but under distracting stimulus conditions. The background of computer screen was set to a complex pattern of dots and dashes, and the sound of Morse Code dots and dashes were played continually throughout the test.

Test 8: Transitive Equivalence 3

On each item in Test 8, a Kanji symbol appeared on the screen and subjects were instructed to draw the corresponding Morse Code character on an answer sheet.

Test 9: Transitive Equivalence 4

On each item in Test 9, a Morse Code character appeared on the screen and subjects were instructed to draw the corresponding Kanji symbol on an answer sheet.

CHAPTER V

RESULTS

Training

Calculation of Net Latencies

Each training session ended with two reflexivity exercises, during which a reflexive latency corresponding to each of the 24 Morse Code symbols was recorded. This latency was a measure of the time between the onset of a choice stimulus and the start of a response. Because the sample stimulus was identical to the correct choice stimulus in each of these trials, the time between the stimulus onset and the start of the response would involve a minimal amount of conditional discrimination.

A net latency was calculated for each trial during training. This latency was calculated by measuring the time between the onset of a choice stimulus and the beginning of a response, and then subtracting the mean reflexive latency for the corresponding Morse Code symbol for that session. For example, if the mean reflexive latency for the •–– : •–– relationship during the first training session was 50/1000 sec, then that amount was subtracted from a measure of the S-R interval for each trial involving the ⁂ : •–– relationship during session one. The resulting net latencies are estimates of the duration of any activity specifically related to paired associate responding.

Performance During Training Sessions

Figure 16 summarizes speed and accuracy data from all training sessions for each subject. When the slow-speed goal was in effect, latencies for all subjects were longer under selection-based conditions than under topography-based. When the fast-speed goal was in effect, three subjects responded fastest under topography-based conditions, whereas two responded fastest under selection-based conditions.

All subjects performed more accurately (usually at or near 100%) under selection-based conditions than under topography-based conditions (ranging from 78% to 99%) across all three training sessions, irrespective of the speed goal. When the slow-speed goal was in effect, there was a correlation of +.16 between mean latency and accuracy under selection-based conditions, and +.11 topography-based conditions, respectively. Thus, responses following long latencies produced accurate responses slightly more often than did those following short latencies under selection-based conditions. When the fast-speed goal was in effect, there was a negative correlation between mean latency and accuracy: -.48 for selection-based responding, and -.45 for topography-based. Under fast-speed conditions, responses following short latencies usually produced accurate responses, whereas those following long latencies usually produced incorrect responses.

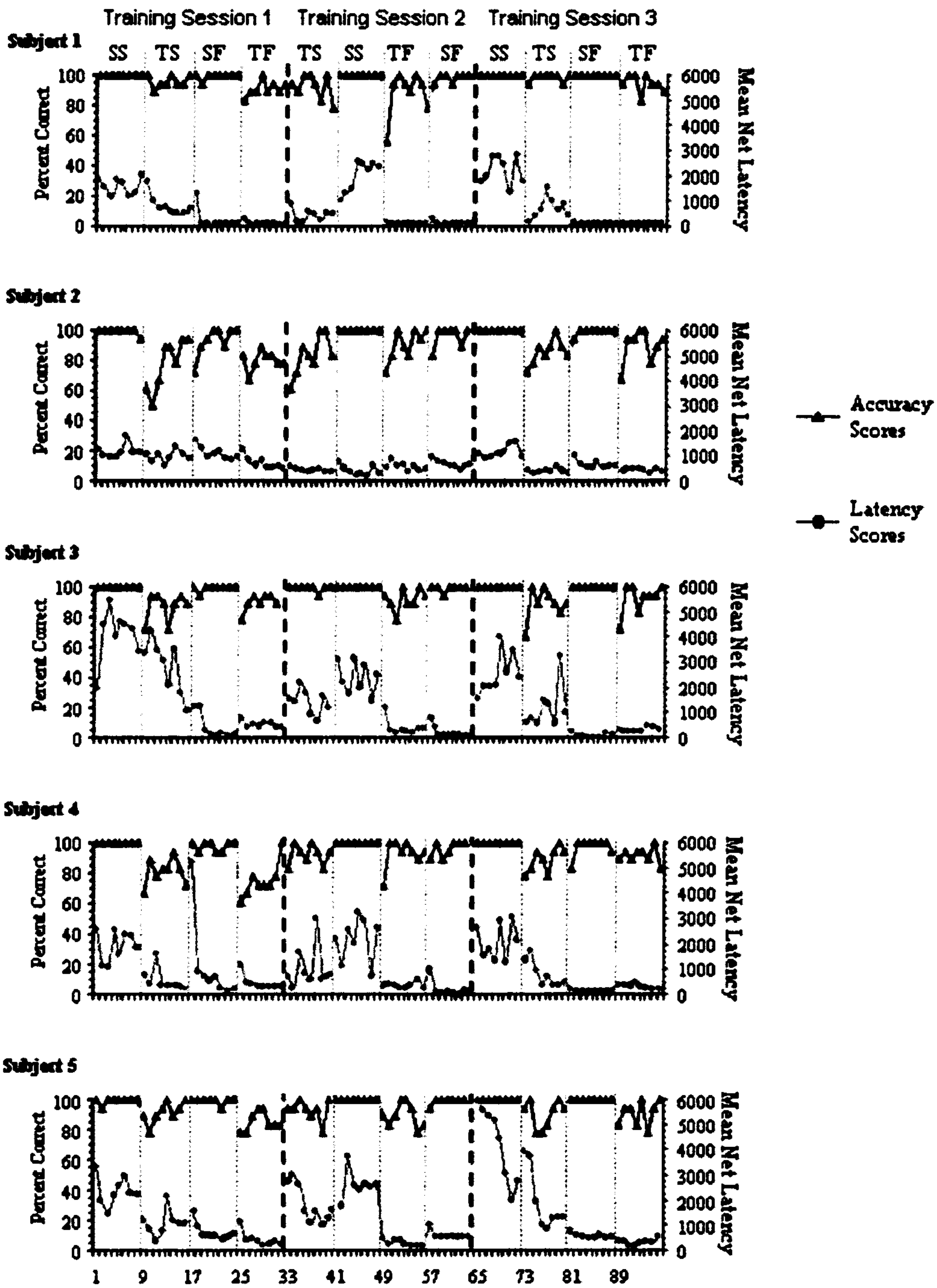


Figure 16. Accuracy and Mean Net Latency for All Training Drills in Chronological Order.

Note: All latencies are in 1/1000ths of a sec. SS=Selection Slow, TS=Topography Slow, SF=Selection Fast, TF=Topography-Fast.

Integrity of the Independent Variable

During training, subjects were encouraged to strive toward stated goals in terms of speed and accuracy, although completion of a condition was not contingent upon meeting the performance goals. When displayed chronologically as in Figure 17, the emphasis on striving toward fast or slow speed goals appears to have resulted in distinctively different latencies between slow-speed conditions and fast-speed conditions.

However, when all relationships for all subjects are sorted first according to response form trained, then according to speed goal, and finally by actual mean latency during training, as depicted Figure 18, the distinction is not as clear. When displayed this way it can be seen that some subjects sometimes responded as fast under slow goal conditions as under the fast ones. To reconcile this, relationships were re-categorized according to actual mean net latency during training. When resorted, first according to response form trained and then according to actual mean net latency during training, the same data fell neatly into four groups as illustrated in Figure 19. When the effects of speed upon dependent variables are discussed later in this paper, the effects of actual mean net latencies during training, rather than speed goal, will be addressed.

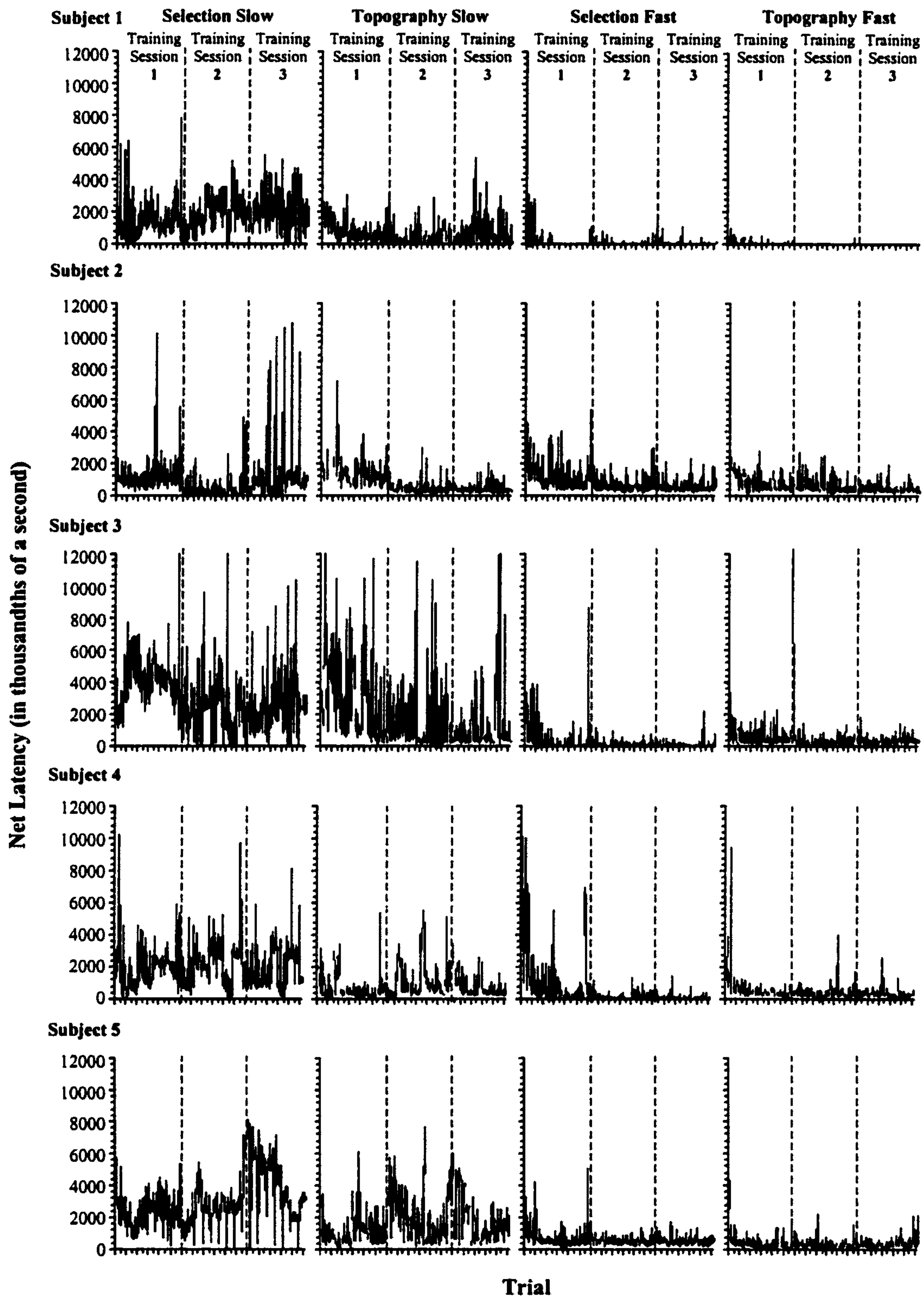


Figure 17. Net Latencies for all Training Trials, in Chronological Order Within Treatment Condition for All Subjects.

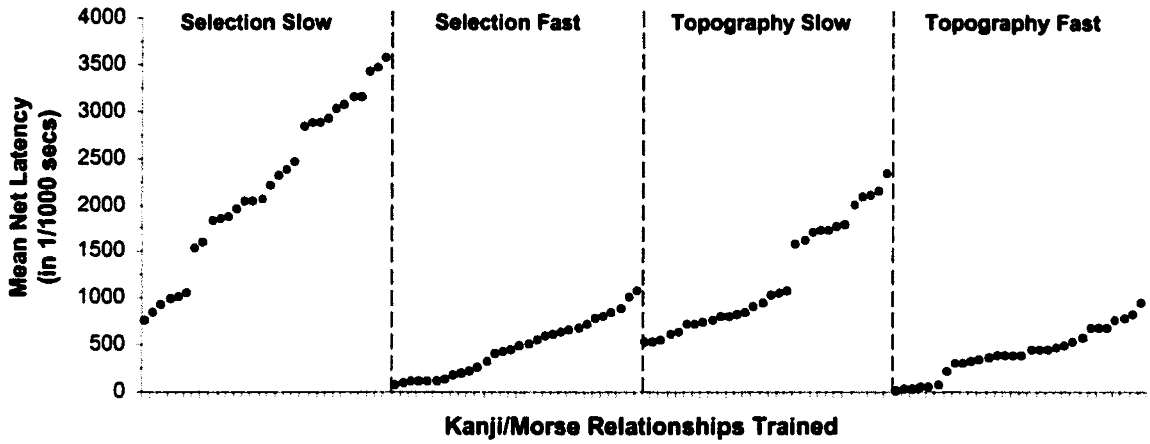


Figure 18. Mean Net Latencies for All Relationships and All Subjects, Sorted First by Training Condition and Then by Length of Latency.

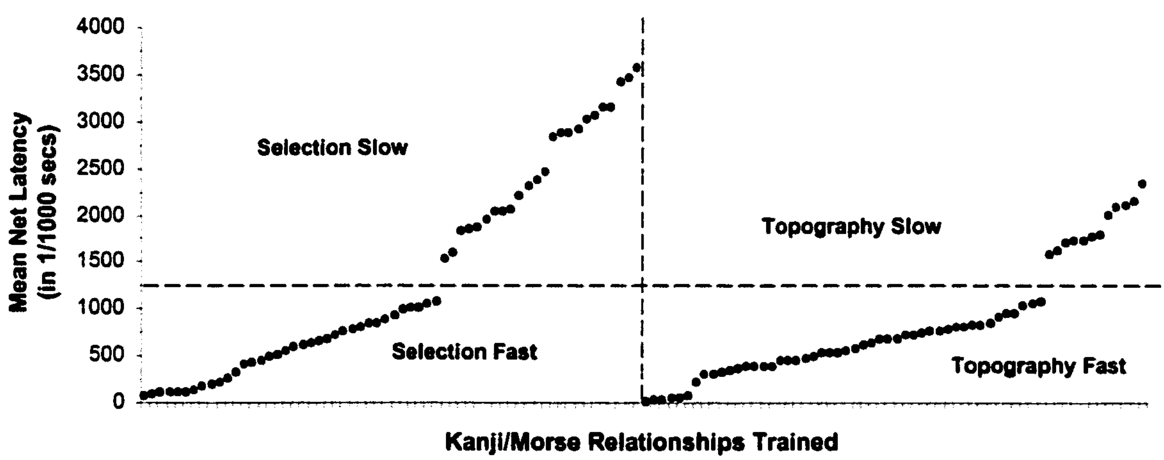


Figure 19. Mean Net Latencies for All Relationships and All Subjects, Sorted First by Response Form Trained and Then by Length of Latency.

Effects of Response Form on Net Latencies

The data in Figure 17 do not show a steady and persistent decrease in net latencies under topography-based conditions as compared to selection-based conditions as was seen in McCarty (1997). No difference of this kind is evident when the data are plotted chronologically or when grouped by individual Kanji/Morse relationships (see Appendix D, Figures D1 through D5 for data displayed in these ways).

Testing

Interobserver Agreement

Seven of the nine tests were graded by the computer program as either correct or incorrect. Tests 8 and 9 required handwritten answers and were graded independently by two observers. On each item on Test 8, a Kanji symbol was displayed on the computer screen and subjects wrote the corresponding Morse Code character onto an answer sheet. Later, observers compared the handwritten answers to an answer key and scored each item as either correct or incorrect. Scores assigned by the two observers for Test 8 were identical; that is, there was perfect agreement between the two observers.

On each item on Test 9, subjects saw a Morse Code character on the computer screen and drew the corresponding Kanji symbol on the answer sheet. Observers later graded the test items on a scale of zero to four, with zero being completely wrong (or left unanswered), and four being perfect or very close to it (all features of the symbol were present, although some may not have been positioned perfectly). A measure of agreement between observers was determined by performing a Pearson product moment correlation on the two sets of scores for each administration of the test. The resulting coefficients, which ranged from .894 to 1.000, are displayed in Table 2.

Table 2

Correlation Coefficients Between Scores by Two Observers for Test 9

Subject	Test Session 1	Test Session 2	Test Session 3
1	0.967	0.959	0.895
2	0.912	0.909	0.900
3	0.903	0.932	0.940
4	1.000	0.937	0.899
5	0.965	0.909	0.894

Effects of Response Form

Combined scores for all subjects, tests, and test sessions are displayed according to response form in Figure 20. When combined in this way, subjects scored 10.9% higher for relationships learned under topography-based conditions than for those learned under selection-based conditions.

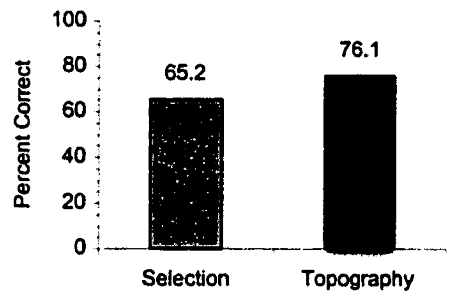


Figure 20. Combined Scores for All Subjects, Tests, and Test Sessions, by Response Form Trained.

When those data are further sorted according to individual subjects as in Figure 21, it can be seen that all subjects performed best under topography-based conditions. However, the extent to which topography-based relationships exceeded selection-based varied a great deal—from 1.4% for Subject 3 to 31% for Subject 1.

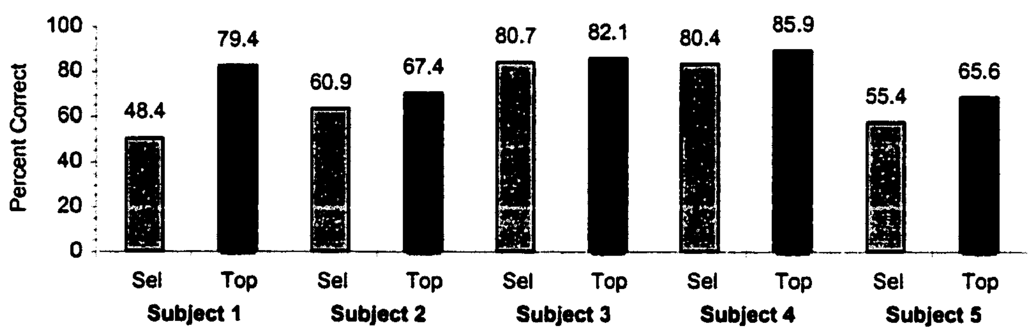


Figure 21. Combined Scores for All Tests and Test Sessions, Sorted by Subject and by Response Form Trained.

Effects of Speed

Combined scores for all subjects, tests, and test sessions sorted according to actual latencies during training are shown in Figure 22. When combined in this way, subjects scored 2.8% higher for relationships learned under slow (long latency) conditions than for those learned under fast conditions.

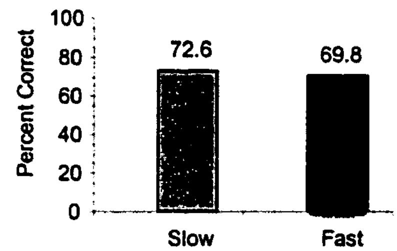


Figure 22. Combined Scores for All Subjects, Tests, and Test Sessions, Sorted According to Speed During Training.

When those data are further sorted according to individual subjects as in Figure 23, it can be seen that three of the five subjects performed best under slow conditions, and one performed best under fast conditions. (Subject 2 did not comply with instructions concerning the speed goal. As a result, all of her relationships were re-categorized as having been learned under fast speed conditions.)

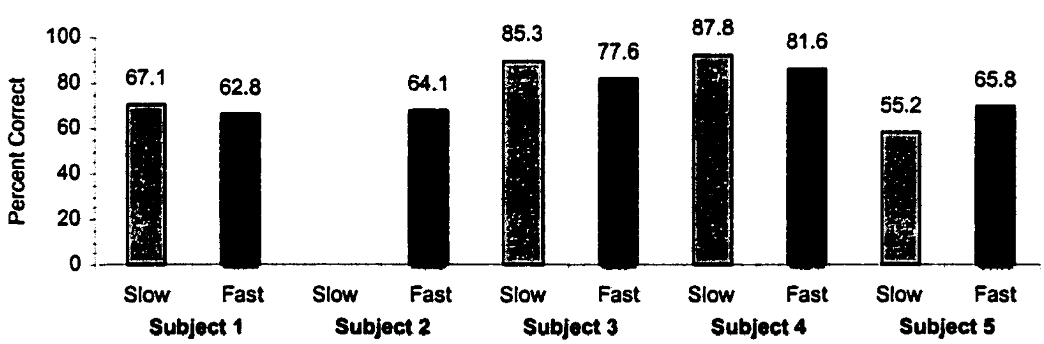


Figure 23. Combined Scores for All Tests and Test Sessions, Sorted by Subject and by Speed During Training.

With respect to the three subjects who performed best on relationships learned under slow conditions, the extent to which slow-speed scores exceeded high-speed scores ranged from 4.3% for Subject 1 to 7.7% for Subject 3. Subject 5 performed 10.6% better on relationships learned under fast conditions than on those learned under slow conditions.

Interaction Between Form and Speed

Combined scores for all subjects, tests, and test sessions are displayed according to response form trained and speed during training in Figure 24. With respect to relationships learned under topography-based conditions, subjects scored the same regardless of the how fast they responded during training. For relationships learned under selection-based conditions, subjects' combined accuracy scores were 13.7% higher for relationships learned under slow conditions during training than for those learned under fast conditions.

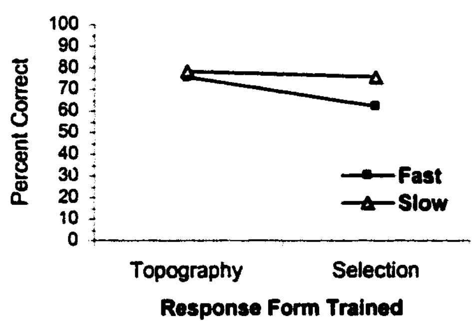


Figure 24. Combined Scores for All Subjects, Tests, and Test Sessions, by Response Form Trained and Speed During Training.

These same data, separated by subject, are displayed in Figure 25. Because Subjects 1, 2, and 4 did not comply with speed goal instructions during training, no interaction between response form trained and speed during training can be evaluated. With respect to Subject 3, there appears to be no interaction between the two independent variables. There does appear to be a clear interaction in Subject 5's scores. For relationships learned under topography-based conditions, shorter latencies are associated with higher test scores. For relationships learned under selection-based conditions, shorter latencies are associated with lower test scores.

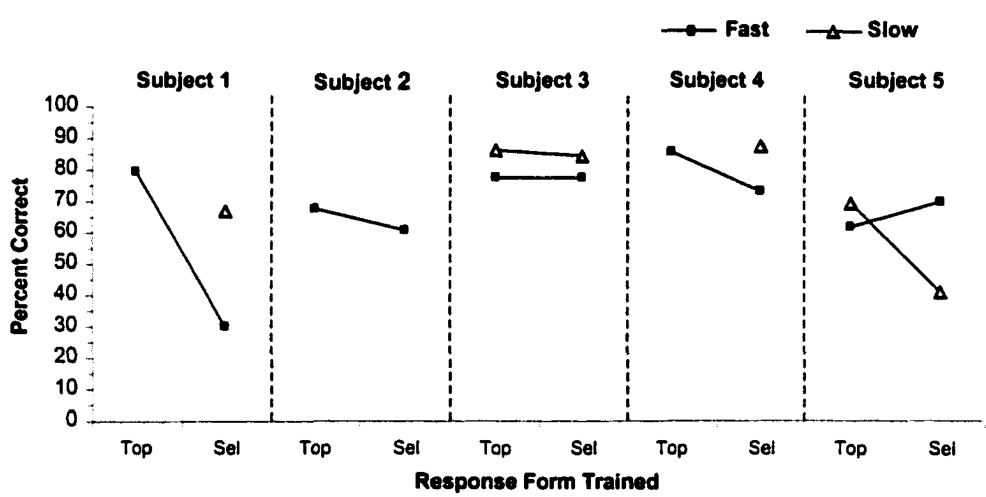


Figure 25. Combined Scores for All Tests and Test Sessions, by Subject, Response Form Trained, and Speed During Training.

Effects of the Passage of Time

Effects of Response Form Trained on Retention Over Time

Test 1 is a simple retention test for relationships learned under selection-based conditions. On this test, subjects saw a Kanji Symbol and selected the corresponding Morse Code character from an array, just as during training under selection-based conditions. Test 2 is a simple retention test for relationships learned under topography-based conditions. On this test, subjects saw a Kanji Symbol and tapped the corresponding Morse Code character, just as during training under topography-based conditions. To enable a meaningful evaluation of retention between the last training session and subsequent testing sessions, Test 1 scores for relationships learned under selection-based conditions and Test 2 scores for relationships learned under topography-based conditions are displayed along with accuracy scores from the last training session in Figure 26.

In the last training session, accuracy scores for relationships learned under selection-based conditions are 8.6% higher than those for relationships learned under topography-based conditions. In each of the subsequent testing sessions, however, scores for relationships learned under topography-based conditions exceed those for relationships learned under selection-based conditions. All of the deterioration occurred between the last training session and the first test session, a period of approximately 24 hours. Although seven days passed between test sessions, no further deterioration appears to have occurred for relationships learned under selection-based conditions. On the contrary, scores for these relationships increased by 6.7% between the first and last test session. Scores for relationships learned under topography-based conditions decreased by 5% over the three test sessions. The overall deterioration between the last training session and the last testing session for relationships learned under selection-based conditions is 28%. The overall deterioration for relationships learned under topography-based conditions is 13.6%.

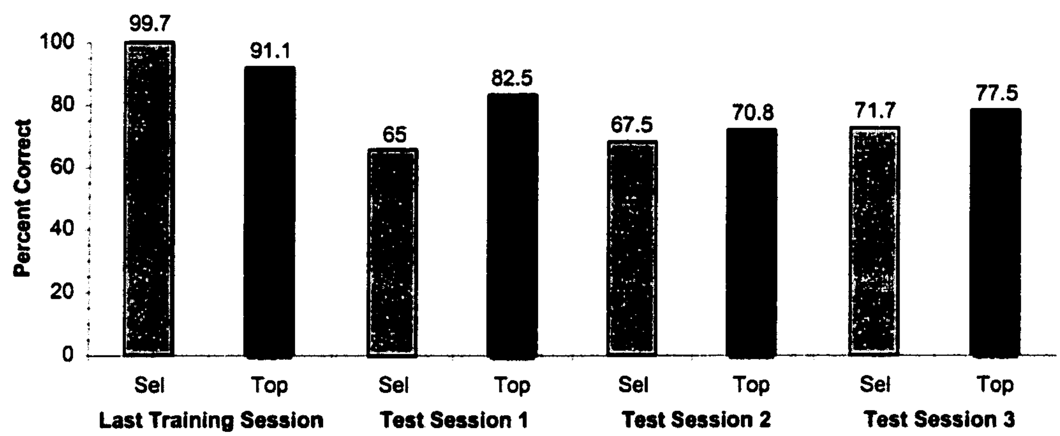


Figure 26. Combined Accuracy Scores for All Subjects During Last Training Session and on Simple Retention Items From Tests 1 and 2, Sorted by Response Form Trained and by Test Session.

Effects of Speed During Training, Over Time

The same combined accuracy scores, now sorted according to speed during training, are displayed in Figure 27. As discussed above, most of the deterioration occurred in the day between the last training session and the first testing session.

Scores for relationships with slow latencies during training decreased by 21.5% over that time. Scores for relationships with fast latencies during training decreased by 15.9%. Little change associated with speed during training appears to have occurred over the three testing sessions.

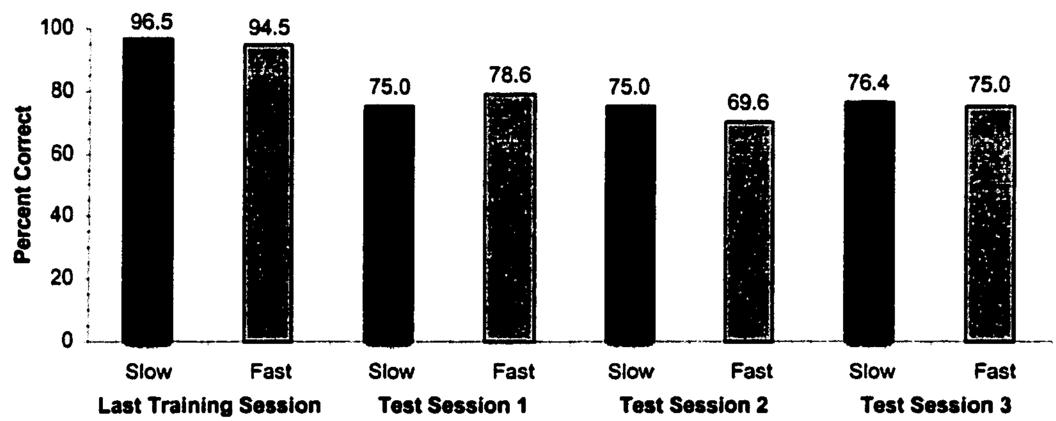


Figure 27. Combined Accuracy Scores for All Subjects During Last Training Session and on Simple Retention Items From Tests 1 and 2, Sorted by Speed During Training and by Test Session.

Effect of Interactions Between Independent Variables on Retention Over Time

Figure 28 depicts the level of interaction during the last training session and in the subsequent training sessions. No apparent interaction occurred during the last training session or during the first test session. By the second test session, there begins to be some interaction favoring topography-based fast relationships and selection-based slow ones. By test session 3, there is a clear interaction between form and speed, such that fast training latencies are associated with better retention than slow ones for relationships learned under topography-based conditions. Just the opposite is true for relationships learned under selection-based conditions.

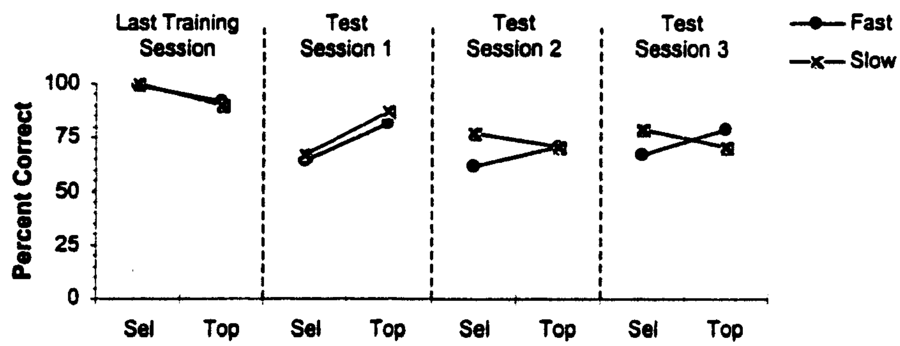


Figure 28. Combined Accuracy Scores for All Subjects During Last Training Session and on Simple Retention Items from Tests 1 and 2, Sorted by Test Session, Response Form Trained, and Speed During Training.

Performance on Individual Tests

Test performance for all subjects and test sessions are summarized in Figures 29 through 37. Although nine different tests were used to measure a variety of dependent variables, there was little variation in performance within subjects. That is, when a particular response form, speed during training, or an interaction of the two variables influenced a subject's test performance, it tended to do so in the same manner on all tests and during all test sessions. Other than the deterioration in performance that occurred between the last training session and the first test session, little or no decrease in scores occurred over the three test sessions for any of the subjects, regardless of training condition.

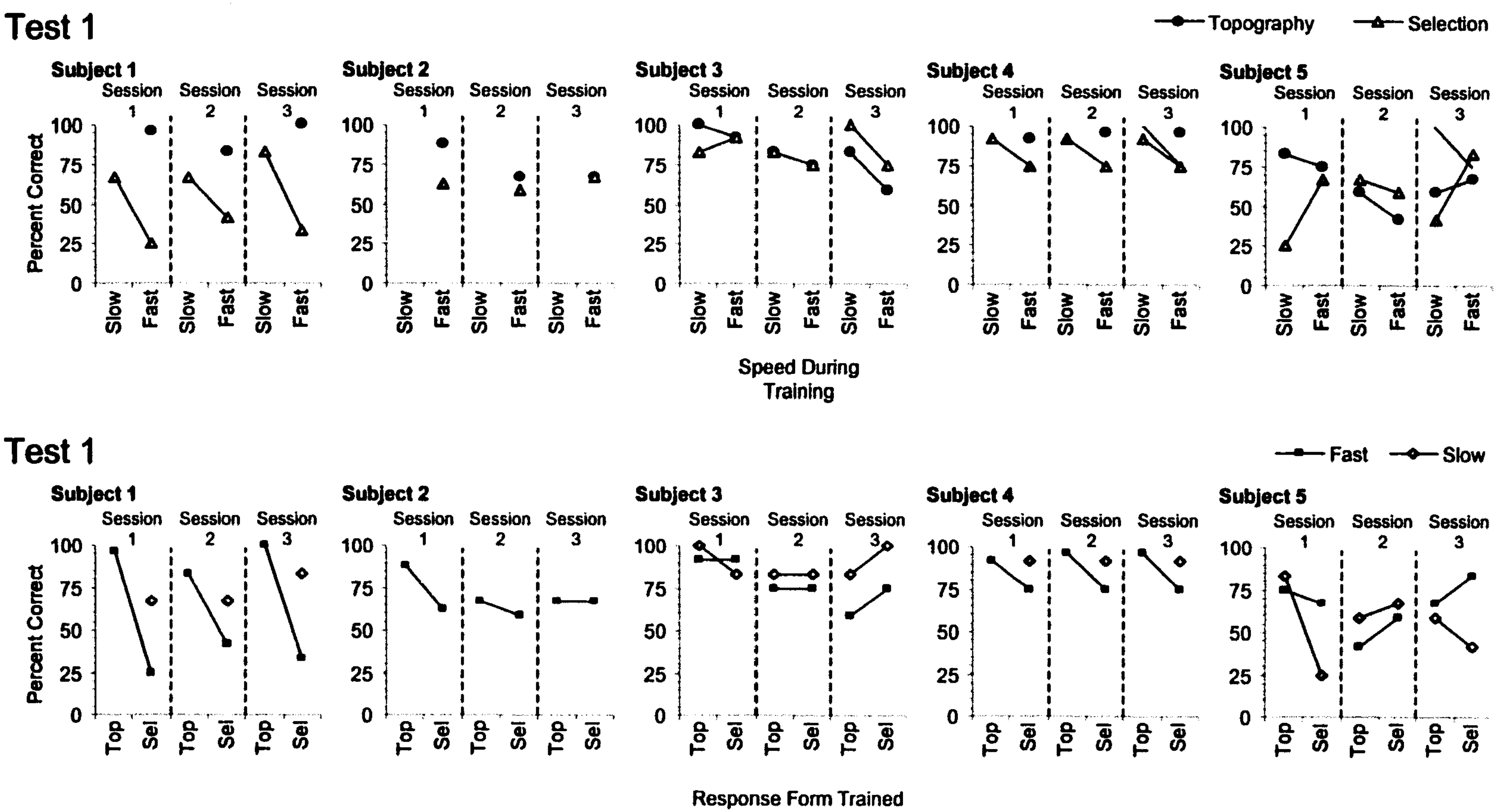
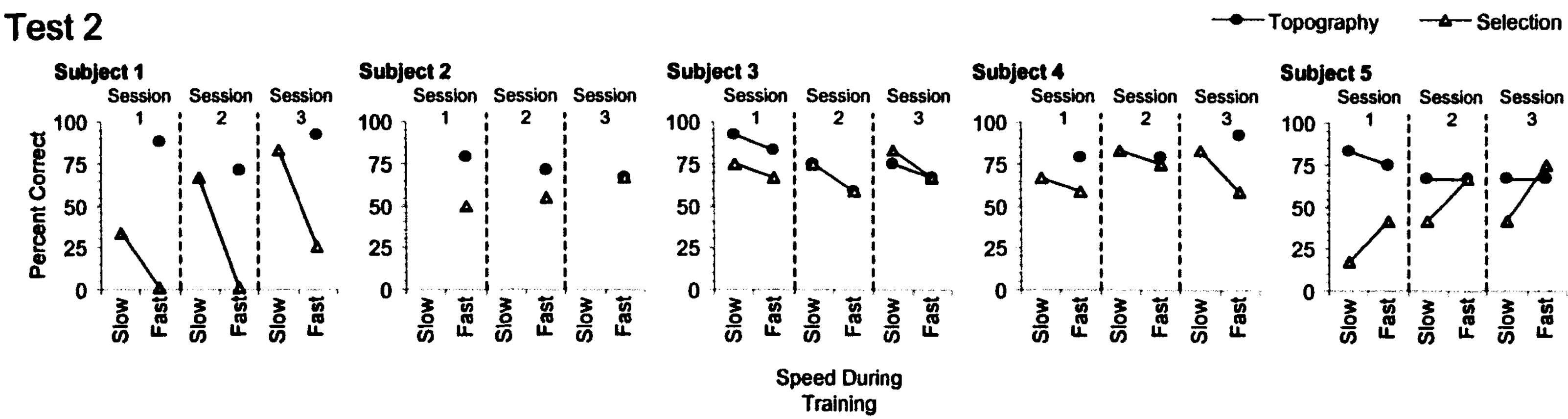


Figure 29. Test Performance for All Subjects and Sessions on Test 1.

Note: On each item in Test 1, subjects saw a Kanji symbol and were instructed to click on the corresponding Morse Code character in an array of choices.



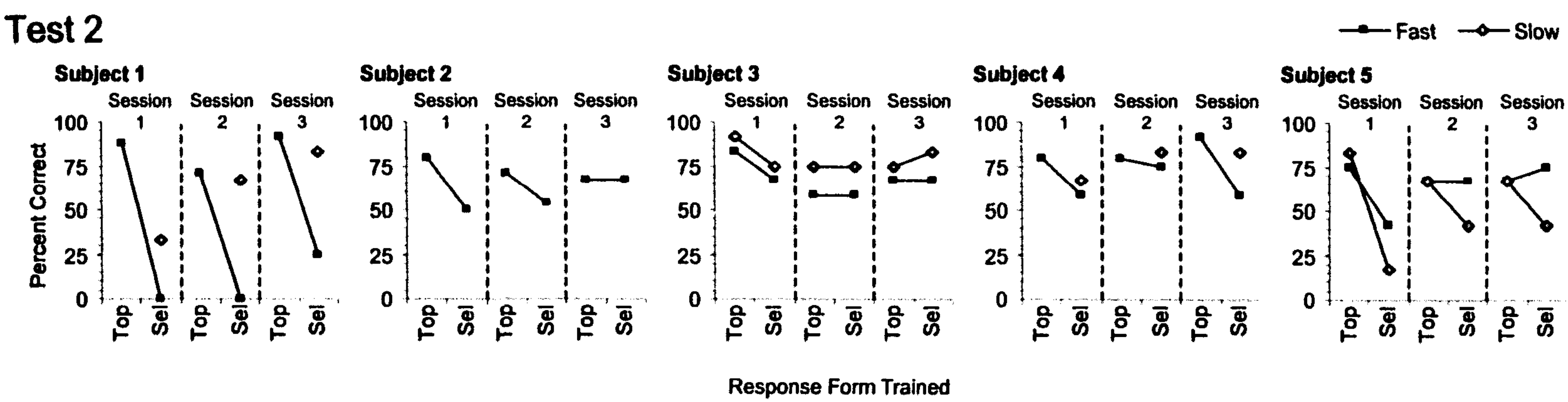


Figure 30. Test Performance for All Subjects and Sessions on Test 2.

Note: On each item in Test 2, subjects saw a Kanji symbol and were instructed to tap out on the corresponding Morse Code character on the computer mouse.

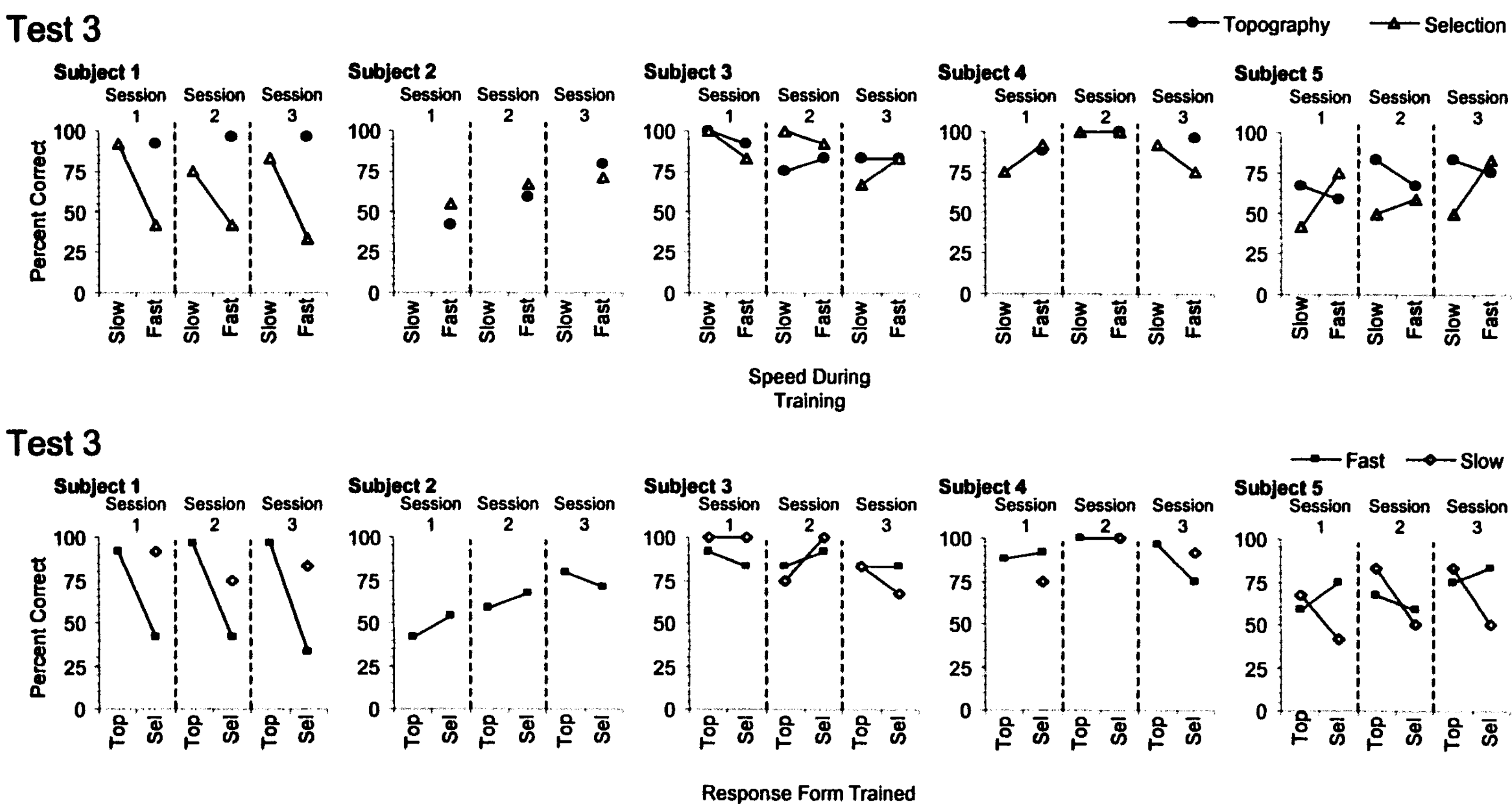


Figure 31. Test Performance for All Subjects and Sessions on Test 3.

Note: On each item in Test 3, subjects saw a Morse Code character and were instructed to click on the corresponding Kanji symbol in an array of choices.

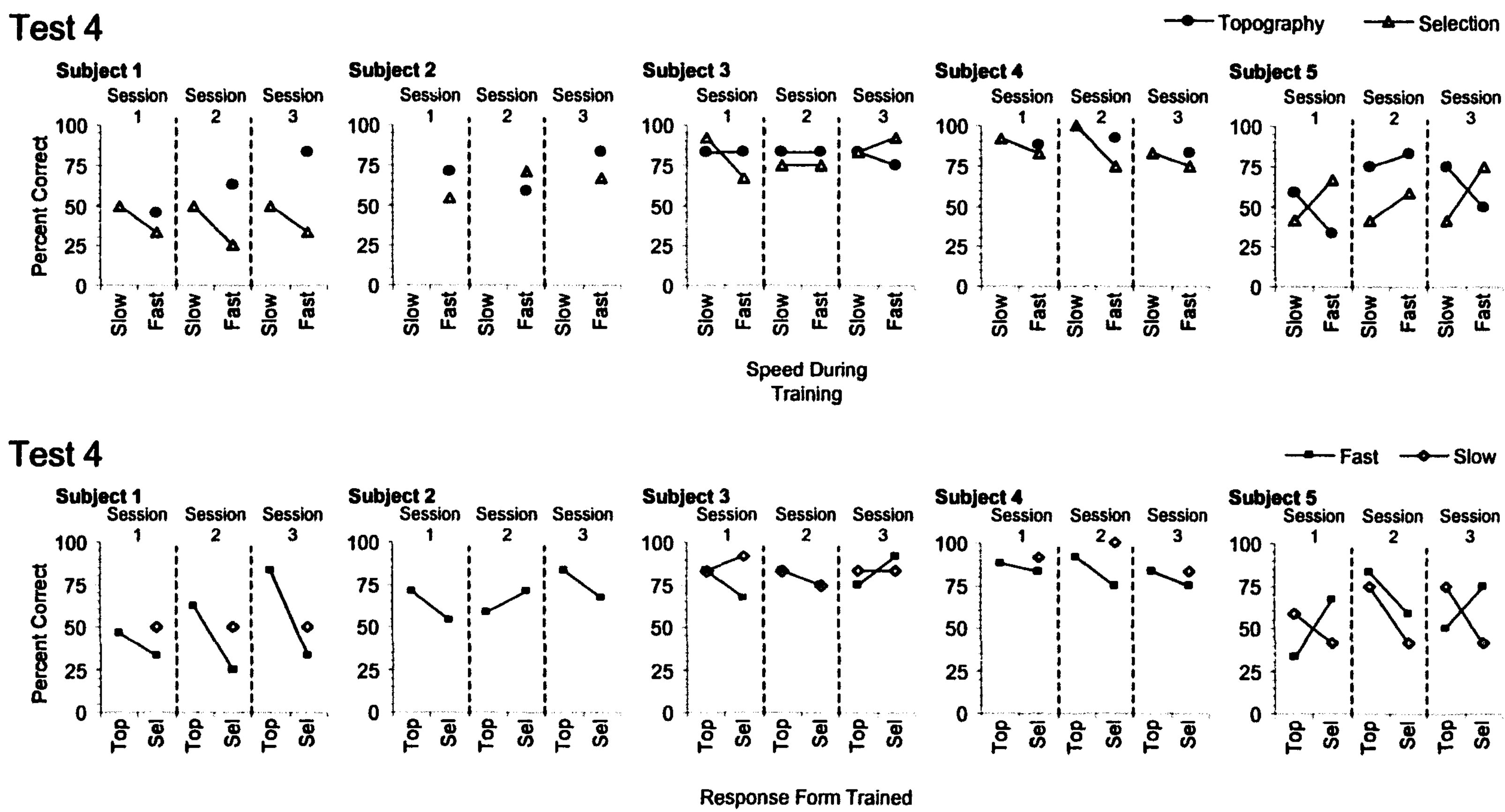


Figure 32. Test Performance for All Subjects and Sessions on Test 4.

Note: On each item in Test 4, subjects saw a Morse Code character flashed out on the computer screen, and were instructed to click on the corresponding Kanji symbol in an array of choices.

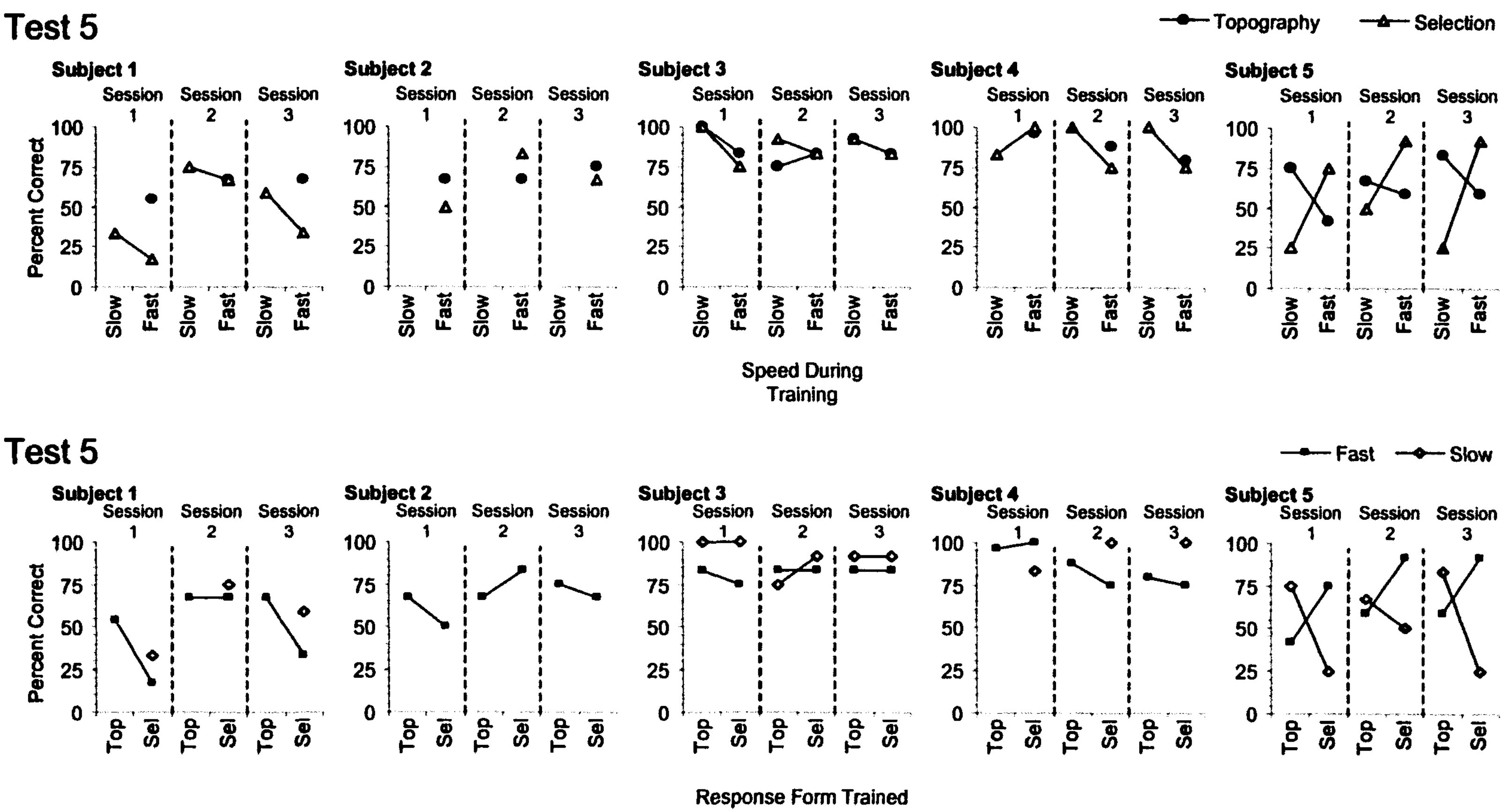


Figure 33. Test Performance for All Subjects and Sessions on Test 5.

Note: On each item in Test 5, subjects heard a Morse Code character played through the computer speakers, and were instructed to click on the corresponding Kanji symbol in an array of choices.

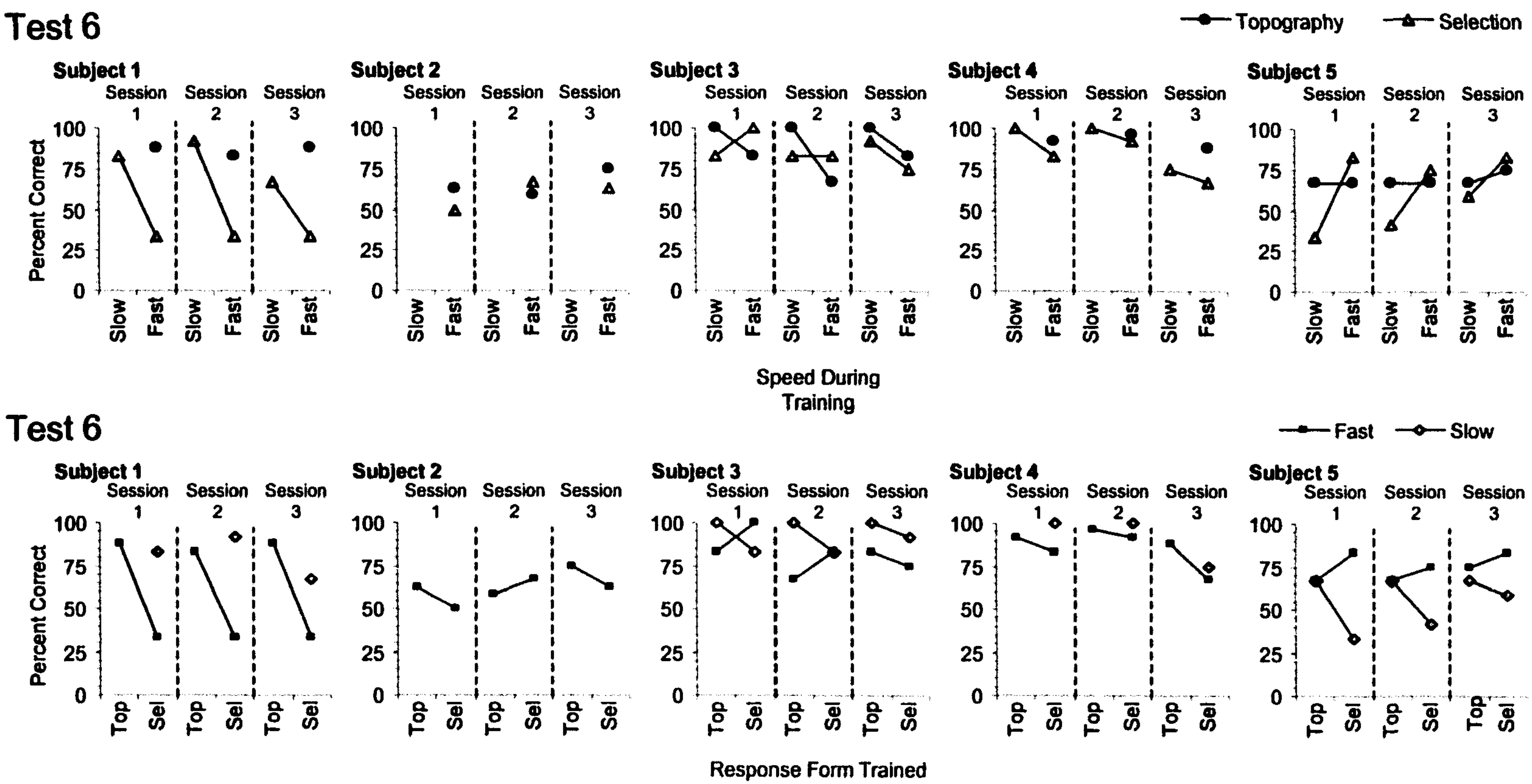


Figure 34. Test Performance for All Subjects and Sessions on Test 6.

Note: On each item in Test 6, a Kanji symbol appeared on the screen along with an array of choice stimuli and a screen object onto which subjects could tap out Morse Code characters. Subjects could choose to use the same response form throughout the test, or to select and click some answers, and tap out others.

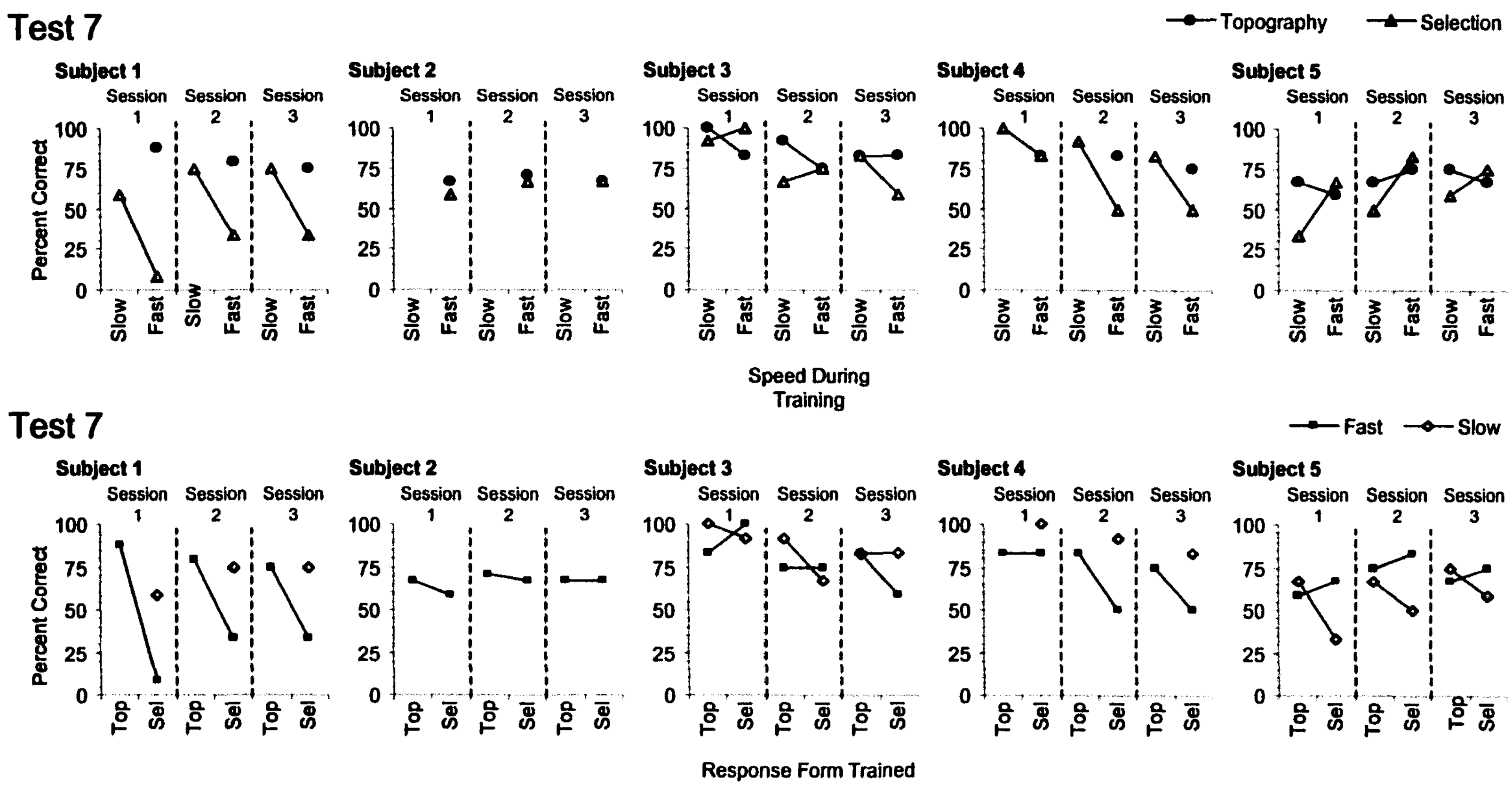


Figure 35. Test Performance for All Subjects and Sessions on Test 7.

Note: On each item in Test 7, a Kanji symbol appeared on the screen along with an array of choice stimuli and a screen object onto subjects could tap out Morse Code characters. As in the previous test, subjects could choose either type of response form on each item. The background was set to a complex pattern of dots and dashes, and the sound of Morse Code dots and dashes were played continually throughout the test.

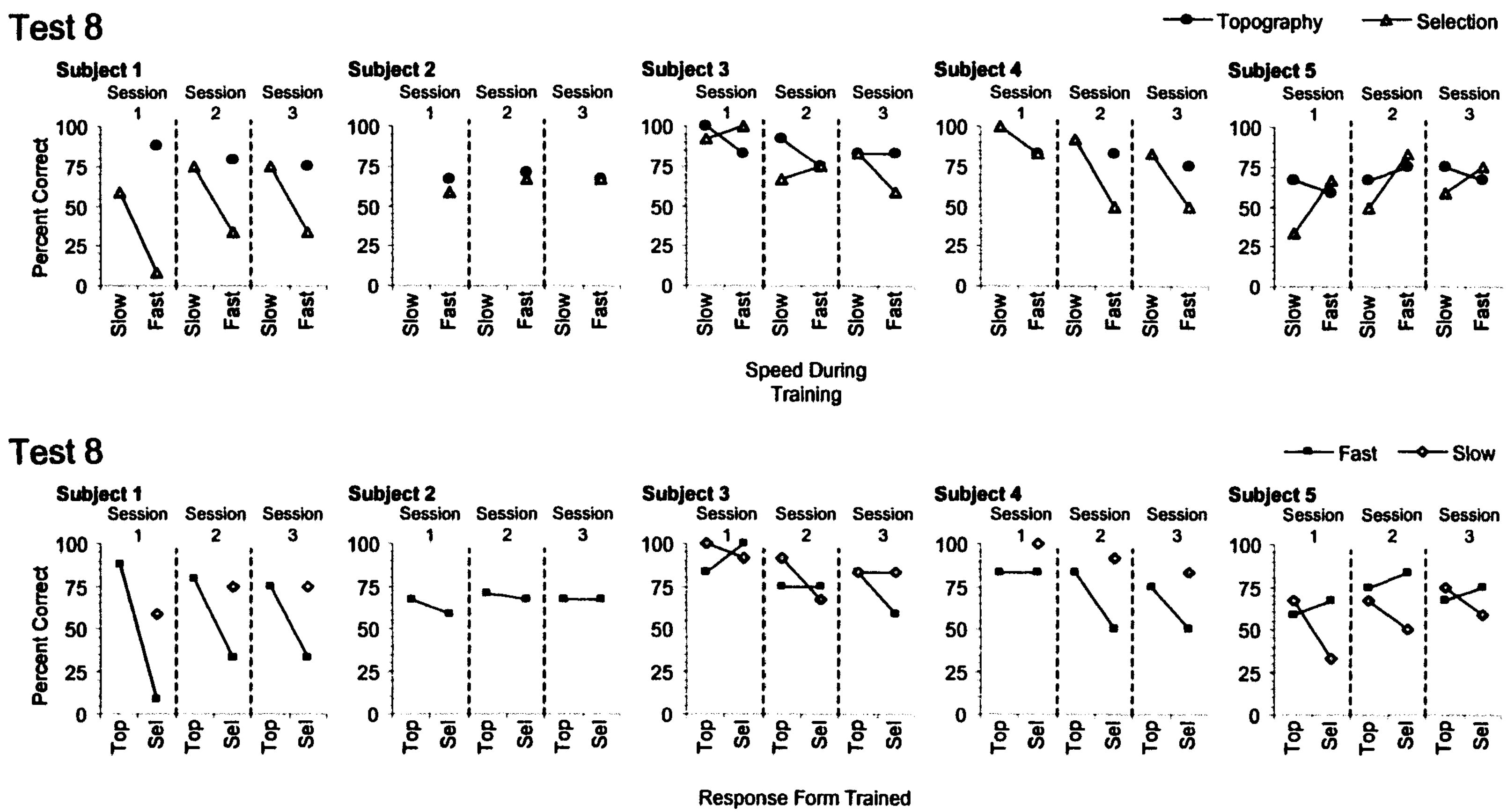


Figure 36. Test Performance for All Subjects and Sessions on Test 8.

Note: On each item in Test 8, subjects saw a Kanji symbol on the screen, and were instructed to draw the corresponding Morse Code character on an answer sheet.

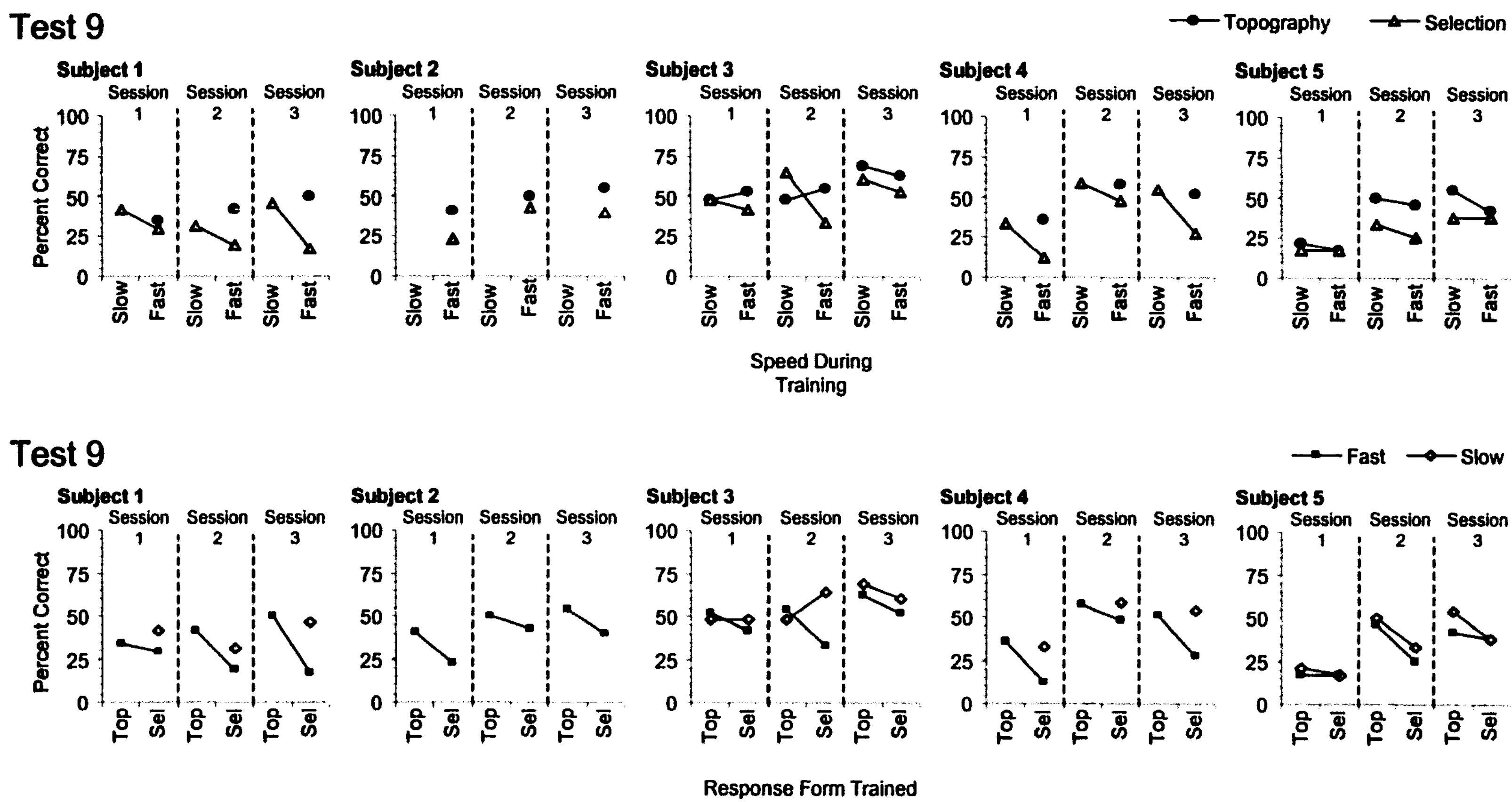


Figure 37. Test Performance for All Subjects and Sessions on Test 9.

Note: On each item in Test 9, subjects saw a Morse Code character on the screen, and were instructed to draw the corresponding Kanji symbol on an answer sheet.

Effects of Response Form Trained

Because some subjects did not respond slowly during all slow-goal training conditions, comparisons between all conditions are not possible. When comparisons are discussed here, they refer to instances where topography- and selection-based relationships were trained under both fast and slow conditions.

On four of the nine tests, the task required a selection-based response. Because of the similarity of this arrangement with selection-based conditions during training, one might expect it to favor relationships learned under those conditions. This did not turn out to be the case. On eight of the nine tests, scores for relationships learned under topography-based conditions exceeded those of selection-based conditions more than half the time, ranging from 52% on Test 3 to 86% on Test 9. On Test 5, scores for relationships learned under topography-based conditions exceeded those of selection-based conditions 48% percent of the time, were about the same 29% of the time, and were less than those for selection-based relationships 24% of the time.

Effects of Speed During Training

Subjects scored higher on relationships learned at a slow speed than those learned at a fast speed 75% of the time. Because no topography-based fast scores are available for three of the subjects, it is impossible to tell whether this would still be the case had they followed instructions. However, given the persistent high scores of relationships learned under topography-based conditions throughout the tests, it is probable that they would have inflated rather than deflated the slow-speed scores.

Effects of Interactions Between Training Speed and Response Form Trained

For Subjects 1, 2 and 4, there appeared to be little or no interaction between training speed and response form trained on most tests and test sessions. There is some evidence in Subject 3's scores of an interaction favoring relationships learned under topography-based slow conditions. Subject 5 scored higher on topography-based relationships learned under slow conditions than fast ones, but for selection-based relationships, she scored higher on those learned under fast conditions than slow ones. As there are no topography-based slow scores for Subjects 1, 2, and 4, this type of interaction cannot be evaluated for those subjects.

Strength of Type Response Form

Johnson and Layng (1992) discussed fluent behavior in terms of being more readily available to the selecting environment than less- or non-fluent behavior. It is conceivable that one type of response form (e.g., selection- or topography-based) might also be more readily available to the selecting environment than the other. To test this, subjects were given the choice of emitting either a topography- or selection-based response on all items on Tests 6 and 7. A tendency to emit responses in the same type of response form (i.e., selection- or topography-based) as originally trained might be evidence that training conditions have influenced the degree to which that response form is available to the selecting environment. On the other hand, a tendency to emit responses in a different response form might indicate that a particular type of response form is more readily available to the selecting environment than the other irrespective of training conditions.

The percent of responses emitted in the original response form trained is displayed in Figure 38. In most cases (23 of out 30), subjects responded with the same type of response form throughout an entire administration of a test, regardless of the training condition. In five cases, after having responded predominantly using selection-based response forms during earlier test sessions, subjects changed to topography-based response forms during later test sessions. No subjects went from emitting topography-based response forms in earlier sessions to selection-based in later sessions. By the last test session, three of the five subjects were emitting strictly topography-based responses on Test 6, one subject only selection-based responses, and one subject a combination of response forms. By the last test session for Test 7, three emitted only topography-based responses, and two emitted only selection-based responses.

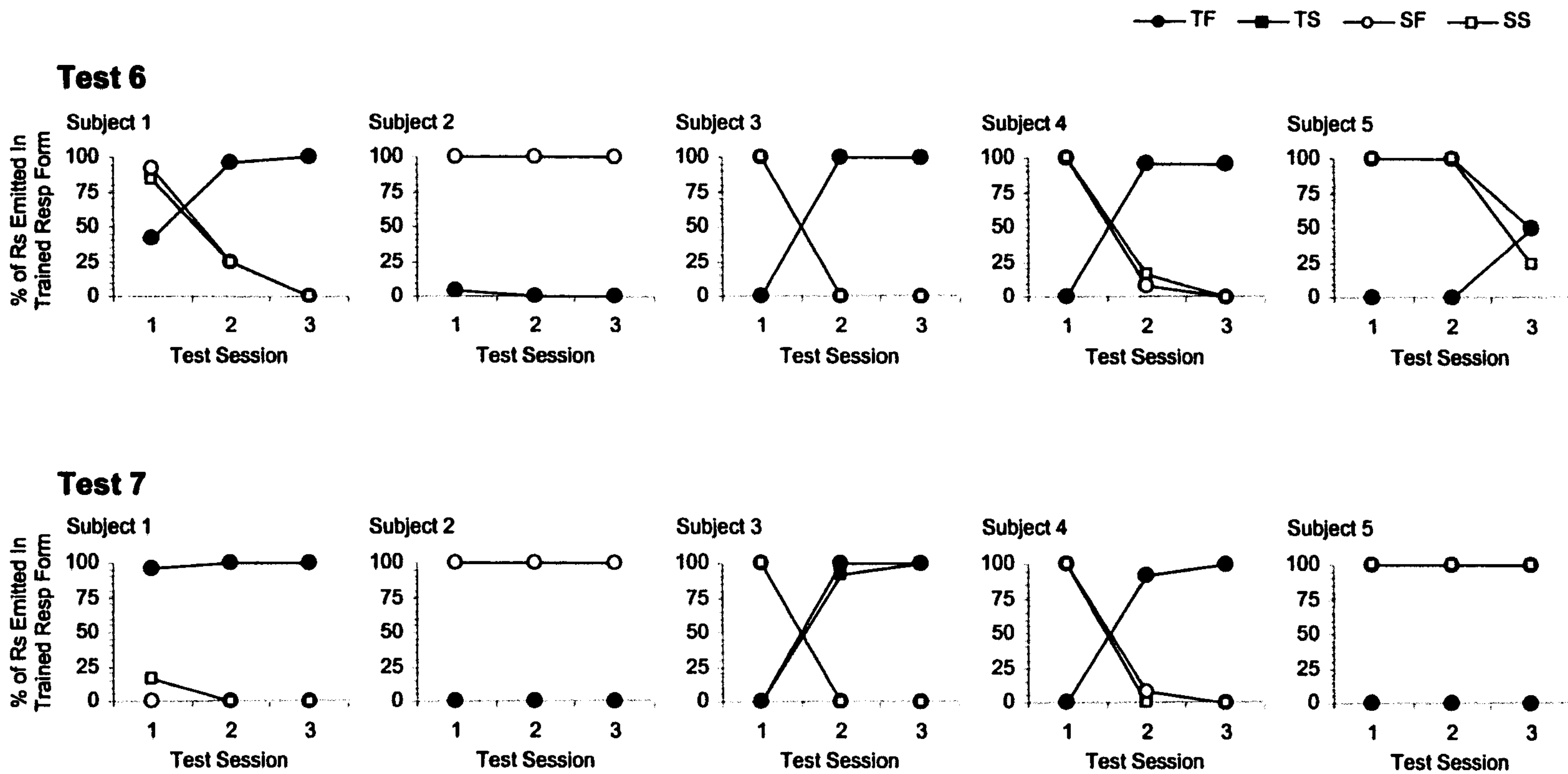


Figure 38. Percent of Responses Emitted in Original Trained Response Form on Tests 6 and 7, by Subject and Session.

CHAPTER VI

DISCUSSION

The purpose of this study was to identify and measure the relative effects of topography- versus selection-based response forms and fast versus slow speed goals on retention, endurance, equivalence, and strength of response form. Three training sessions consisted of four sets of practice drills during which subjects performed matching-to-sample tasks under four distinct conditions: topography-based with slow speed goal, topography-based with fast speed goal, selection-based with slow speed goal, and selection-based with fast speed goal. A different group of six relationships between Kanji symbols and Morse Code characters was taught under each of the four conditions. Because net latencies for some subjects were as fast under slow-goal conditions as under fast-goal conditions, outcomes were interpreted in terms of the effects of actual net latencies during training rather than speed goal during training.

Main Conclusions

Results clearly favored topography-based learning conditions. Although subjects performed more accurately under selection-based conditions during training, performance on relationships learned under topography-based conditions exceeded that of relationships learned under selection-based conditions on three follow-up test sessions, and deteriorated less over the two-weeks during which follow-up tests were conducted.

The effects of speed during training produced inconclusive results. Only two subjects followed instructions to the letter with respect to speed goal during training. Of these, one subject's follow-up test performance was better for relationships learned under slow training conditions, and the other was better for those learned under fast conditions. Two subjects responded at a fast rate under both (fast- and slow-speed goal) topography-based conditions during training, but responded in accordance the speed goal during selection-based training conditions. For relationships learned under selection-based conditions, both subjects' follow-up test performances were better for relationships learned under slow conditions than for those learned under fast conditions. One subject responded at a fast rate throughout all training sessions; therefore, no comparison based on speed is possible for that subject's scores.

An interaction between response form trained and speed during training was observed only in Subject 5's scores. Topography-based relationships learned under slow conditions produced higher follow-up test scores than did those learned under fast conditions. Follow-up test scores for relationships learned under selection-based conditions were just the opposite—scores for relationships learned at a fast rate were higher than those learned under at a slow rate.

The effects described above persisted throughout all nine tests over all three test sessions. That is, for a given relationship (within subject), when retention was high, endurance and equivalence scores were also high. When retention was low, endurance and equivalence scores were low. With respect to strength of response form, when subjects could choose between response forms, results favored topography-based response forms regardless of conditions under which individual relationships were trained.

Theoretical Implications

Sundberg (1990), Wraikat (1990, 1991), and McCarty (1997) demonstrated that topography-based relationships are acquired more easily than selection-based ones. Wraikat (1991) further demonstrated that equivalence relations are learned more readily with topography-based behavior than with selection-based behavior. This study extends that research, demonstrating that topography-based training produces better retention, endurance, and equivalence. Results of this study also suggest that over time, topography-based response forms might be more readily available to the selecting environment than selection-based response forms.

The results of this study do not support the findings of Berquam (1981), Ivarie (1986), Kelly (1996), and Shirley and Pennypacker (1994). In those studies, when practice was held constant, higher rates during training produced better learning gains than slower rates. While practice was also held constant here, higher rates did not produce better learning outcomes.

In this study, it was predicted that net latencies of topography-based responses would decrease steadily and persistently compared to those of selection-based responses. Although net latencies of topography-based responses underwent greater decreases than did those of selection-based responses, there was far more variation than was seen in McCarty (1997). This might have been due to the difference in types of topographies. In the current study, topography-based responses involved tapping out a Morse Code character on the computer mouse, whereas in McCarty (1997), topography-based responses consisted of making handshapes from the manual alphabet. In terms of the complexity of the response, the making of a handshape is simple compared to the tapping out of a Morse Code character, which is actually a cluster of two or more responses in quick session. That difference in complexity might account for the difference in variation of net latencies seen between the two studies.

It is also possible that covert intervening behavior (i.e., mnemonic behavior) was more complex and persisted longer for the Morse Code response than for the handshape response. For example, the only types of intervening behavior mentioned by subjects making the handshapes were descriptions of the similarities in appearance between the sample and choice stimuli (e.g., the head of a bull seen in a Kanji symbol relating to the  handshape). With the Morse Code symbols, there might have been covert auditory as well as covert visual intervening responses. For example, after responding covertly to similarity between the Kanji, 名, and the Morse, −•−•, a subject might have made a covert auditory response corresponding to the target response (e.g., "daaah-duh-daaah-duh").

If combinations of intervening responses such as these did occur, and if one type faded from the repertoire more quickly than the other did, it might account for the variation in net latencies seen throughout topography-based training sessions. If training had continued for more sessions, a pattern of steady and persistent decreases in response latencies similar to that seen in McCarty (1997) might have occurred here as well. Further, if this pattern of behavior had occurred sooner under fast speed conditions than under slow speed ones, more distinctive outcomes associated with training speed might have been produced.

Practical Implications

The results of this study continue to support the use of topography-based behavior during training, which is the general practice among practitioners of the fluency procedure. As computer-based instruction becomes more commonplace in education at all levels, there is a temptation to use selection-based response forms due to the ease of development of instructional materials. Input devices used in computer-based instruction are usually limited to the keyboard and the mouse, neither of which is conducive to producing topography-based response forms. Responses using the computer mouse are virtually always selection-based. Short answers can be entered using the keyboard, but unless the student is a very fluent typist, typed answers are also selection-based. When responses are typed, issues such as spelling, tense, number, word form, and synonyms, make effective evaluation by the computer program more difficult to program.

At the time this study was designed, voice-activation software capable of enabling the effective use of vocal responses during computer-based fluency drills was not available. This type of software has undergone tremendous technological advances recently, and may now be appropriate for use in fluency research. A suggestion for future research would be to replicate this study using voice-activation software with vocal responses, rather than Morse Code, as the topography-based response form. One-word vocal responses would be more comparable to the handshape response used in the McCarty (1997) study than Morse Code responses, and might produce patterns of net latencies similar to those seen in that study. Such a study might demonstrate (a) whether intervening behavior in the stimulus-response chain can be eliminated through the use of standard topography-based response forms, (b) whether decreased latencies are indicators that such an event has occurred, (c) whether the imposition of a high speed criterion facilitates the occurrence of such an event, and (d) whether such a procedure produces learning gains superior to those seen without a speed criterion.

Appendix A

Protocol Clearance From the Human Subjects Institutional Review Board

Human Subjects Institutional Review Board

Kalamazoo, Michigan 49008-3899

WESTERN MICHIGAN UNIVERSITY

Date: 9 April 1999

To: Jack Michael, Principal Investigator
Amy McCarty, Student Investigator for dissertation

From: Sylvia Culp, Chair Sylvia Culp

Re: HSIRB Project Number 99-03-16

This letter will serve as confirmation that your research project entitled "A Comparison of the Effects of Temporal Requirement and Response Form on Retention, Equivalence, and Endurance" has been **approved** under the **exempt** category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note that you may **only** conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: 9 April 1999

Appendix B

Consent Form

Western Michigan University
Department of Psychology
Principal Investigator: Jack Michael, Ph.D.
Student Investigator: Amy McCarty, M.A.

You are invited to participate in a research project entitled "A Comparison of the Effects of Temporal Requirement and Response Form on Retention, Equivalence, and Endurance" designed to analyze various features of instructional design and how they interact to affect learning outcomes, being conducted by Dr. Jack Michael and Ms. Amy McCarty from Western University, Department of Psychology. This research is being conducted as part of the dissertation requirements for Amy McCarty.

Participants will attend eight research sessions with Amy McCarty at her home office at 740 Strawberry Valley Ave., NW, Comstock Park, Michigan. The first session will involve prerequisite skill training and testing to determine whether participants' computer mouse skills are sufficient for this study. If the results of the prerequisite tests indicate insufficient skill using the computer mouse, you will be excluded from the remainder of the study. Participants will be asked to provide general information about themselves, such as their age, and level of education. The next four sessions will involve matching Morse Code characters to Kanji symbols by selecting them on the screen or by tapping them out on the computer mouse. Several drill and practice exercises will be conducted during each of these four sessions. During three testing sessions, you will take a variety of computer-based and written tests covering the material learned during the training sessions. Each session will last approximately one hour.

As in all research, there may be unforeseen risks to the participant. If an accidental injury occurs, appropriate emergency measures will be taken; however, no compensation or treatment will be made available except as otherwise specified in this consent form. One potential risk of participation in this project is that you may experience the type of stress that sometimes accompanies test taking; however, you can take a break at any time to relax.

Although participants may not gain directly and personally from the results of this research, others may benefit if the research results lead to increased efficiency and effectiveness in instructional design methods in the educational community. You will receive $5 compensation for each session you attend.

All of the information collected is confidential. That means that your name will not appear anywhere this information is recorded. Instead of actual names, pseudonyms will be used, and only Ms. McCarty will know the actual names of participants. All forms will be retained for three years in a locked file in the principal investigator's laboratory.

You may refuse to participate or quit at any time during the study without prejudice or penalty. If you have any questions or concerns about this study, you may contact either Amy McCarty at (616) 784-7743, or Dr. Jack Michael at (616) 372-3075. You may also contact the chair of Human Subjects Institutional Review Board at 387-8293 or the vice president for research at 387-8298 with any concerns that you have.

This consent document has been approved for use for one year by the Human Subjects Institutional Review Board as indicated by the stamped date and signature of the board chair in the upper right corner. You should not participate in this project if the corner does not have a stamped date and signature.

Appendix C

Study Sheet of All Kanji/Morse Relationships

Over the next few sessions, you will practice matching Kanji symbols to Morse Code characters. The purpose of this study sheet is to introduce you to all of the pairings. Study each pair below, and try to memorize which Kanji symbol goes with which Morse Code character. Feel free to make up mnemonics to help you remember. Don't worry if you can't memorize them all. You will have plenty of opportunity to practice.

女 = –●	町 = ●–	月 = ––	赤 = ●●
男 = ●––	右 = ––●	音 = –●●	早 = ●●–
車 = –●–	草 = ●●––	学 = ●●●–	金 = ●–●
村 = ●●–●	森 = ●–●●	名 = –●–●	虫 = ●––●
糸 = –●●●	竹 = ●–●–	左 = –●––	文 = ●–––
玉 = –––●	四 = –●●–	天 = ––●●	空 = ––●–

Appendix D

Individual Net Latencies for All Subjects and Training Trials by Individual Morse/Kanji Relationship

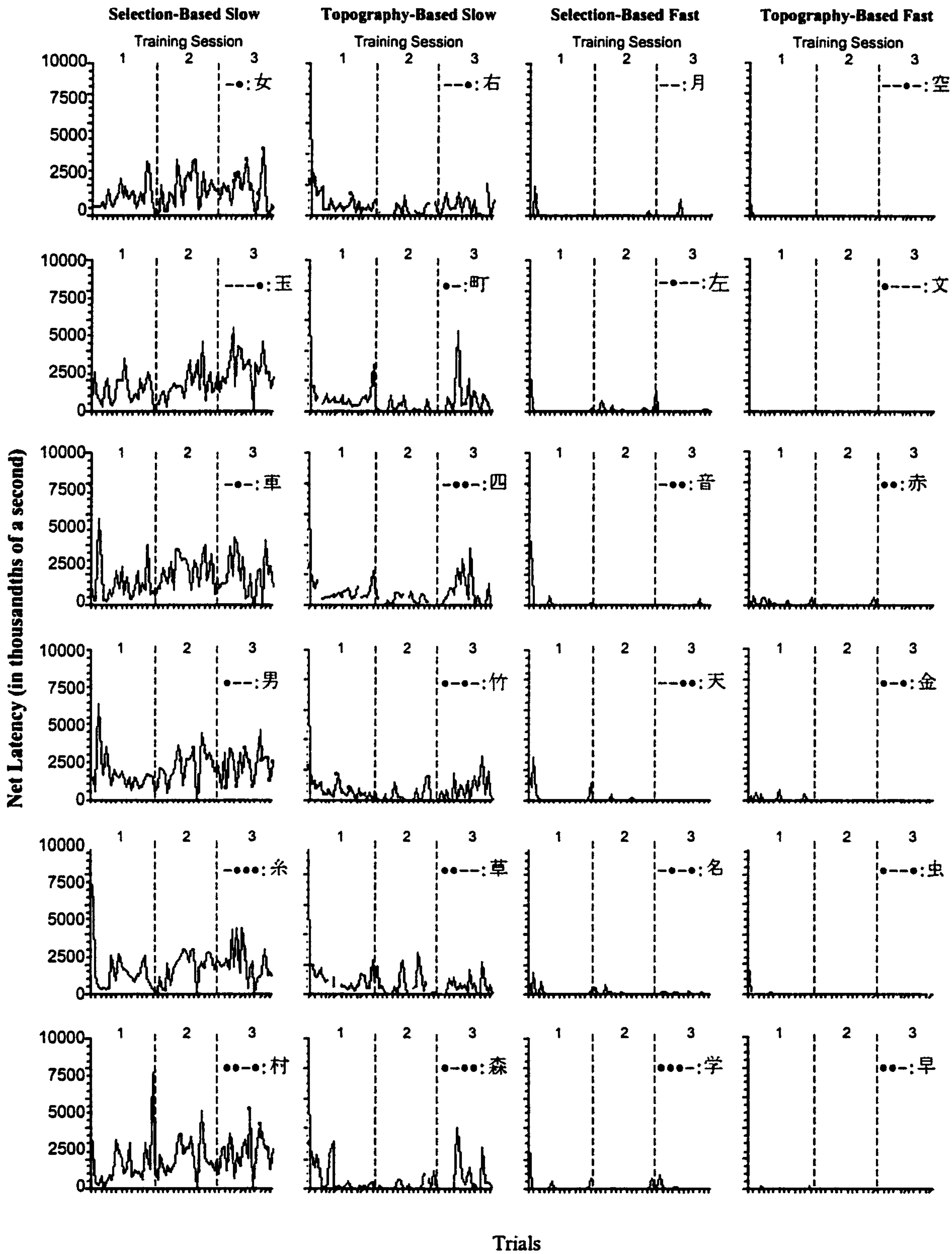


Figure D1. Net Latencies for All Training Trials for Subject 1, in Chronological Order by Individual Morse/Kanji Relationship.

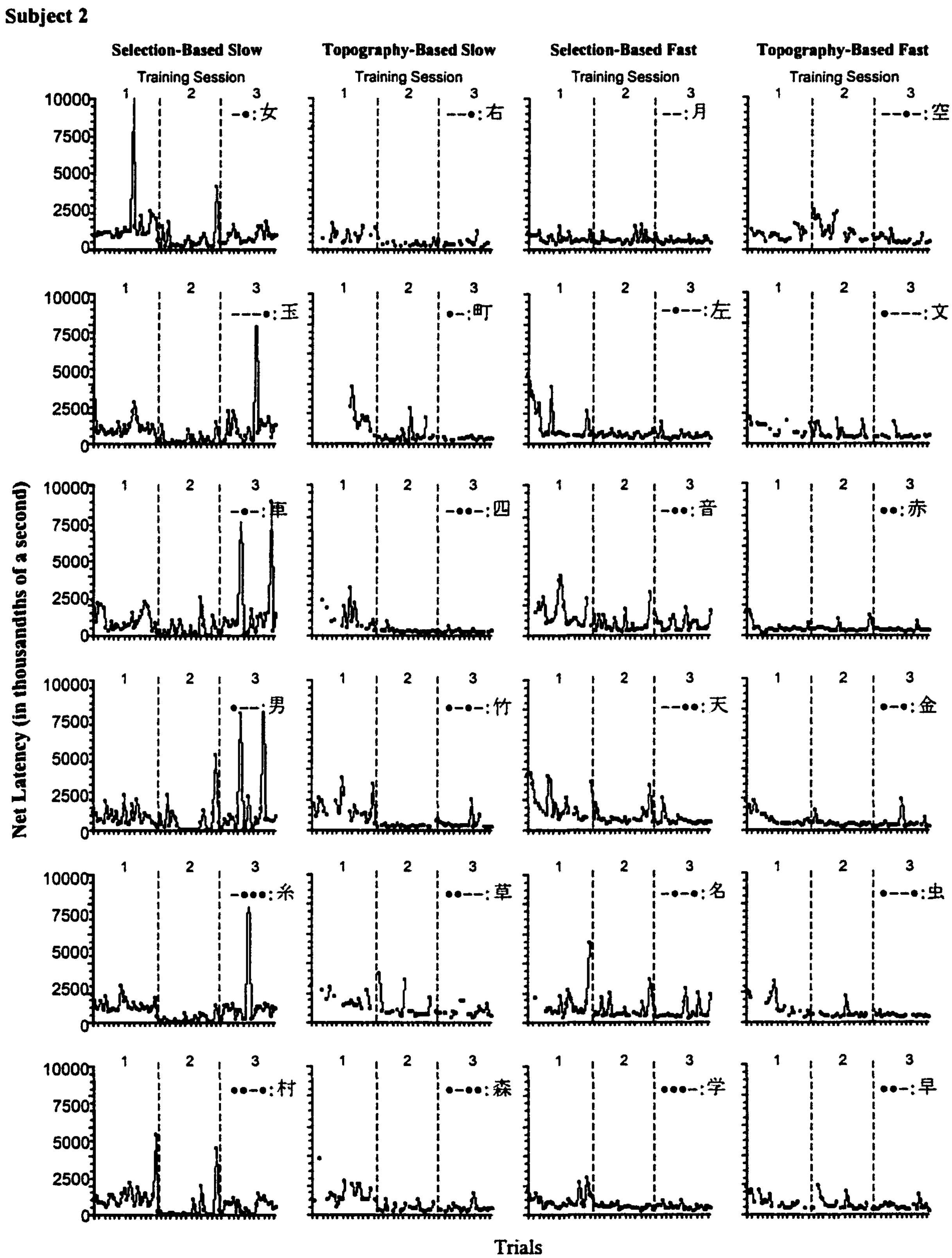


Figure D2. Net Latencies for All Training Trials for Subject 2, in Chronological Order by Individual Morse/Kanji Relationship.

Subject 3

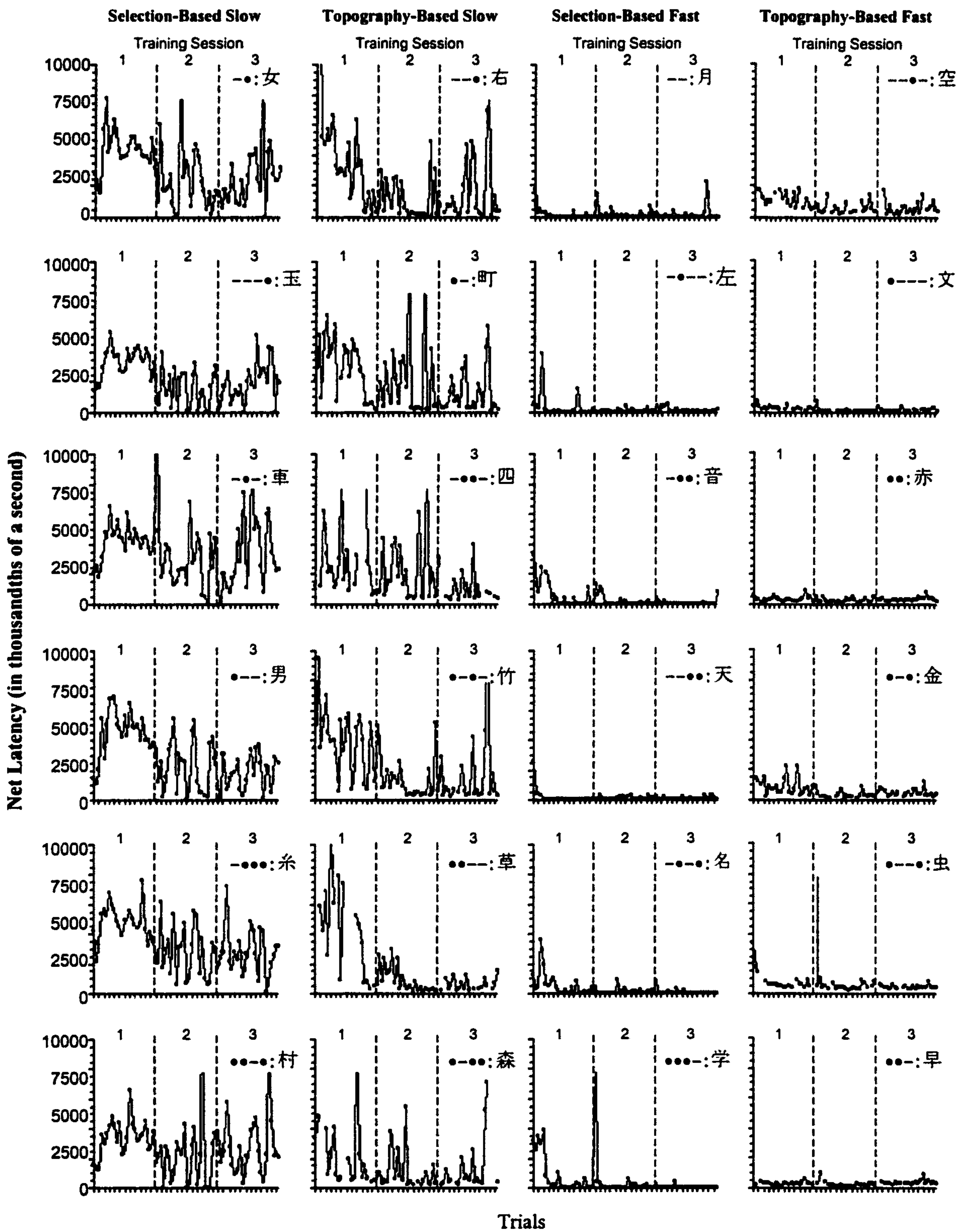


Figure D3. Net Latencies for All Training Trials for Subject 3, in Chronological Order by Individual Morse/Kanji Relationship.

Subject 4

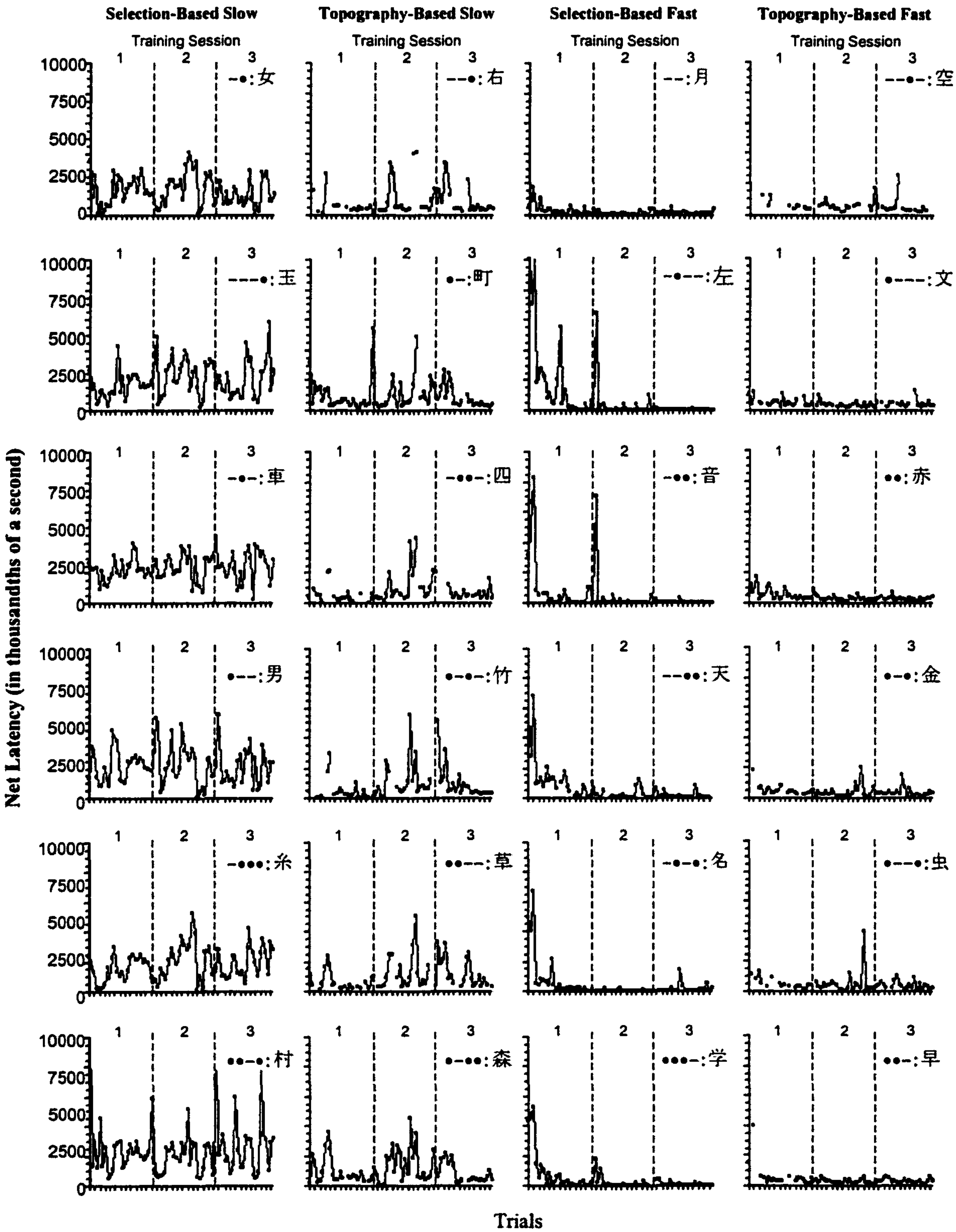


Figure D4. Net Latencies for All Training Trials for Subject 4, in Chronological Order by Individual Morse/Kanji Relationship.

Subject 5

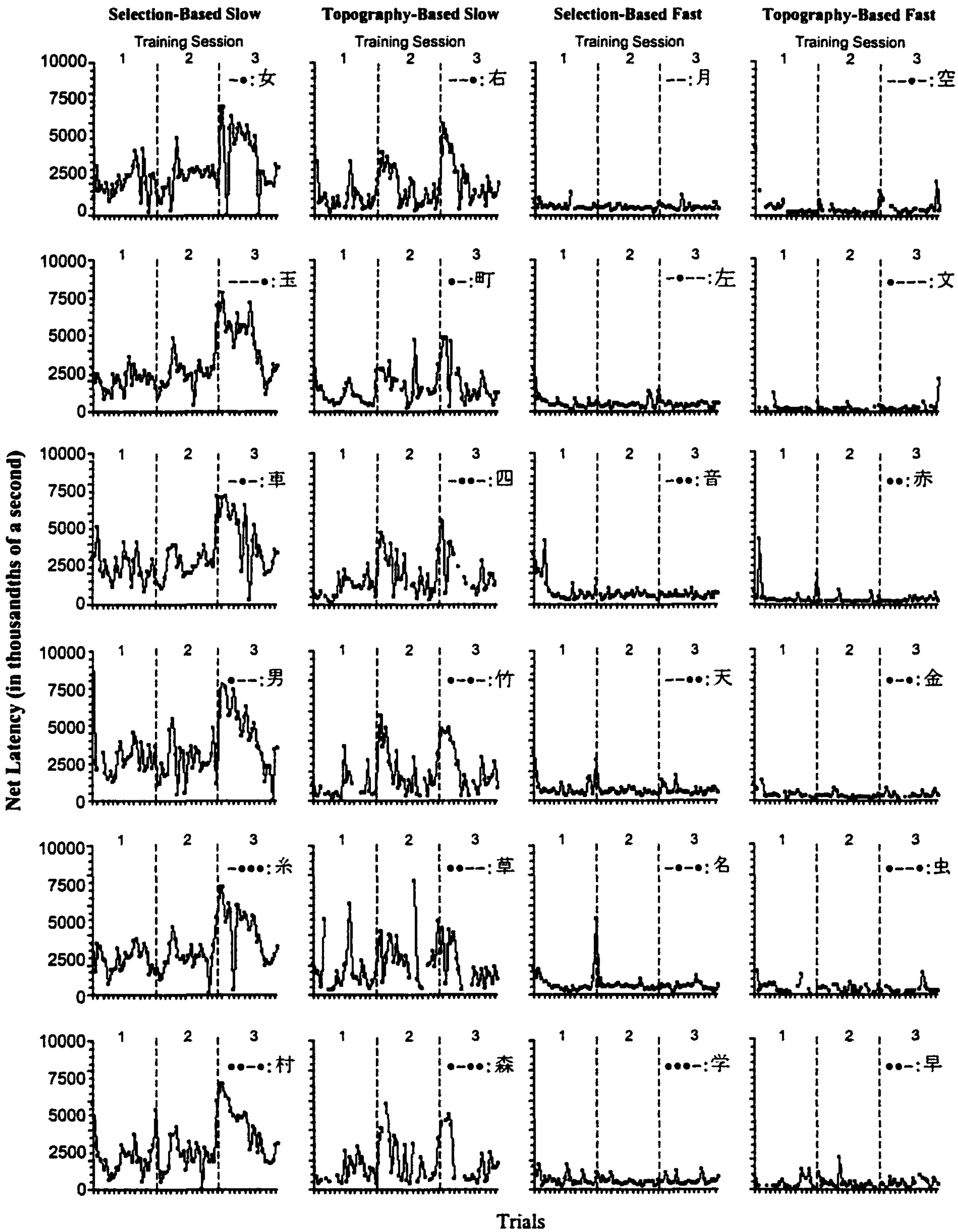


Figure D5. Net Latencies for All Training Trials for Subject 5, in Chronological Order by Individual Morse/Kanji Relationship.

BIBLIOGRAPHY

Alessi, G. (1987). Generative strategies and teaching for generalization. The Analysis of Verbal Behavior, 5, 15-27.

Barbetta, P. M., Heward, W. L., Bradley, D. M., & Miller, A. D. (1994). Effects of immediate and delayed error correction on the acquisition and maintenance of sight words by students with developmental disabilities. Journal of Applied Behavior Analysis, 27, 177-178.

Beck, R., & Clement, R. (1991). The Great Falls Precision Teaching Project: An historical examination. Journal of Precision Teaching, 8(2), 8-12.

Berquam, E. (1981). The relation between frequency of response and retention on a paired associates task. Unpublished doctoral dissertation, University of Florida, Gainesville.

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

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. The Behavior Analyst, 19, 163-197.

Binder, C., & Bloom, C. (1989). Fluent product knowledge: Application in the financial services industry. Performance and Instruction, 28(2), 17-21.

Binder, C., Haughton, E., & Van Eyk, D. (1990). Precision teaching attention span: Increasing endurance by building fluency. Teaching Exceptional Children, 22(3), 24-27.

Buckingham, B. R. & Dolch, E. W. (1936). A Combined Word List, Compiled by B. R. Buckingham and E. W. Dolch. Boston, MA: Ginn and Company.

Cohen, M. A., Gentry, D. N., Hulten, W. J., & Martin, G. L. (1972). Measures of classroom performance. In N. G. Haring & A. Hayden (Eds.), The Improvement of Instruction. Seattle, WA: Special Child Publications.

Comunidad Los Horcones (1987). The concept of consequences in the analysis of behavior. The Behavior Analyst, 10(2), 291-294.

Cresson, O. (1994). The writing response in studies of topography-based and selection-based verbal behavior. Unpublished doctoral dissertation, Western Michigan University, Kalamazoo, Michigan.

Engelmann, S., & Carnine, D. W. (1982). Theory of Instruction. New York: Irvington.

Evans, S. S., Mercer, C. D., & Evans, W. H. (1983). The relationship of frequency to subsequent skill acquisition. Journal of Precision Teaching, 4(2), 28-34.

Grindle, A. C. (1997). The effects of fluency level on retention. Unpublished master's thesis, Western Michigan University, Kalamazoo.

Haughton, E. (1971). Great gains from small starts. The Teaching Exceptional Children, Spring, 141-146.

Haughton, E. (1972). Aims—growing and sharing. In J. B. Jordan & L. S. Robbins (Eds.), Let's try doing something else kind of thing: Behavior principles and the exceptional child (pp. 20-39). Reston, VA: The Council for Exceptional Children.

Haughton, E. (1980). Practicing practices: Learning by activity. Journal of Precision Teaching, 1(3), 3-20.

Ivarie, J. J. (1986). Effects of proficiency rates on later performance of a recall and writing behavior. Remedial and Special Education, 7(5), 25-30.

Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. American Psychologist, 47, 1475-1490.

Johnson, K. R., & Layng, T. V. J. (1994). The Morningside Model of generative instruction. In R. Gardner, III, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. W. Eshelman, & T. A. Grossi (Eds.), Behavior analysis in education: Focus on measurably superior instruction (pp. 173-197). Belmont, CA: Brooks/Cole.

Kelly, R. L. (1996). A functional analysis of the effects of mastery and fluency on maintenance. Unpublished doctoral dissertation, Columbia University Graduate School of Arts and Sciences, New York.

Lindsley, O. R. (1956). Operant Conditioning Methods Applied to Research in Chronic Schizophrenia. Psychiatric Research Reports, 5, 118-139.

Lindsley, O. R. (1990). Precision teaching: By teachers for children. Teaching Exceptional Children, 22(3), 10-15.

Lindsley, O. R. (1995). Ten products of fluency. Journal of Precision Teaching, 13(1), 2-11.

Makepeace, M. D. (1998). The effect of rate vs. number of emissions of an operant behavior on behavior probability. Unpublished doctoral dissertation, Western Michigan University, Kalamazoo.

McCarty, A. (1997). Comparing selection- and topography-based fluency using repeated acquisition. Unpublished master's thesis, Western Michigan University, Kalamazoo.

McDade, C. E., Austin, D. M., & Olander, C. P. (1985). Technological advances in precision teaching: A comparison between computer-testing and SAFMEDS. Journal of Precision Teaching, 6(3), 49-54.

McDade, C. E. & Goggans, L. A. (1993). Computer-based precision learning: Achieving fluency with college students. Education and Treatment of Children, 16, 290-305.

Michael, J. (1985). Two kinds of verbal behavior plus a possible third. The Analysis of Verbal Behavior, 3, 1-4.

Shirley, M. J., & Pennypacker, H. S. (1994). The effects of performance criteria on learning and retention of spelling words. Journal of Precision Teaching, 12(1), 73-86.

Stratton, M. (1992). Comparing selection-based and topography-based language systems with verbal adults learning Japanese words. Unpublished masters thesis, Western Michigan University, Kalamazoo, Michigan.

Sundberg, C. T., & Sundberg, M. L. (1990). Comparing topography-based verbal behavior with stimulus selection-based verbal behavior. The Analysis of Verbal Behavior, 8, 31-41.

Toolbook II Instructor (Version 5.0) [Computer software]. (1990-1996). Bellevue, WA: Asymetrix Corporation.

Van Houten, R., Hill, S., & Parsons, M. (1975). An analysis of a performance feedback system: The effects of timing and feedback, public posting, and praise upon academic performance and peer interaction. Journal of Applied Behavior Analysis, 8, 449-457.

Van Houten, R., Morrison, E., Jarvis, R., & McDonald, M. (1974). The effects of explicit timing and feedback on compositional response rate in elementary school children. Jounal of Applied Behavior Analysis, 7, 547-555.

White, O. R. (1986). Precision teaching—precision learning. Exceptional Children, 52, 522-534.

Wraikat, R. M. (1990). Teaching tact and intraverbal behavior to developmentally disabled adults: A comparison of topography-based and selection-based paradigms. Unpublished project, Western Michigan University, Kalamazoo, Michigan.

Wraikat, R. M. (1991). Ease of learning and the emergence of equivalence class formation: A comparison of topography-based and selection-based paradigms. Unpublished doctoral dissertation, Western Michigan University, Kalamazoo, Michigan.