4-2002

An Examination of the Effects of Fluency Training on Retention, Distractibility, and Generativity

Victoria Mary Pellettiere
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Cognitive Psychology Commons, Experimental Analysis of Behavior Commons, and the Theory and Philosophy Commons

Recommended Citation
Pellettiere, Victoria Mary, "An Examination of the Effects of Fluency Training on Retention, Distractibility, and Generativity" (2002). Dissertations. 1298.
https://scholarworks.wmich.edu/dissertations/1298

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
AN EXAMINATION OF THE EFFECTS OF FLUENCY TRAINING ON RETENTION, DISTRACTIBILITY, AND GENERATIVITY

by

Victoria Mary Pelletiere

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
April 2002
AN EXAMINATION OF THE EFFECTS OF FLUENCY TRAINING ON RETENTION, DISTRACTIBILITY, AND GENERATIVITY

Victoria Mary Pelletiere, Ph.D.
Western Michigan University, 2002

The Precision Teaching movement grew out of a commitment to use frequency as a universal measure of behavior, as well as the desire to employ research methods derived from the experimental analysis of behavior in education (Lindsley, 1991). One component of instruction employing the precision teaching model is fluency training that typically involves exposing learners to the training materials until they have met criteria for both accuracy and speed. Proponents of fluency training ascribe a number of specific benefits to this instructional tool. Lindsley (1992, 1995) and others (Binder, 1993, 1996; Haughton, 1981b) suggested that fluency training enhances retention, endurance, application, understanding, and generativity. Unfortunately, the empirical data to support the proclaimed benefits is severely limited (Berquam, 1985). The purpose of the present research was to evaluate the effects of training using an accuracy criterion alone versus accuracy and rate criteria on retention, distractibility, and generativity. The first study examined the effects of training component skills to different levels of fluency on the retention of those skills and the acquisition and retention of more complex, composite skills (generativity). The second study examined the effects of training under an accuracy-alone criterion or accuracy plus rate criteria on distractibility. The results of these studies suggest that fluency training does not produce superior retention or generativity, and does not produce superior performance in the face of distracting stimuli.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
ACKNOWLEDGMENTS

I wish to acknowledge and thank the many friends and colleagues who gave of their time to participate in these experiments. I would also like to thank the staff at Living Ways who, by doing their jobs as well as they do, permitted me the flexibility to complete this dissertation without concerns about our services. I would especially like to thank Fi Spalvieri for her support as my colleague, boss, and most importantly, friend. I could not have completed this without your on-going support.

I would also like to thank the members of my committee, Dr. Ron Crowell, Dr. Wayne Fuqua, Dr. Jack Michael, and Dr. Alan Poling for reviewing my work. Special thanks to Dr. Jack Michael for his patience, support, and teaching over these many years. I can say without hesitation that I have learned more from you than any other professor or colleague.

Finally, I would like to thank my family and friends for their support over the many years that I have been in graduate school. A special thanks to my sister Amy for countless pep talks and a lot of good advice. And finally, special thanks must go to my mother who has been my greatest supporter, cheerleader, role model, teacher, and friend through this and all my other journeys in life. I am proud to be your daughter. This dissertation is dedicated to my mother for her on-going love and support and to the memory of my father who supported me in all of my endeavors.

Victoria Mary Pelletiere

ii
## TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................ ii

LIST OF TABLES ................................................................................................................ vii

LIST OF FIGURES ............................................................................................................... viii

CHAPTER

I. INTRODUCTION ........................................................................................................ 1

Response Rate as a Critical Measure of Behavior .......................................................... 1

The Precision Teaching Model ...................................................................................... 7

Fluency Training ............................................................................................................. 15

Retention ....................................................................................................................... 16

Endurance / Distractibility ............................................................................................ 34

Generativity ................................................................................................................... 41

Purpose of Current Research ....................................................................................... 50

II. EXPERIMENT I: EFFECTS OF FLUENCY TRAINING ON RETENTION AND GENERATIVITY ................................................................. 53

Participants .................................................................................................................. 53

Setting ........................................................................................................................... 53

Apparatus ..................................................................................................................... 54

Experimental Design ................................................................................................... 54

Procedure ..................................................................................................................... 55

Training ........................................................................................................................ 56

Retention Probes ......................................................................................................... 60
Table of Contents—continued

Generativity Probes................................................................. 61
Results / Discussion............................................................... 62
Analysis of the Effects of Character Sets Independent of the Experimental Conditions ........................................... 69
Analysis of the Effects of Order in Training Independent of the Experimental Conditions ..................................... 71
Overall Results for Rate Achieved During Training............. 76
Effects of the Number of Practice Trials............................... 92
Loss of Performance Over Time.............................................. 96
General Discussion................................................................. 99

III. EXPERIMENT II: EFFECTS OF FLUENCY TRAINING ON DISTRACTIBILITY......................................................... 103

Participants........................................................................... 103
Setting.................................................................................... 103
Apparatus.............................................................................. 104
Experimental Design............................................................ 104
Procedure............................................................................. 105
Arithmetic Task................................................................. 107
Memory Task................................................................. 108
Visual Monitoring Task.................................................. 109
Auditory Monitoring Task............................................ 110
Results / Discussion........................................................... 111
Total Score Per Minute.................................................... 112
Percent Correct............................................................ 114

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table of Contents—continued

Correct Responses Per Minute .............................................................. 115
Incorrect Responses Per Minute ........................................................... 116
Performance on Distracting Tasks ....................................................... 117

IV. OVERALL DISCUSSION ........................................................................................ 123

APPENDICES

A. Informed Consent Completed by All Participants .............................................. 126
B. Protocol for Clearance From the Human Subjects Institutional Review Board ................................................................. 129
C. Presentation of Experimental Conditions for Each Participant in the Retention / Generativity Experiment .................................................. 131
D. Character Sets Utilized During Training in the Retention / Generativity Experiment ................................................................................ 133
E. Overview of Individual Sessions in the Retention / Generativity Study ...... 135
F. Mathematical Equations Used During Generativity Probes in the Retention / Generativity Study ................................................................. 138
G. Designation of High, Moderate-, and Low-Rate Groups Based Upon Average Rate of Correct Responding Achieved During the Final Three Sprints of Training ................................................................. 140
H. Individual Results for Retention / Generativity Experiment Corrected With Respect to Character Set and Order in Training ......................... 142
I. Individual Results for Retention / Generativity Experiment Uncorrected With Respect to Character Set and Order in Training .................. 203
J. Average Percent Correct and Latencies for All Participants on all Retention and Generativity Probes Based Upon the Character Set Associated With That Condition ................................................................. 236
K. Average Percent Correct and Latencies for All Participants on all Retention and Generativity Probes Based Upon the Character Set Associated With That Condition ................................................................. 238

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table of Contents—continued

BIBLIOGRAPHY ........................................................................................................... 240

vi
LIST OF TABLES

1. Total Number of Sprints, Practice Trials, and Average Number of Practice Trials per Sprint for All Subjects for All Conditions ............................................ 55

2. Average Number of Correct Responses Per Minute for Each Participant Under All Three Experimental Conditions During the Final 3 Minutes of Training ........................................................................................................... 65

3. Frequency of Association of the Rate Required Under the Nominal Conditions and the Rates Achieved Under Those Conditions .............................. 66

4. Correction Factors for Character Set and Order in Training Employed in the Retention / Generativity Experiment ................................................................. 75

5. Number of Practice Trials Completed for All Subjects in the Fluency and Accuracy-Alone Groups ................................................................................................ 106

6. Comparison of Average Final Training and Testing Scores on the Arithmetic Task for Fluency and Accuracy-Alone Groups ........................................ 113

7. Comparison of Average Final Test Scores on Distracting Tasks for Fluency and Accuracy-Alone Groups ................................................................. 121
## LIST OF FIGURES

1. Screen Template for Relationships Presented as a Group ............................................... 57
2. Screen Template for Match-to-Sample Trials and Retention Probes ................................. 58
3. Screen Template for Graph of Performance Presented Between Sprints......................... 60
4. Screen Template for Generativity Probes ........................................................................ 62
5. Average Percent Correct (upper panel) and Latencies (lower panel) on Retention Probes for All Participants For the Three Character Sets ........................................................................ 73
6. Average Percent Correct (upper panel) and Latencies (lower panel) on Generativity Probes for All Participants For the Three Character Sets .................................................. 74
7. Average Percent Correct (upper panel) and Average Latencies (lower panel) on Retention Probes for All Participants Based Upon Order in Training .............................................................................................................................. 79
8. Average Percent Correct (upper panel) and Average Latencies (lower panel) on Generativity Probes for All Participants Based Upon Order in Training .............................................................................................................................. 80
9. Average Percent Correct (upper panel) and Average Latencies (lower panel) for Nine Participants for High, Moderate and Low-Rate Achieved During Training on Retention Probes .............................................................................................................................. 81
10. Average Percent Correct (upper panel) and Average Latencies (lower panel) for Nine Participants for High, Moderate and Low-Rate Achieved During Training on Generativity Probes .............................................................................................................................. 86
11. Screen Template for Testing Phase of Distractibility Study ............................................. 108
12. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 1 on Retention Probes .............................................................................................................................. 145
13. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 1 on Generativity Probes .............................................................................................................................. 147
14. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 2 on Retention Probes .............................................................................................................................. 151
List of Figures—continued

15. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 2 on Generativity Probes ................................................................. 152

16. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 3 on Retention Probes ................................................................. 155

17. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 3 on Generativity Probes ................................................................. 156

18. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 4 on Retention Probes ................................................................. 160

19. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 4 on Generativity Probes ................................................................. 161

20. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 5 on Retention Probes ................................................................. 165

21. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 5 on Generativity Probes ................................................................. 166

22. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 6 on Retention Probes ................................................................. 169

23. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 6 on Generativity Probes ................................................................. 170

24. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 7 on Retention Probes ................................................................. 173

25. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 7 on Generativity Probes ................................................................. 175

26. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 8 on Retention Probes ................................................................. 180

27. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 8 on Generativity Probes ................................................................. 181

28. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 9 on Retention Probes ................................................................. 185
List of Figures—continued

29. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 9 on Generativity Probes .................................................. 186

30. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 10 on Retention Probes ..................................................... 190

31. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 10 on Generativity Probes ............................................. 191

32. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 11 on Retention Probes ..................................................... 194

33. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 11 on Generativity Probes ............................................. 195

34. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 12 on Retention Probes ..................................................... 199

35. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 12 on Generativity Probes ............................................. 202

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
CHAPTER I

INTRODUCTION

Response Rate as a Critical Measure of Behavior

One goal of the analysis of behavior is to understand how environmental variables, be they motive, discriminative, or consequent, affect the learning of organisms. Learning may be defined as a change in performance, or a change in the occurrence of a behavior (Skinner, 1972; West & Young, 1992). While a variety of measures have been employed to evaluate learning, Skinner asserted that "rate of responding appears to be the only datum which varies significantly and in the expected direction under conditions which are relevant to the learning process" (Skinner, 1972, p. 75). Binder also suggested that most of the principles and methods derived from the experimental analysis of behavior and subsequently applied in classrooms, including schedules of reinforcement, extinction, response shaping, stimulus fading, and discrimination, were derived from experiments in which the primary data were response rate measures (Binder, 1993, p. 8).

Response rate, or frequency, may be defined as behavioral events per unit time (Grindle, 1996; Lindsley, 1991). Frequency has been employed widely as a measure of learning in laboratory settings (Ferster & Skinner, 1957; Keller & Schoenfeld, 1950; Skinner, 1938, 1953). However, when early behavior analysts began to extend...
basic laboratory findings to education and training, the use of rate of response as a
dependent measure declined, and emphasis shifted toward more traditional measures
like accuracy (Lindsley, 1990; Skinner, 1954, 1968). This practice has continued
such that behavior analysts working in the field of education seldom use frequency as
a measure of learning by (Binder, 1993; Binder & Watkins, 1990; Lindsley, 1992).

Proponents of the precision teaching model believe that the neglect of
frequency as a critical measure of learning has significant negative consequences for
White 1985, 1986). Their concerns revolve around the imprecision of percent correct
and other traditional measures and the implications of that imprecision for the
accurate measurement of learning both simple and complex behavior. Researchers
and practitioners of this model make a variety of assertions about the superiority of
frequency as a measure of learning based upon specific assumptions about the
learning process. Although the empirical data available to support these assumptions
are limited, their assertions will be summarized briefly.

One concern is that accuracy measures alone can not distinguish between
merely accurate performance and mastery because accuracy measures tell little about
the quantity of performance - only large differences in quality are revealed (Binder,
1993, 1995; West, Young, & Spooner, 1990; White, 1986). The presence of accurate
behavior alone, however, may not be sufficient to ensure that a learner will acquire
more complex behaviors of which the accurate behavior is a component (Gagne,
1970; Haughton, 1980). White (1985) suggested further that accuracy measures alone
may even be misleading indicators of whether a pupil is really using the measured skill at all. For example, a student may be able to complete simple addition problems with 100% accuracy, but only at a rate of four problems per minute. One might predict that this behavior would be exhibited infrequently outside of the classroom setting. Because frequency measures can differentiate between performances that may be equally accurate, but vary with respect to how quickly the behavior is emitted, they can quantify ranges in performance, thereby differentiating between mastery and merely accurate performance (Binder & Bloom, 1989; Grindle, 1996; Shirley & Pennypacker, 1994; West & Young, 1992; West et al., 1990).

Similarly, since accuracy measures alone cannot gauge performance beyond 100% accuracy, these measures cannot identify differences in performance between learners who achieve the same level of accuracy, or for the same individual over time (Binder, 1998, 1993; Binder & Bloom, 1989; Howell & Lorson-Howell, 1990; West & Young, 1992). These differences may be critical in determining the future success of students in learning simple skills and more complex skills of which the original skills are components (Binder, 1996; Gagne, 1970; Haughton, 1972; Makepeace, 1998). Frequency measures, however can measure those differences. For example, Lindsley (1991) found behavioral frequencies to be ten to one hundred times more sensitive to the effects of drugs and different reinforcers than percent correct. Similarly, frequency measures revealed 40 times more effects of curricular changes than did percent correct on performance on the same practice sheets in the classroom (Holzschuh and Dobbs, 1966). If frequency measures permit a more careful
evaluation of the effects of different independent variables on learning than accuracy alone, their use may be critical for the careful analysis of learning both simple and complex skills (Berquam, 1985, p.324).

A third concern revolves around the issue of cumulative dysfluency (Binder, 1993, 1996). Many researchers distinguish between component skills and more complex composite skills that are comprised of the component skills (Binder, 1993; Grindle, 1996; Haughton, 1972; Johnson & Layng, 1992, 1994). For example, saying sounds and words may be viewed as component skills necessary to emit the more complex, composite behavior of oral reading. The relationship between these two types of behavior has been widely studied (Barrett, 1979; Bilodeau & Bilodeau, 1954; Gagne & Foster, 1949; Gagne, Baker, & Foster, 1950; Johnson & Layng, 1992, 1994; Haughton 1971, 1972, 1980; Layng & Andronis, 1984). Many of these researchers argue that before a learner is able to exhibit competent performance of a composite skill, that learner must achieve fluency on the component skills that comprise it (Binder, 1993, 1996; Haughton, 1972, 1981b; Johnson & Layng, 1992, 1994; Medade, 1992). Binder (1996, p. 163) defined behavioral fluency as “that combination of accuracy plus speed that enables competent individuals to function efficiently and effectively in their natural environments”. He maintains that “behavioral components with frequency deficits, despite their accuracy, accumulate when they are layered on top of one another in a curriculum sequence. This accumulation of dysfluent skills limits, and may prevent, the acquisition of composites that depend on them” (Binder, 1996, p. 183; Johnson & Layng, 1992).
The overall effect of the repeated acquisition of skills that are accurate, but have not been trained to fluency on future learning has been termed cumulative dysfluency (Binder, 1993, 1996). Many fluency-oriented educators see this phenomenon as the single most important factor in long-term student failure (Binder, 1988b, 1996; Johnson & Layng, 1992; Pennypacker & Binder, 1992). Unfortunately, traditional measures of accuracy do not detect the speed with which component skills are performed. It is, therefore, impossible to detect cumulative dysfluency until the final effect is observed - the inability to learn or perform complex skills due to multiple dysfluent (but possibly accurate) prerequisites or components (Binder, 1996, p. 184). Since frequency measures can detect deficits in the fluency of the component skills, they allow instructors to intervene and correct those deficits before they affect future learning. The concept of cumulative dysfluency is intriguing, but this area would benefit from empirical research.

Proponents of fluency training also argue that the limitations produced by the dysfluent performance of component skills may render previously effective reinforcers ineffective (Binder, 1993, 1996; Haughton, 1971, 1972, 1980; Johnson & Layng, 1992). Binder asserted that "unlike what traditional or strict Skinnerians might predict, precision teachers have found that students' performances hit "ceilings" imposed by non-fluent prerequisites or component skills, and that reinforcement procedures did not enable students to break through those ceilings. Only additional practice on components, and attainment of higher rates, allowed students to "lift" the ceilings and achieve fluency on more advanced or composite
performances based on non-fluent components” (Binder, 1993, p. 9). Based upon their work, Binder (1993) and Johnson & Layng (1994) suggested that if a learner does not exhibit a composite behavior competently, instructors should look at the fluencies of the component behaviors before instituting reinforcement procedures for the composite behavior. If these observations were confirmed through empirical research, it would seem critical that we measure both accuracy and speed when teaching new behavior to ensure the efficacy of instructional interventions.

The achievement of fluent performance on component skills may also reduce the amount of instruction necessary to achieve competent performance of composite skills (Alessi, 1987; Binder, 1993; Epstein, 1991; Johnson & Layng, 1992, 1994). Johnson & Layng (1994, p. 181) stated that “when practitioners apply this technology (fluency training), new and complex repertoires emerge as a function of simply presenting a context for their combination, such as an activity, game, or simulation; component behaviors taught in the instructional sequence combine into new, untaught complex behaviors as a result of the requirements imposed by the game, activity, or simulation”. That is, more complex behaviors that are comprised of the component skills emerge without explicit instruction. If, in fact, fluency training of component skills can help the learner acquire more complex skills efficiently and competently, then we must measure speed as well as accuracy when teaching those skills.

Additional empirical research on the learning process may help clarify this important issue.

Proponents of the precision teaching model clearly argue that frequency
measures should be employed in education because accuracy measures alone are insufficient to accurately assess learning and alter instructional methods based upon those assessments. Although few researchers would argue against the precise measurement of the behavior being studied, these specific arguments are based upon certain assumptions about the process through which individuals learn simple and complex behaviors. The data available with respect to the acquisition and retention of component and composite skills will be reviewed later in this section.

The Precision Teaching Model

The precision teaching movement grew out of a commitment to using frequency as a universal measure of behavior, as well as the desire to employ research methods derived from the experimental analysis of behavior in education (Lindsley, 1991). Barlow, Hayes and Nelson (1984) maintain that precision teaching offers educators an alternative to traditional research methods and helps them become “scientists-practitioners” (cited in West and Young, 1992). An overview of the history and development of this model is beyond the scope of this review, but several are available (Eshleman, 1990; Lindsley, 1972, 1991, 1992; Potts, Eshleman, & Cooper, 1993; West, Young, and Spooner, 1990). This review will be restricted to a brief overview of the critical beliefs that underpin the model and a review of relevant research.

The precision teaching model of evaluation and decision-making in education is based upon several critical tenets derived from skinner’s experimental analysis of
behavior (Lindsley, 1971b, 1990; Potts et al., 1993; West & Young; West, Young, & Spooner, 1990; White, 1986). These tenets have been described variously, but may be summarized as follows:

1. the student knows best -- the effect of any instructional strategy must be evaluated based upon the effects of that procedure on the behavior of the learner;

2. an emphasis on the direct measurement of observable behavior and continuous monitoring (daily performance assessment);

3. the use of rate of response as a universal measure of behavior;

4. the use of a standard graphic format for the visual display of data that can be used to study behavioral patterns (the standard celeration chart);

5. the use of descriptive and functional definitions of behavior and processes;

6. ongoing analytical investigations of the effects of environmental influences (teaching tactics) on individual behavior (student learning); and

7. an emphasis on building appropriate and useful behavior, rather than focusing exclusively on eliminating undesired or inappropriate behavior.

The implementation of this model often involves frequent, brief, timed measurements of the frequency of a behavior that are plotted on a standard celeration chart. The slope of the line that best fits the data is then evaluated based upon decision rules to determine if the current instructional method should be continued or altered (White, 1985, 1986). Although it is not explicitly described in the model, implementation frequently also includes fluency building - training a specific behavior or behaviors until the learner meets specified criteria for both accuracy and
speed.

Although instructional approaches that employ the precision teaching model may involve a number of very specific procedures, the model itself is not a specific set of instructional techniques. Precision teaching does not prescribe any particular teaching methodology, but rather provides overall decision rules on when to continue with similar teaching or practice on a particular skill, when to try something different, and when to introduce teaching or practice of advanced skills—all based on rate data (Binder, 1996; Lindsley, 1971, 1990, 1992; Makepeace, 1998; West, Young, & Spooner, 1990; White 1985, 1986). Similarly, the use of this model does not require the instructor to abandon current teaching techniques, it simply provides a precise system for evaluating the effects of instructional strategies on the behavior of the learner (Potts et al., 1993; Lindsley, 1991, 1992). The precision teaching model has reportedly been employed successfully with a variety of instructional techniques (Potts et al., 1983).

Despite widespread discussions about the positive impact of the precision teaching model on learning behavior, there are few published empirical data to support these claims. Berquam described the lack of published data to support precision teaching techniques, noting that much of the data come from doctoral dissertations, university research studies, or “internal” reports. While results of many of these studies are common knowledge to persons closely involved in precision teaching, they hardly comprise what one might call a solid “professional literature” (Berquam, 1985, p. 326-327). Eshleman completed an analysis of the research
available with respect to precision teaching and concluded that precision teaching has become largely an “oral tradition” (Eshleman, 1990, p. 22).

Binder (1996, p. 164) identified several reasons for the lack of published empirical research: most early researchers were practitioners and did not pursue publication for career advancement; Lindsley, an early leader in the movement, discouraged early researchers from seeking to publish their data based upon his experience that publications did not change professional behavior sufficiently to be worth the effort required to do so; discoveries were occurring more rapidly in the field than the typical journal and book publication cycle could accommodate, thus discouraging people from publishing reports which would be obsolete by the time of publication; and finally, rather than rely upon mainstream behavioral journals, which binder suggested refused to publish data presented in standard behavior charts, these researchers developed an informal system for the dissemination of information through the “behavior bank.” Although most of the research in the area of fluency has not been collected through controlled scientific methods and reported in professional journals, binder asserts that much of the work represents huge numbers of replications across participants, and that many of the empirical generalizations derived by fluency researchers and practitioners over the last 30 years suggested opportunities for systematic research (Binder, 1996, p. 164). The following section provides an overview of research that reports the effects of the precision teaching model, and several teaching strategies that are often employed with it, on the behavior of a wide range of learners.
Early precision teaching efforts began with children. The precision teaching project in Great Falls, Montana is one of the most renowned examples of the effects of the precision teaching model. This project involved several studies with a range of participants. In one study, teachers added 20-30 minutes per day of timed practice, charting, and decision-making in basic skills to a conventional teaching program for first, second and third graders. Over three years, student scores on the Iowa test of basic skills rose between twenty and forty-one percentile points compared with other schools in the district that did not employ the precision teaching model (Binder, 1988, 1993; Beck & Clement, 1991; White, 1986). In a second study, sophomore and junior high school students who were performing at least three years behind grade level on a standardized test received instruction using either a precision teaching model or conventional approaches. Researchers reported that the precision teaching groups demonstrated significantly superior performance in each of the four parts of the standardized test (White, 1986). Although these results are impressive, neither study controlled for the possible effects of increased exposure to the learning materials (number of practice trials) associated with the precision teaching techniques.

The Morningside Academy and Malcolm X College also demonstrated many successes with the implementation of precision teaching methods. These programs emphasize the basics: mathematics concepts, calculation and problem-solving, reading, decoding and comprehension; language arts including handwriting, composing, grammar, spelling and mechanics; computer basics, including keyboarding, document organization, and word processing; time management;
materials organization and learning; critical thinking, reasoning, and argumentation (Johnson & Layng, 1996, p. 174). Educators focus on the attainment of fluent performance on each step in the curriculum. Children diagnosed as learning disabled, who have never gained more than half a year in any one academic year, typically gain between two and three years in each academic skill per year. Adults below the U.S. Government-defined eighth grade literacy level advance at the rate of two academic years for every 20 hours of instruction in each skill (Johnson & Layng, 1992, p. 1482; 1994).

The Morningside Academy demonstrated similar successes through their adult literacy program in reading, writing and mathematics for agencies eligible for federal monies dispersed by the job training and partnership act (JTPA, 1985). Thirty-two at-risk young men ranging from 16 to 26 years old participated in the first project. Participants entered with skills ranging from second to eighth grade as measured on the MAT6 (Prescott, Barlow, Hogan, & Farr, 1986). Twenty-nine of the 32 students completed the program and gained an average of 1.7 grades per 20 hours of instruction (Johnson & Layng, 1992, 1994). In a second project, 20 Asian-American women, aged 25-40 learned prerequisite skills in math, reading, spelling, and writing for office and computer-related skills training programs. Participants entered with skills between the fifth and ninth grades. Nineteen students completed the program, gaining an average of 2.1 grades per 19 hours of instruction (Johnson & Layng, 1992, 1994).

Spence and Hively (1993) reported success using elements of precision
teaching and direct instruction with learning disabled students at Ben Bronz Academy in Hartford, Connecticut. Twenty-one students whose achievement scores in reading averaged three and one half years below grade level were given rate-building exercises in reading at least three times daily, and had classes in decoding, linguistic-phonemic word study, and comprehension. All students’ reading scores (comprehension) rose an average of 2.5 years per year of instruction on the gates-MacGinitie reading test (1978) over a two-year period. Seven students’ scores also rose between one and three years above their current grade level. Fluency levels (correct words read per minute) were also assessed through sixth grade reading samples. The data indicated that all students exhibited increases in both fluency rates and comprehension (this measure was not clearly defined). Students who began the study reading at higher rates than other participants made the greatest gains in fluency rates. Students who read slowly prior to the study, however, achieved the greatest gains in comprehension. The former result is not surprising, but the latter may add to the support for this model.

Several other studies reported that the precision teaching model and related teaching techniques have been effective in enhancing the learning and performance of undergraduate students learning Psychological (McDade, 1992) and Pathophysiological (Olander, Collins, McArthur, Watts, and McDade, 1986) course content, bank employees learning complex financial products and services (Binder & Bloom, 1989), and employees in several business settings including a computer chip manufacturing company; a medical insurance company; General Electric Plastics; and
Ohio Bell learning job-related skills (Snyder, 1992). Although the results of these studies are suggestive, the lack of experimental controls employed prohibits empirical conclusions. In each study, the experimenters failed to either employ a control condition or group with which to compare the performance of participants exposed to precision teaching methods, or control for the effects of repeated practice independent of the specific teaching techniques employed.

The studies reviewed above provide a brief overview of the successes that have been reported using precision teaching model with a wide range of participants. Each of these studies involved a number of instructional strategies, including fluency-building, self-paced practice, brief timings, and other techniques that are commonly employed with the precision teaching model. They also examined the effect of those strategies on a variety of dependent variables using a wide range of research methods. However, in many cases it is difficult to determine what, if any, experimental controls were employed to ensure the integrity of the data. In particular, many researchers failed to control for the effects of practice alone. Finally, the lack of component analyses in these reports does not allow one to determine which components of the precision teaching model or teaching strategies contribute to the reported results. The remainder of this review will focus on fluency training as a strategy that is frequently employed in successful applications of the precision teaching model.
Binder (1988, p.12) defined fluency as “a combination of accuracy (or quality) plus speed which ensures that students will be able to perform easily in the presence of distraction, will be able to retain newly-acquired knowledge, and will be able to apply what they have learned to acquire new skills or to real life situations”. Johnson and Layng expanded this definition to include the selection of behavior by the environment. They noted that “the definition of fluency requires that the skills 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.” (Johnson & Layng, 1992, p. 1476). Fluency training typically involves exposing learners to the training materials until they have met a criterion for both accuracy and speed of responding. For example, learners may be required to practice completing simple addition problems until they are able to complete them both with 100% accuracy, and at a rate of 40 correct responses per minute. The benefits of fluency training have been widely described and will be reviewed in this section. Lindsley (1995) describes these benefits as the ten products of fluency, which he summarizes with the acronym “reaps fun cg”. Lindsley (1992, 1995) and others (Binder, 1993, 1996; Haughton, 1981b) suggested that fluency training enhances retention, endurance, application, performance standards, stability, fun, understanding, no cheating, confidence, and generativity. The following section will review the proclaimed benefits and empirical data related to them with respect to retention, endurance, stability (distractibility) and generativity.
Retention

Retention may be defined as the relation between behavior frequencies at two points in time, between which the individual has had no opportunity to emit the behavior (Binder, 1996, p. 164). The importance of adequate retention in the successful learning and maintenance of behavior has been widely documented (Binder, 1996; Johnson & Layng, 1992, 1994; Olander, et. Al., 1986). Proponents of fluency training maintain that teaching a behavior under conditions that require both accuracy and a high rate of responding can facilitate the rapid development of superior retention (Berquam, 1985; Johnson & Layng, 1994; McDade, 1992; West & Young, 1992). Although published empirical studies are limited, several studies that investigated the effects of different learning contingencies, including rate and accuracy-alone criteria during training on retention will be reviewed below.

One procedural variable that may affect the level of retention produced by training, independent of the imposition of a rate criterion, is the number of exposures to the relationship to be learned (number of practice trials). A thorough review of the research examining the effects of number of practice trials alone on future performance is beyond the scope of this paper, however Makepeace (1998) provides an extensive review of this literature. The current review will restrict itself to studies that examined the effects of an accuracy-alone criterion versus criteria that imposed both accuracy and a minimum rate requirements on behavior, but will be divided based upon whether the number of exposures to the learned relationships were equal under the two conditions. The first set of studies did not explicitly control the number
of exposures to the learned relationship, but rather allowed it to vary depending upon the formal criteria imposed (accuracy-alone or accuracy and rate). Studies that demonstrated positive findings with respect to fluency training will be summarized, however, studies that failed to control for the effects of number of practice trials and suggest that fluency training did not produce superior results will be reviewed in detail.

**Number of Exposures Uncontrolled:**

Ivarie (1986) exposed 120 fourth grade students categorized as above average, average, and below average with respect to math computation skills to two different fluency criteria during training. Students were randomly assigned to groups wherein a 70 responses per minute or 35 responses per minute criterion (with approximately 10% or fewer errors) was imposed. Students completed a paired-associates task that required them to “recall” and write a roman numeral when presented with an Arabic numeral until they met the identified criterion. The actual number of learning trials varied between conditions. Participants received four retention tests, one immediately after training, and one at thirty, sixty and ninety days after training. The results showed that regardless of previous achievement levels, participants in the 70 responses per minute group (high fluency) made significantly fewer errors than individuals in the 35 responses per minute group (low fluency) on retention tests. However, while both groups of participants demonstrated increased retention, participants exposed to the lower fluency criterion actually increased their retention.
ratio (mean performance rate on posttests two, three, and four divided by the mean performance rate on the posttest given immediately after training) significantly more than individuals in the 75 responses per minute group. This finding would suggest that the low-fluency training criteria actually produced greater gains in retention than the high-fluency criteria.

The increased posttest measures for both groups may be partially explained by the practice effects of taking each posttest, however, this does not explain the greater gains in retention obtained by the low-fluency group. Ivarie suggested that the superior performance of this group may be because the 35 responses per minute fluency rate was easier to attain and maintain, or because researchers used different bases to compute the retention ratios. As such, an increase in 10 on a base of 35 represents 28%, while increases in 10 on a base of 70 represents 14%. This would suggest that the apparent differential increases are due to data analysis procedures rather than performance. These results provide mixed support for the importance of a high rate requirement in producing superior retention.

Using a within-subject design, Kelly (1996) conducted two experiments evaluating the effects of fluency training versus training to accuracy-alone on the retention of sight-reading of Dolch words (the 220 most frequently found words in books that children read). Although the number of practice trials was not controlled during the first experiment, it was during the second, and therefore both studies will be described. Three five-year-old children diagnosed as retarded, learning disabled, and speech and language impaired participated in the first study. Each participant was
exposed to each of the experimental conditions. During the accuracy-alone condition, participants received training until they achieved 90% accuracy for two consecutive sessions. The fluency condition required participants to achieve both 90% accuracy and a high rate of correct responding defined as two consecutive sessions wherein each session lasted not more than 13 seconds (or an average of 1.3 seconds per trial). Participants were taught to read two sets of dolch words containing three words each while working one-on-one with the trainer. Retention tests were conducted periodically from one week to 194 days following the end of training, however, the number of retention tests was not consistent across conditions. Retention tests were similar to the training trials, however, the experimenter did not provide any feedback for correct or incorrect responses.

Participants averaged 100% accuracy on retention tests for sets of Dolch words learned during the fluency training condition, and less than 50% accuracy for sets of words learned during the accuracy-only condition. In addition, the average rate of correct responding on retention tests was always higher for sets of words trained to meet the fluency criteria than for words trained during the accuracy-alone condition. These results would seem to support the notion that fluency training produces better retention, however, the experimenter did not control for the amount of exposure to the learning relationships. With the exception of one set of Dolch words for one participant, the fluency condition involved a greater number of training trials than the accuracy-alone condition. We cannot, therefore, determine which variable may have produced the observed results.
In her second experiment, Kelly (1996) attempted to assess whether the additional instructional time that participants received during fluency training was a critical factor in the results of her first experiment. This study replicated the first experiment, however, the procedure was altered to ensure that the accuracy-alone condition, the condition expected to have the least effect on the dependent variable, involved more instructional time than the fluency condition. A "yoked" accuracy-alone control condition was introduced which extended the amount of time that the instructor spent between trials such that the duration of the control condition session was twice as long as the previous fluency condition session. For example, if it took a participant 15 seconds to complete ten trials during the fluency condition, the experimenter assured that the next yoked condition took 30 seconds by "increasing the student's inter-response time" (Kelly, 1996, p. 90). Retention tests were completed periodically from five to 63 days after training.

The data indicated that fluency training significantly affected accuracy (percent correct) on retention tests. Both participants achieved between 98% and 100% accuracy on sets of words trained under the fluency conditions, but only between 0% to 36% correct on words taught during the yoked control condition. The rate of correct responding on retention tests also varied depending upon the training condition. Rates of correct responding were 50% lower on word sets learned during the yoked control condition than those learned during the fluency condition. These results appear to lend support to the contention that the imposition of a rate requirement significantly enhances retention independent of the number of practice
trials, however, two procedural variables complicate a meaningful interpretation of these data.

One factor that may explain the results of the first experiment is the number of exposures to the learning relationship, that is the number of practice trials. In an attempt to control for possible effects during the second study, Kelly controlled the amount of time the participant spent with the experimenter during training. She did not increase the number of practice trials during the yoked condition, but rather increased the student’s inter-response time. Participants still received a greater number of training trials during the fluency condition. These data may suggest that the amount of time spent during training is not a functional variable in retention, but do not address the issue of number of practice trials. Kelly also included the experimenter-imposed inter-response time delays in her calculations of the participants’ response rates during the yoked-control condition, thus artificially reducing those rates. Taken together, these studies suggest that a rate requirement may be important during training, but do not provide concrete support for that conclusion.

Grindle (1996) evaluated the effects of training to three different terminal fluency rates on fluency rates (average number of correct responses per minute) on retention tests. Although she did not control for the number of practice trials across conditions, the contradictory nature of her results warrants description. Thirty college students completed part of an OSHA (Occupational Safety and Health Act) computer-based training program comprised of multiple choice, fill-in-the-blank, and true/false
questions. Three groups of participants were exposed to a fluency conditions requiring 95% accuracy and 34-39, 26-31, or 18-23 correct responses per minute. The number of practice trials varied between the experimental conditions, with the greatest number being associated with the condition requiring the highest rate of correct responding. Retention tests wherein students were required to complete the module two times were administered at intervals from two to 21 weeks after training.

The results showed that accuracy remained high for all three groups on all retention tests, with no significant differences between them. These findings do not support the contention that requiring a high rate of correct responding will produce superior retention, however, the differences between the rate requirements for each condition were relatively small, and may not have been functionally different for the participants. Differences in fluency rates were, however, observed. Participants exposed to the higher fluency rate criteria during training emitted a greater number of correct responses per minute during training. Evaluation of the data from retention testing, however, revealed that fluency rates increased over time for all groups. That is, the level of fluency achieved during training did not differentially affect how quickly fluency was lost after training.

In addition to the retention tests described above, participants completed retention tests comprised of novel presentations, or combinations of the material used during training. Neither the accuracy nor the fluency rates achieved on these tests were differentially affected by the criterion imposed during training. Overall, Grindle’s (1996) results suggest that the imposition of criteria for both accuracy and a
high-rate during training does not produce significantly better retention than a lower rate criterion. This study did not, however, compare these conditions to one that only required accuracy.

Bucklin (1998) examined the effects of fluency training on retention and application using a between-groups design. Thirty adult college students completed stimulus equivalence tasks that involved pairing nonsense syllables with both Hebrew characters and Arabic numbers as separate relations. All three groups received training on these tasks until they achieved 100% accuracy. Two fluency groups then received additional training on the task until they achieved 100% accuracy with a rate of either 50 or 100 correct responses per minute. The number of practice trials varied with the rate requirements of the condition, with the greatest number of trials associated with the 100 correct responses per minute criterion. Retention tests on a composite task, addition problems comprised of Hebrew letters, were completed immediately after the completion of training, and at two or four-week intervals up to sixteen weeks. A final retention test completed sixteen weeks after training examined the effects of the three training conditions on the retention of the component skills. Participants completed the stimulus equivalence worksheets for the original training relations: nonsense syllables and Hebrew characters, and nonsense syllables and Arabic numbers, in two one-minute timings.

The data suggested that training two component skills to meet both accuracy and rate criteria increased immediate post-training performance on the composite task. Participants exposed to conditions with both requirements completed
significantly more correct items per minute on retention tests completed immediately after training, and their accuracy was as high as participants trained to be accurate only. Participants trained under fluency conditions also completed more items correctly per minute on both component and composite tasks, and demonstrated greater accuracy on the composite skill on the retention tests. Moreover, they lost less accuracy on both the composite and component skills four and sixteen weeks after training. These results differ from those obtained by Grindle (1996), who found that participants from both high-fluency and low-fluency groups performed comparably in terms of accuracy on retention tests. The differential effects of training to the two different levels of fluency (100% accuracy + 50 or 100 correct responses per minute) were not presented separately or evaluated in this study.

As an extension of Grindle’s (1996) work, this study also examined whether the effects of the different training requirements on retention would vary depending upon the interval of time between training and retention testing. The type of training did not influence the difference between immediately post-training fluency rates and fluency rates on retention tests administered four and/or sixteen weeks after training for either component or composite skills. That is, training to fluency did not influence how quickly fluency deteriorated on successive retention tests. These results are consistent with those reported by Grindle (1996) who found that fluency rates increased over time for both the high-fluency and low-fluency groups. Ivarie (1986) found that although fluency conditions requiring both a high and low rate of correct responding produced greater retention than the accuracy-alone condition,
participants exposed to the low-fluency condition demonstrated the greatest gains in fluency. These data do not support the theoretical claim that training behavior until a high rate of correct responding is achieved produces retention beyond that which can be produced by accuracy training alone or a low rate requirement. Bucklin (1998), Grindle (1996), and Ivarie (1986) acknowledge that procedural and statistical variables, rather than the training conditions, may have produced the contradictory results.

Orgel (1984; cited in Binder, 1993 and Binder, 1996) reported that fluency training enhanced students’ performance in a university calculus class. White (1985) also reported a case study wherein fluency training was effective in training a learning disabled student to pronounce short vowels correctly. Each of these studies, however, lacked adequate experimental controls to conclude that the training alone produced the observed results. For example, they did not control for the number of exposures to the learning relationship, exposure to the training materials outside of the experimental setting, or, in Orgel’s experiment, individual differences.

Taken together, the studies described above suggest that the imposition of criterion for both accuracy and rate may enhance learner’s performance on retention probes. However, the number of exposures to the learned relationship varied greatly between conditions and participants, with the greatest number of trials presented during the fluency condition requiring the highest rate of correct responding. Additional empirical research that differentiates between the effects of the criteria imposed and the effects of number of exposures to the learned relationship may
produce a more precise understanding. The following studies explicitly controlled the number of exposures to the learning relationships such that the number of exposures was equal across conditions, or the condition that was believed to have the least effect on the dependent variable involved a larger number of learning trials.

**Number of Exposures Controlled:**

Berquam (1981) employed a between-groups design to examine the effects of un-paced practice and fluency training on the performance of fourth grade students on a paired-associates task. Thirty-four participants learned to match a set of ten pairs of items (a nonsense syllable paired with a single digit number) until all of the students achieved 100% accuracy on all ten pairs (phase one). Participants were then presented with five additional practice periods (phase two). The control group engaged in un-paced practice, with no time limits, while the experimental group completed one-minute timed trials (probes) daily that increased correct frequencies. Participants in the control group spent almost 70% more time in un-paced practice and completed about 60% more practice items than the other group. Retention probes were completed immediately after practice sessions and at two weeks after training. During the third phase of the study, both groups received one-minute timed probes on a daily basis that emphasized frequency followed by retention probes immediately after practice sessions and at two weeks after training.

The data indicated that the group that received one-minute timed practice during the second phase demonstrated greater accuracy and speed during all retention
probes. The differences between the effects of un-paced practice and fluency training were significant for each of the four retention probes except the probe completed two weeks after the end of the second phase. The experimental group demonstrated better retention despite the fact that these participants received fewer opportunities to emit correct responses. These results lend support to the contention that training conditions that require the learner to meet criteria for both accuracy and rate produce greater retention than accuracy-alone criterion independent of the number of practice trials.

Shirley and Pennypacker (1994) examined the effects of fluency and accuracy-alone training on learning efficiency and retention of spelling words using a within-subject design. Two eighth-grade boys with learning disabilities learned lists of ten spelling words under 3 different conditions. During baseline, participants learned either one or two lists of ten words each week and received new word lists each week regardless of their performance. During the accuracy-alone condition, participants received two word lists and were required to practice both lists until they achieved 100% accuracy on one list. The second list was “yoked” to the first list such that participants were required to continue to practice that list until they met the accuracy criterion on the first list. During the yoked fluency condition, participants received two lists and were required to practice both lists until they met a 100% criterion for both lists and an additional fluency criterion for one list. The number of practice trials was the same for both lists under both experimental conditions. Each participant received retention probes ten days after the final performance session for
each list.

The data showed that a 100% accuracy criterion increased both correct frequencies and accuracy ratios (total number of correct letters / total number of incorrect letters) slightly above baseline rates, while implementation of the fluency criterion almost doubled the correct frequencies for both participants. Retention data for the experimental conditions were inconclusive. Between-phase data suggested fairly stable performance on the follow-up sessions throughout the study, with one participant demonstrating stable accuracy throughout study, and the other showing progressive declines over time. Within-phase data, however, suggested that the accuracy-alone and fluency conditions produced higher correct frequencies than the yoked practice conditions, and that the fluency condition produced higher accuracy ratios. Within-phase data also suggested that the fluency criteria produced better retention than accuracy-alone criterion. These conclusions are not, however, supported by a comparison of the between-phase data for these conditions. Given the conflicting nature of the between- and within-phase data, this study offers mixed support for the contention that fluency training produces superior retention.

Makepeace (1998) examined the effects of number of training trials vs. Fluency training on retention, generalization, and generativity using a within-subject design. In a series of three experiments, the researcher exposed 18 adult participants to three different training conditions: an accuracy-only condition which “encouraged” a low rate of responding; a rate-criterion-alone condition which “encouraged” a high rate of responding; and a condition which required both accuracy and a specific rate
of responding. The number of training trials was equated for the rate-alone and accuracy and rate conditions. The main difference between the three experiments was how participants switched from condition to condition.

During training in all three experiments, participants were required to "match" a particular response square with a specific nonsense syllable (e.g., "hef") by clicking a mouse on one of a series of squares on a computer screen when presented with a nonsense syllable. Participants received immediate feedback for correct and incorrect responses, and received information relevant to the experimental condition between groups of trials. Training continued until participants met the criteria for each condition. After training, each participant was exposed to a "distraction phase" for 30 minutes outside of the experimental room. Following the distraction phase, each participant received three different tests. No feedback for correct or incorrect responses was given during testing, and the primary dependent variable was accuracy.

During the "plenary" test, participants were required to click the correct response-square when presented with one of the 24 nonsense syllable used during training. During the "moving reference point test", which Makepeace described as a test of stimulus and response generalization, participants were presented with a grid that contained many more response squares than presented during training. The position of the darkened square was varied within the grid during test trials. When presented with the 24 nonsense syllables from training, participants were required to click on the response-square in the position relative to the darkened square that the correct square held during training. The final test, the "concatenation test", was
designed to test for co-adduction (generativity). This test employed the same response grid used during the moving reference point test, but the darkened square was returned to its central, fixed position. Participants were then presented with pairs of nonsense syllables from training separated by a “+” sign (e.g., “usk + fap”) and required to click on the response square that represented the final position produced by chaining the two movements. For example, if the correct response square for “usk” was two to the left of the darkened square, and the correct response square for “fap” was one square to the left of the darkened square, the correct response to “usk + fap” would be three squares to the left of the darkened square.

The data from all three experiments suggested that the condition that imposed an accuracy-alone criterion and encouraged a low rate of responding produced the lowest accuracy during retention (plenary) testing, but the data for the other two conditions were inconclusive. Results of the moving reference point tests demonstrated a statistically significant difference between the effects of the condition that imposed accuracy and rate criteria and the accuracy-alone condition on accuracy. Similarly, results of the concatenation tests revealed a statistically significant difference between the effects of the rate-alone condition and the accuracy-alone condition, with the rate-alone condition producing superior accuracy. The differences between the effects of the condition which required accuracy and encouraged a high rate of responding and the condition which encouraged a high rate of responding without an accuracy criterion were not, however, significant. Makepeace concluded that the condition which required the greater number of responses during training
virtually always led to improved performance in tests for generalized and co-adduced behavior based upon those responses (Makepeace, 1998, p.185).

The relative superiority of the rate-alone and rate and accuracy conditions varied between experiments one and two, but the results could not be validated or invalidated statistically due to the small n. Statistical analyses of experiment three revealed no statistically significant differences between the effects produced by the two conditions where the number of practice trials were equated (rate-alone and rate and accuracy conditions). Other than the one described above, there were no significant differences between the effects of the low- and high-rate conditions on accuracy.

Although these results support the conclusion that repeated practice during training may positively affect retention, generalization, and co-adduction (generativity), the data do not support the contention that it is necessary to train behavior until a high rate of responding is achieved to produce such results (Makepeace, 1998). Any conclusions about the differences between the rate-alone and the rate and accuracy conditions must be interpreted cautiously. The experimenter reported that due to the low rate of responding observed during the first experiment, participants were encouraged to be accurate during early training runs of the rate-alone conditions in experiments two and three. These experimental contingencies may have minimized the functional differences between these conditions for the participants.

Taken together, these studies present a somewhat mixed picture of the effect
of the imposition of criteria for both rate and accuracy during training on the retention of behavior. Additional research is necessary to further distinguish between the effects of fluency training per se, and the number of exposures to the learning stimuli.

**Fluency Training and the Length of the Retention Interval:**

In each of the studies that did not control for the number of practice trials described above, the fluency conditions involved many more exposures to the learning relationships, and presumably more time, than the accuracy-alone conditions. If the results of these studies are consistent with applications in education and business, fluency training as an educational strategy can be expected to involve greater time, and therefore greater cost to implement. Given the potential increased cost related to high-rate fluency training, we must consider the following question: if training a response until it is accurate at a relatively low rate produces the same level of retention as that produced when the response is trained until it is accurate at a high rate, under what circumstances might it still be important to achieve a high rate of correct responding during training?

To some extent, the answer to this question depends upon the nature of the response that is trained. If the response in question must be performed with absolute accuracy and rapidly to be effective, for example the use of a defibrillator to resuscitate a heart attack victim, one might argue that it is critical that initial training continue until a high rate of correct responding is achieved. If, however, a machine operator uses a stamping machine approximately once every six months and has a
written outline to follow when he does so. Such intensive initial training may be unnecessary and costly. If one determines that a high rate of correct responding is necessary for a response to be effective in the environment, two relevant questions may be: how long after training can we expect the positive effects of high-rate fluency training to last?; And is there a point in time after training when the positive effects of achieving accurate responding at a high rate during initial training are lost?

Only two studies comparing fluency training with other techniques explicitly evaluated the relationship between type of training, and the length of the retention interval, and retention (Bucklin, 1998; Grindle, 1996). As was noted above, both of these researchers found that the rate of correct responding during training did not differentially affect how quickly fluency deteriorated on successive retention tests. Grindle (1996) found that accuracy remained high, and fluency rates increased across retention tests (two to 21 weeks after training) for groups exposed to high, medium, and low rate criterion during training. Similarly, Bucklin (1998) found that the rate of responding required during training did not influence how quickly fluency rates declined on retention tests completed four and/or sixteen weeks after training.

In an extensive review and statistical analysis of the overlearning literature, Driskell, Willis, & Cooper (1992) reported a statistically significant relationship between the amount of overlearning, the length of the retention interval, and the degree of retention for certain tasks. These researchers distinguished between two types of tasks, cognitive (verbal) and physical (motor skills). Their analysis revealed a positive relationship between the length of the retention interval and the
performance of physical tasks. That is, as the interval between the practice sessions and subsequent performance increased for physical tasks, retention was enhanced. The authors suggest that this unusual finding may be attributed to practice that may have occurred during the interval. For cognitive tasks, however, a negative relationship emerged: the effects of overlearning on retention of cognitive tasks declined as the retention interval increased. Their analysis also suggested that while the strongest effect of overlearning was observed immediately after training, the positive effects of overlearning decreased by one half when the retention interval increased to 19 days, and the initial benefit of overlearning may be expected to be lost when the retention interval exceeds 38 days (Driskell, et al., 1992, p. 620).

Additional empirical research that examines the relationship between the type of training, the length of the retention interval and retention may prove to be fruitful in determining whether fluency training is cost-efficient. That is, are the effects of fluency training on retention robust and long lasting enough to justify the additional time and cost necessary to complete such training?

Endurance / Distractibility:

In the precision teaching literature, the term "endurance" seems to describe two different phenomena: sustaining the same level of performance for extended periods of time ("endurance"); and sustaining a stable level of performance in the face of distracting stimuli ("distractibility") (Binder, 1996). Binder (1996, p. 178) defined endurance as "the ability to continue performing over increasing durations and in the
face of environmental distractions. Others (Binder, Haughton, & Van Eyk, 1990) suggest that what has traditionally been described as attention span may also be viewed as endurance. More recently, Lindsley (1995) described “distractibility” as stability. This section will provide an overview of both endurance and distractibility.

Proponents of fluency training as an educational tool suggest that once individuals have attained fluent performance, they can remain on task, or endure, for sufficient periods of time to meet real-world requirements, even in the face of distraction (Binder, 1984, 1996; Binder, Haughton, & Van Eyk, 1990; Cohen, Gentry, Hulten, & Martin, 1972). Binder et al. (1990) asserted further that, in general, until students attain a certain level of speed and accuracy on individual curriculum tasks, they typically lack the ability to maintain steady performance levels for extended periods of time without slowing down or stopping. Furthermore, requiring performance that is accurate, but non-fluent for long periods of time may result in increased error rates, negative emotional behaviors, and depressed learning. In contrast, Johnson & Layng, (1994) have found that sometimes reaching a higher fluency aim on component skills produces endurance, while at other times learners need to practice the skill over the period of time that will be required in the environment.

Binder et al. (1990) also suggested that many of the “off-task” behaviors and “behavior problems” which are often exhibited by children who have been labeled as having attention deficit disorder or as “learning disabled” might be successfully addressed by examining the fluency and endurance levels of the component skills.
rather than teaching on-task behavior. These researchers found that when students diagnosed with attention deficit disorder were given extensive endurance training on a variety of tasks, they were able to greatly increase their “attention span”. Similarly, Binder (1996) reported that when the time that they were required to perform a task was decreased, disabled students exhibited less off-task and disruptive behavior, and more stability in their performance. The Morningside Academy has also helped individuals diagnosed with attention deficit disorder increase their time on-task from their typical of one to three-minute spans, to 20 minutes or more by building fluency on component skills (Johnson & Layng, 1994).

Although fluency training on component skills may enhance “attention span” on those skills and composite skills comprised of them, there are no published empirical studies that explicitly examined the effects of imposing criteria for accuracy and rate during training on endurance. Binder et. al (1990) present a series of unpublished studies completed at the behavior prosthesis lab at Fernald State School to support their assertions about endurance. These studies involved: developmentally disabled children counting objects and matching numbers; children writing the numbers 0 through 9; and children completing pre-multiplication skills (writing the digits from 0 to 20, by twos). The results of these studies suggested that if a behavior must be exhibited for long periods of time in the environment, one should ensure that the behavior can be emitted at high fluency rates for brief periods of time during training before lengthening the required interval for responding. Each of these experiments, however, failed to control for individual differences in performance that
existed prior to, or developed independently of, the independent variable studied. It is, therefore, impossible to form any concrete conclusions about the impact of fluency training on endurance without additional research conducted with adequate experimental controls. The remainder of this section will review research on stability, or distractibility – the sustenance of stable level of performance in the face of distracting stimuli. Research in the precision teaching literature is limited to a series of published and unpublished studies completed by Carl Binder.

Binder (1979a) conducted a pilot study based upon a free-operant analogue of automaticity experiments conducted by cognitive psychologists (Laberge & Samuels, 1974). Two adult participants completed five different tasks in successive three-minute intervals including: reading numbers; saying answers to simple addition problems (sums to 18); reading printed anglicized names of Hebrew characters; saying numbers when presented with the names of the Hebrew characters (a stimulus equivalence task previously learned through a paired-associates procedure); and “adding” Hebrew characters (an example of generativity). During training trials participants performed all tasks by reading aloud from practice sheets into a microphone. The number of practice trials varied for each participant. A final probe, which followed a brief break, required participants to repeat the same tasks while wearing headphones through which an individual read random numbers for 30 seconds halfway through each session.

Both participants performed equally well on the simple reading and adding skills, however, the participant who completed more practice trials performed at a
higher correct frequency on the paired-associates task (saying numbers when presented with the names of the Hebrew characters) and the composite task ("adding" Hebrew names). Suppression ratios were calculated by dividing the performance frequencies while the distracting stimulus was on by frequencies averaged for the period before and after the stimulus was present. The data suggested that although there was little disruption in performance of either participant on the simple tasks (reading and adding), the participant who performed the new paired-associates task and component tasks at lower frequencies experienced significant disruption in performance during the distracting stimulus. Although binder does not describe a specific rate criterion during the 3-minute timings, one assumes that participants were asked to answer quickly. The results of this study may suggest that if an individual is able to perform a task correctly and at a high rate, it is probable that the performance of that task will be sustained in the presence of distracting stimuli. They do not, however, provide compelling evidence that the imposition of a rate criterion during training is critical to establishing performance that will be resistant to distractions. Although binder controlled for the amount of time that each participant was exposed to training, he did not control for the number of training trials received or individual differences which may have existed prior to the introduction of the independent variable.

In a related study, Binder (1987) examined the effects of terminal training response rate on fluency, endurance, and distractibility. Using a paired-associates task, children learned to say specific numbers when presented with Hebrew letters.
until they met a criterion of 100% accuracy. The number of practice trials was uncontrolled and varied by student. Participants were then asked to "add" Hebrew characters while listening to a tape recording of voice reading random numbers. Participants who had a low rate of saying numbers when presented with Hebrew letters exhibited significant disruptions in performance while the random numbers were presented, while participants who were able to complete the original learning task fluently performed the adding task at a consistent pace. Again, although these results may suggest that accurate and high-rate performance is more resistant to distractions, the lack of controls for individual differences which are unrelated to the experimental manipulations, the number of exposures to the learning relationship, and the lack of a control condition or group, make it impossible to form any concrete conclusions about the effects of fluency training on distractibility.

Finally, Binder (1995) described an unpublished study that evaluated the effects of fluency training on distractibility during a paired-associates math task. Participants were taught to say numbers when they saw Hebrew characters using a paired-associates procedure. A probe for "emergent" behavior demonstrated that participants were able to add Hebrew characters without explicit training. All participants then completed fluency building exercises on the "see Hebrew character, say number" relationship. It is not clear how many practice trials were completed by any of the participants. In a final probe, participants were asked to complete the prerequisite skills (read numbers and Hebrew characters and answer simple math problems), the training task (saying numbers when presented with Hebrew
characters), and the final addition task ("adding" Hebrew characters) while listening to a tape recording of a voice reading random numbers. The data suggested that on both the paired-associates training task and the final addition task, lower performance rates were associated with greater distractibility. These results, however, must be considered inconclusive given the lack of experimental controls employed.

Taken together, Binder's (1979a, 1987, 1995) studies are intriguing. These studies suggest that behavior that is trained to the point of being accurate and occurring at high rates may be more likely to remain stable in the face of distracting stimuli in the environment. The lack of experimental controls, however, makes it impossible to identify any specific causal relationships. The observed differences in performance may be a function of many variables including response rate, number of exposures to the learning relationships (practice trials), procedural peculiarities, individual differences, or the tasks themselves.

Several studies found in the overlearning literature demonstrate that repeated practice may decrease the disruption in performance produced by the presentation of distracting stimuli (Peterson & Peterson, 1959; Richardson (1956); Slamecka, 1959). Although these researchers did not explicitly manipulate response rate, the data may allow us to speculate on the effect of rate. Since the imposition of a rate criterion in training typically involves a greater number of learning trials to meet criterion, one would assume that these results would be replicated under those conditions.

Overall, the empirical data that support the contention that the imposition of criteria that require a high rate of correct responding during training will produce
performance that is sustained for long periods of time and in the face of distracting stimuli are severely limited. Additional empirical research in this area may help determine the true effect of the imposition of rate and accuracy criteria on endurance and stability.

**Generativity:**

Component skills that are initially shaped under separate circumstances may be recruited in a substantially different context into a new composite skill (Johnson & Layng, 1994, p.182). For example, students may be taught to recognize and define behavioral terms, and to diagram simple four-term contingencies on a computer such that they are able to complete both skills accurately and at a high rate. If the students are then able to diagram behavioral contingencies with the correct terms when presented with story problems on a paper and pencil test, without any explicit training, one might suggest that the two component skills have combined to comprise a new composite skill. This sudden combination of component behaviors into a new composite skill, without explicit training, has been called contingency adduction (Andronis, 1983; Layng & Andronis, 1984; Johnson & Layng, 1992), generativity (Binder, 1993, 1996; Lindsley, 1995), and application (Binder, 1996).

Proponents of fluency training as an educational tool suggest that students who exhibit fluent performance on component skills can apply, adapt, or combine what they have learned in new situations, in some cases without explicit instruction (Binder, 1979b, 1993; 1996; Binder & Bloom, 1989; Haughton, 1972; Johnson &
Layng, 1992, 1994, 1996). Johnson & Layng (1996) report that at the Morningside Academy they have found that “skills making up as much as one third of a course of instruction may emerge ‘for free’ along the way, as the component skills that make up the emergent skill are mastered” (Morningside Academy, 1996, p.3). Further, they contend that “generative instruction emphasizes making new or latent repertoires available to the environment, so that new contingencies can select solutions and curriculum leaps that have been adduced from former related and unrelated component performances, rather than explicitly trained, sequenced, or chained” (Johnson & Layng, 1992, p. 1476). This process may increase the probability that creative and problem-solving repertoires will develop, and may emerge without training (Binder, 1996; Johnson & Layng, 1994, 1996; Haughton, 1972). If these assertions are supported by empirical data, fluency training may introduce significant efficiencies into the educational process. If through efficient teaching practices we are able to reduce the explicit instruction required for individual concepts, educators will be freed to provide students with greater and more diverse educational opportunities. This section will review the empirical research related to these assertions.

Johnson and Layng (1992, 1994) provided an example of generativity with respect to solving word-problems involving fractions. Four students in the Summer Program at Malcolm X College in Chicago were given a pre-test in mathematics composed of word problems involving fractions. Students’ scores ranged from 3 to 7 correct out of 14. These students also demonstrated deficits in lower level whole
number word problems and computation of fractions. Participants received instruction and fluency-building exercises in whole number word-problems and fraction computation. After these skills were emitted at a high rate, the students were given another test of word-problems involving fractions. One of the four students scored 13 correct out of 14; the others scored 14 out of 14. These improvements occurred without direct instruction in solving word-problems involving fractions.

Binder (1979b) reported the appearance of direction-following behavior without specific training in four disabled adults. Participants completed paired-associates training wherein they learned to read printed words corresponding to words already in their speaking vocabularies until they met criteria for both 100% accuracy and a low-rate (no more than twelve words per minute). Participants could already name the actions and objects and follow spoken directions using the words. After training, all participants were able to match the printed words to objects and actions, and two were able to follow four-word written instructions. Each participant then completed daily frequency building exercises for oral reading. Researchers found that participants were able to match printed words with actions and objects at higher frequencies, and all four could follow written directions using the words. These results would seem to support the contention that more complex behaviors (direction following) may emerge without training when component skills are trained to meet fluency criteria. The lack of experimental controls, however, makes any conclusion tentative.

Evans, Mercer, and Evans (1983) studied the effects of three different fluency
criteria for rate of saying letter sounds on the later rate and acceleration of saying consonant-vowel--consonant trigrams using a between-groups design. Nine elementary school students diagnosed as "learning disabled" completed an un-timed pre-test on saying trigrams and two one-minute timings of the same task. Participants were then divided into three groups of three that practiced saying letter sounds during one-minute timed sessions until they reached a criteria of 80, 60, or 40 correct sounds per minute with five or less errors. Each group received 940 presentations of the letters. If participants reached the rate criteria in less than 940 trials, the experimenter altered the pace at which participants could respond to the letter presentations to ensure that rates did not exceed the criterion for a given group. During the final phase, each participant was given an un-timed test of saying trigrams followed by fifteen one-minute timings of saying trigrams.

The median correct scores of the two one-minute timings for saying trigrams at the beginning of the study and the median correct scores of the last three one-minute timings for saying trigrams in the final phase were compared. The group exposed to the highest rate requirement (80 correct responses per minute) showed the greatest gain in number of correct trigrams said, while the medium-rate-group (60 correct responses per minute) showed the least gain. The number of correct responses during the initial and final un-timed tests of saying trigrams was also compared. Again, the group exposed to the 80 correct letters per minute criteria showed the greatest increase in correct trigrams said, while the group exposed to the 40 correct letters criterion showed the least improvement. These results support the contention

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
that training conditions that require a high rate of correct responding on component skills may engender composite skills without explicit training. It is notable, however, that the group exposed to a criterion requiring 60 correct responses per minute (the medium group) showed the smallest gain in number of correct responses. If we assume that a higher correct rate requirement, independent of number of presentations, positively affects the rate and accuracy of performance, one would predict that this group’s accuracy would improve more than the group exposed to the lower correct rate requirement (40 responses per minute). It is not clear, however, how the much the experimenter’s control of the presentation of stimuli may have impacted the performance of the participants in each of the groups.

McDade, Rubenstein, and Olander (1983) examined the effects of fluency testing using either flash cards or computer-generated tests on the performance of college students on essay questions in an undergraduate psychology course. The overall course was structured using a Personalized Systems of Instruction (PSI) involving self-paced and module-by-module learning. Students were also required to complete fluency testing over the description or identification of basic concepts, terms, and definitions related to different personality theorists and then answer essay questions designed to “assess the student’s knowledge of more complex concepts” (McDade, et. Al., 1983, p. 2). Fluency evaluations continued until participants were able to achieve 80% correct at a minimum rate of ten correct concepts per minute.

Researchers evaluated the number of concepts written correctly on essay questions. The data revealed a consistent and concomitant increase in the frequency
of correct concepts in both frequency testing and on written essay questions. In general, as the frequency of correct concepts answered during fluency evaluation increased, the frequency of correct concepts in the answers to the essay questions increased. The frequency of words and correct concepts written per minute on essay questions also typically increased over the course. McDade et al. (1983) concluded that frequency testing on basic concepts that help the student become fluent and accurate in those basics facilitates students’ use of those concepts in essay questions. Although it was not explicitly stated, one assumes that the participants in this study were required to complete more practice trials than typically required in the course to meet the identified criterion. In the absence of a control group or condition that would equate the number of exposures to the instructional material, it is difficult to conclude that the achievement of high correct rates on concepts alone increased students’ performance on essay questions.

Grindle (1996) evaluated the effects of fluency training to three different terminal fluency rates on retention fluency rates for college students completing part of an OSHA computer-based training program. Three groups of participants were exposed to fluency conditions requiring 95% accuracy and 34-39, 26-31, or 18-23 correct responses per minute for three consecutive trials. The number of practice trials varied between the experimental conditions. Participants completed retention probes comprised of the component skills, novel presentations, and combinations of the material used during training tests at intervals from two to 21 weeks after the completion of training. Neither the accuracy nor the rates achieved on these tests
were differentially affected by the criteria imposed during training. That is, participants exposed to both the rate and accuracy criteria in training of component skills were no more likely than other participants to exhibit the composite behavior more accurately or rapidly. These results are not consistent with those described above (Binder, 1979b; Evans, Mercer, & Evans, 1983, McDade et al., 1983).

Bucklin (1998) examined the effects of fluency training on retention and generativity using stimulus equivalence tasks which involved pairing nonsense syllables with both Hebrew characters and Arabic numbers as separate relations. Three groups received training on these tasks until they achieved 100% accuracy. Two fluency groups then received additional training on the task until they achieved 100% accuracy and a rate of 50 or 100 correct responses per minute. The number of practice trials varied with the rate requirements of the condition. Retention tests of a composite skill, addition problems comprised of Hebrew letters, were completed immediately after the completion of training, and at two or four-week intervals up to sixteen weeks. The data suggested that training two component skills to fluency produced greater accuracy and correct responses per minute on the composite task both immediately post-training and at retention intervals up to 16 weeks. Moreover, these participants lost less accuracy on both the composite and component skills four and sixteen weeks after training. These results differ from those obtained by Grindle (1996), who found that participants from both fluency and accuracy-alone groups performed comparably in terms of accuracy on tests for generativity. The lack of control for the number of practice trials makes it impossible to determine which
variable may have produced the current results.

Makepeace (1998) examined the effects of number of practice trials vs. The imposition of a rate requirement on retention, generalization and generativity. Participants were exposed to three different training conditions involving an accuracy-only criterion that encouraged a low rate of responding, a rate criterion alone, and a condition that required both accuracy and a high rate of responding. The number of practice trials was equated for the rate-alone and accuracy and rate conditions. All participants completed training that required them to "match" a particular response square on a grid on a computer monitor with a specific nonsense syllable (e.g., "hef").

During the "concatenation test", a test for co-adduction (generativity), participants were presented with pairs of nonsense syllables from training separated by a "+" sign (e.g., "usk + fap") and required to click on the response square that represented the final position produced by chaining the two movements. For example, if the correct response square for "usk" was two to the left of the darkened square, and the correct response square for "fap" was one square to the left of the darkened square, the correct response to "usk + fap" would be three squares to the left of the darkened square. The results of this test revealed a statistically significant difference between the effects of the rate-alone condition and the accuracy-alone condition, with the rate-alone condition producing superior accuracy. Statistical analyses of the data did not, however, reveal statistically significant differences between the effects produced by the two conditions where the number of practice trials were equated (rate-alone and rate and accuracy conditions). The results of this
study provide some support for the contention that a greater number of practice trials in training may lead to better performance on tests for generativity, however, they do not provide evidence that the imposition of criteria for both rate and accuracy during training is necessary to produce superior generativity.

Young, West, Howard and Whitney (1986) taught two mentally retarded, handicapped preschool children three dressing skills using a multiple baseline design wherein initial teaching for each skill was introduced at different points in time. Preliminary training involved whole-task training and graduated guidance. After a dressing skill was occurring with consistent accuracy, participants received additional practice sessions until the they were able to complete these skills accurately and at a specific, relatively high rate. The data indicated that high-rate training was almost always associated with accurate dressing skills in situations where the behaviors did not occur before training (in different rooms, in the presence of different personnel, at home). Dressing skills were also maintained four weeks after formal training. These results support the conclusion that training a behavior to meet both a rate and accuracy criterion may produce behaviors that generalize to other environmental conditions without explicit training.

When taken as a group, the results of these studies provide mixed support for the contention that the imposition of criteria for both accuracy and a high rate of responding during training of component skills will increase the probability that more complex composite skills comprised of those components will be selected under unique environmental conditions without training. However, the lack of experimental
controls around individual differences and the number of practice trials, and/or the unique experimental parameters employed make it difficult to form concrete conclusions based upon these data. Additional empirical research will help determine if fluency training can provide educators with a tool that promotes more efficient learning, thus allowing them to provide students with opportunities for broader and more creative education.

Purpose of the Current Research

The purpose of the present research was to further evaluate the effects of training under an accuracy criterion alone versus criteria for both accuracy and the rate of correct responding on retention, generativity, and distractibility. It is hoped that the data obtained will contribute to the limited body of empirical evidence related to the effects of fluency training. In addition, this research provided additional information with respect to two of the procedural and experimental parameters employed in this area of research. Specifically, this research explored whether any differential effects observed were due to the imposition of a rate criterion or the number of exposures to the learning relationship (number of practice trials v. Fluency). It also examined whether any observed benefits of fluency training on retention and generativity persist across repeated measurements over time, or diminish as the time after initial training increases.

Specifically, these two studies examined:

1. The effects of training basic component skills under a criterion for
accuracy alone or criteria for both accuracy and the rate of correct responding on the retention of those skills (retention);

2. The effects of training basic component skills under a criterion for accuracy alone or criteria for both accuracy and the rate of correct responding on the acquisition and retention of more complex, composite tasks (generativity);

3. The effects of training basic component skills under a criterion for accuracy alone or criteria for both accuracy and the rate of correct responding on the performance of those skills in the presence of distracting stimuli.

Several different dependent variables were examined during these two experiments. Percent correct was employed as the measure of accuracy during the retention / generativity experiment. Due to technical limitations of the computer program, it was not possible to obtain absolute measurements of the rate of correct responding achieved during retention and generativity probes. Fortunately, because of the nature of the experimental task, the length of the latency observed had a direct, negative relationship to the rate of correct responding achieved. That is, decreases in the length of latencies indicated increases in the rate of correct responding, and increases in latencies indicated that the rate of correct responding decreased.

Latencies measured the time between when the trial was initiated (stimuli presented) and the participant’s response. Under the procedures during both training and the retention and generativity probes, the response was a single mouse click, a response
with essentially no duration. The next trial was presented immediately after the mouse click occurred, with no inter-trial-interval. Therefore, the only thing that varied with the rate of responding for a participant was the latency between the presentation of the trial and their response. As such, latency was employed as the measure of the rate of correct responding, or speed during this experiment.

During the distractibility experiment, percent correct was employed as the measure of accuracy and the rates of correct and incorrect responding achieved were employed as measures of speed. An additional measure, total score per minute, provided an overall assessment of performance in terms of both correct and incorrect responding. Additionally, unique dependent variables were examined in relation to the individual distracting tasks and will be described with those tasks.
CHAPTER II

EXPERIMENT I: EFFECTS OF FLUENCY TRAINING ON RETENTION AND GENERATIVITY

Participants

Participants were recruited through in-person solicitations and public notices posted at Living Ways, Inc. Twelve adult participants completed this experiment and 5 of them completed both experiments. Each participant was between 25 and 65 years old, was a high school graduate, and had taken at least some college courses. All participants could read, write, and manipulate the computer mouse.

Participants were offered $5.00 for each session that they completed, and a $15.00 bonus for completing all 7 sessions. All of the participants declined to be compensated. Participants completed an informed consent form prior to starting either experiment. A copy of the informed consent form is included in Appendix A, and the letter validating the approval of the human subjects institutional review board is included in Appendix B.

Setting

All sessions were conducted in two similar offices at the main office of Living Ways, Inc. at 626 Reed St., Kalamazoo, Michigan. Each office is approximately eight feet by ten feet and contains one large window. Participants were alone in the office during both training and probe sessions.
Apparatus

All sessions were conducted on a dell computer system equipped with a full-sized keyboard, mouse, and mouse pad. A custom software program was used.

Experimental Design

A single-subject, multi-element design was employed. Each of the twelve participants was exposed to the accuracy-alone condition as well as each of the two different fluency conditions. The experimental conditions were: accuracy-alone – 95% correct for three consecutive 2-minute sprints; high fluency – 95% correct and 30 correct responses per minute for three consecutive sprints; and low fluency -- 95% correct and 20 correct responses per minute for three consecutive sprints. In an effort to differentiate between the effects of the imposition of different performance criteria and the effects of the number of exposures to the learned relationship (number of practice trials), a yoked procedure was employed. During each of the accuracy-alone and low-fluency conditions presented after the high-fluency condition, participants were presented with the same number of sprints that they required to meet criterion during the high-fluency condition. Table 1 presents the number of sprints and practice trials completed for each participant for each experimental condition. The order of presentation of the conditions was also varied and balanced across participants (see Appendix C).
Table 1

Total Number of Sprints, Practice Trials, and Average Number of Practice Trials per Sprint for All Participants for All Conditions.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Total Number of Sprints</th>
<th>Total Number of Practice Trials</th>
<th>Average Number of Trials per Sprint</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy Alone</td>
<td>High-Fluency</td>
<td>Low-Fluency</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>33</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>26</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>132</td>
<td>153</td>
<td>126</td>
</tr>
<tr>
<td>Average</td>
<td>11</td>
<td>13</td>
<td>11</td>
</tr>
</tbody>
</table>

Procedure

The training phase for each experimental condition involved learning relationships between a set of Japanese characters and the Arabic numbers 1 - 9 through a paired-associates procedure. The experimenter established three distinct sets of characters from a group of Japanese characters balanced for difficulty. The character set associated with each experimental condition was varied and balanced across participants. The three Japanese character sets are presented in Appendix D.
The Arabic numbers were used during the pre-training, and character sets 2, 3, and 4 were used during formal training and probes.

Each participant completed a pre-training exercise during the first session. This exercise included a single presentation of each procedure (group presentation of relationships, individual presentations of relationships, single sprint, retention probe (six trials), and generativity probe (six trials). Following this exercise, the experimenter evaluated whether participants met the minimum requirements for participation in the study, and answered any questions. All participants met the minimum requirements.

Within each experimental condition, participants received matching-to-sample training and retention and generativity probes. Retention and generativity probes occurred immediately after each participant completed training on a given character set and at approximately seven, fourteen, and twenty-eight days after the training session for each character set. Appendix E presents an overview of the experimental procedures for each participant during each session.

Each of the three experimental conditions involved the following procedures:

**Training**

1. Relationships presented as a group: participants were presented with the Arabic numbers 1 through 9 and the Japanese characters associated with each number as a group (see figure 1). There was no time limit for the participants to study these relationships and they moved onto the next training phase by clicking the mouse.
Figure 1. Screen Template for Relationships Presented as a Group.

2. Relationship between a given number and character presented individually and with a matching-to-sample trial:

   A. Participants were presented with the numbers 1 through 9 individually, in random order, with the character associated with that number. Participants studied that relationship with no time limit and then clicked the mouse to move on to the match-to-sample part of the trial.

   B. Match-to-sample trial: participants were presented with the character involved in the relationship presented in the previous part of the trial surrounded by the numbers 1 through 9 in a circle (see figure 2). Each trial required that participants select the correct number by clicking on that
number with the mouse. To control for the possible effects of position, the numbers in the circle were presented in order from 1 to 9, but rotated one position clockwise after each trial. The arrow pointer with which participants select the correct number also returned to the center of the circle after each trial.

![Screen Template for Match-to-Sample Trials and Retention Probes.](image)

Figure 2. Screen Template for Match-to-Sample Trials and Retention Probes.

Immediately after an incorrect response occurred, the incorrect number was crossed out, the correct response was highlighted in blue for two seconds, and the same trial was presented again until the participant responded correctly. Correct responses resulted in the correct number being highlighted in blue for two seconds and the initiation of the next trial. There was no time limit for the participant to respond during a particular trial. The individual relationship phase continued until the participant responded correctly to each of the nine relationships presented in this fashion.
3. **Break:** participants were presented with a 60-second break that could be extended or shortened by clicking a button with the mouse. Due to concerns about the accuracy of the data, participants also recorded their correct responses per minute for each sprint on a form. This form was present during each condition.

4. **Matching-to-sample sprints:** participants completed a series of 2-minute “sprints” composed of matching-to-sample trials involving the Japanese characters and Arabic numerals presented in the previous part of the training phase. The procedure and feedback during this phase was the same as those employed during the matching-to-sample trials described above.

   Each 2-minute sprint was followed by a break during which participants were presented with a graph showing performance information relevant to that condition. For example, during the high-fluency condition, participants were presented with a graph showing the target criteria for accuracy and response rate, as well as the individual’s performance on each of those measures (see figure 3). During the accuracy-only condition, the graph presented accuracy information alone.

   Participants were able to extend or terminate the break at any time. Matching-to-sample sprints continued until the participant met the experimental criteria identified for that condition for three consecutive sprints. For example, during the accuracy-alone criterion condition, participants continued with sprints until they achieved 95% accuracy for three consecutive 2-minute sprints.
Retention Probes:

Retention probes were presented both within-session for relationships learned during that session, and between-sessions for the relationships learned in previous sessions. Probes were completed immediately after training and at approximately 7, 14, and 28 days after the training session for each condition. Each probe involved twenty-four trials wherein participants were presented with the characters learned during previous conditions presented individually, and the numbers 1 - 9 in a circle around the character (see figure 2). Participants were required to select the correct number using the mouse. No feedback was presented for correct or incorrect responses, and the next trial was presented immediately after the participant responded. Performance was evaluated in terms of both percent correct and average...
latencies.

Generativity Probes:

Generativity probes were presented both within-session for relationships learned during that session, and between-sessions for the relationships learned in previous sessions at the same intervals as retention probes. Participants were presented with a series of mathematical problems (addition, subtraction, multiplication, division) using the characters learned under previous conditions with a correct answer of 1 - 18. For example, participants might be presented with the following problem: “\( \frac{2}{3} \) + “福” = _____. Participants then selected the correct answer from the numerals 1 through 18 presented in a circle around the equation (see figure 4). No feedback was presented for correct or incorrect responses, and the next trial began immediately after the participant responded. Performance was evaluated in terms of both percent correct and average latencies.

Generativity probes involved twenty-four trials including six equations each of addition, subtraction, multiplication, and division. Each of the four generativity probes (immediate, 7-day, 14-day, and 28-day) contained a unique set of equations with the exception that the following division equations appeared in two different probes: 6/2; 8/2; 8/1; 5/1; 4/2; and 2/1. No equations were duplicated within the same probe. Appendix f presents the equation sets employed during all generativity probes.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Results / Discussion

Each participant completed both training and retention and generativity probes for each of the three experimental conditions. These conditions involved the following performance criteria: accuracy-alone - 95% accuracy for 3 consecutive 2-minute sprints; high-fluency - 95% accuracy and 30 correct responses per minute for 3 consecutive 2-minute sprints; and low-fluency - 95% accuracy and 20 correct responses per minute for 3 consecutive 2-minute sprints. Although participants were required to achieve a minimum rate of correct responding to meet the performance criteria for 6 minutes, the procedure did not prevent them from exceeding the required rate during those sprints, and many participants did so. The dependent variables examined during probes were percent correct and average latency. During each of the
accuracy-alone and low-fluency training sessions completed after the high-fluency condition, participants were presented with the same number of sprints that they required to meet criteria during the high-fluency condition.

An analysis of the average rate of correct responding achieved by each participant during the final 3 sprints of training under each experimental condition revealed that many participants achieved correct-rates that exceeded the performance criteria imposed by the experimenter. For example, participant nine achieved average correct response rates of 29.8 per minute under the accuracy-alone criterion, 31.8 per minute under the high-fluency criterion, and 34.4 per minute under the low-fluency condition. If the experimenter-imposed criteria alone determined performance, the low-fluency condition should have produced accurate responding at a low to moderate rate of speed. In actuality, however, the low-fluency condition was associated with the highest rate of correct responding, thus making it the functional "high-fluency" condition for this participant. Table 2 presents the average number of correct responses per minute (CPM) achieved for each participant under each of the three experimental conditions.

For a majority of the participants, the nominal high-fluency condition did not produce the highest rate of correct responding achieved during training. If the actual rate of correct responding is a critical determinant of subsequent performance, as is espoused by advocates of fluency training, an examination of the data based upon the experimenter-imposed, or nominal criteria alone may be inadequate. As such, an analysis was completed to examine the relationship between the nominal rate required
under the experimental conditions and the actual rate achieved during that condition. For purposes of evaluation, each experimental condition was designated as high, moderate or low based upon the rate of correct responding achieved under that condition relative to the other conditions for each participant. Three participants (2, 7 and 10) were excluded from the analysis because the rates of correct responding achieved under all three of the experimental conditions were essentially equal. For example, participant seven achieved rates of 35.9 correct per minute under the accuracy-alone and low-fluency conditions, and 35.5 per minute under the high-fluency condition. In terms of her performance in real time, this represents approximately one additional mouse click every two minutes. As such, these conditions were considered to be functionally equivalent.

For the remaining 9 participants, two conditions were considered to have produced equal achieved rates of correct responding if the difference between the conditions was 1.5 correct responses per minute or less. For three participants (3, 8, and 12) two experimental conditions produced equal correct rates during training. These conditions were considered either high or low, depending upon the condition that they were most similar to, and the participant was counted as not having a condition that produced a moderate achieved-rate of correct responding. Appendix g presents the participants’ designations into the high, moderate or low-achieved-rate groups. Although assigning participants to groups based upon absolute differences in performance would have been simpler, the current approach produced grouped data that represented the individual results more accurately. Table 3 presents the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
frequency with which the high-, moderate- and low-achieved rates were associated with the nominal required rates for the 9 participants evaluated.

Table 2

Average Number of Correct Responses per Minute for Each Participant Under all Three Experimental Conditions During the Final 3 Minutes of Training.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Accuracy-Only</th>
<th>Low-Fluency</th>
<th>High-Fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25.4</td>
<td>32.8</td>
<td>30.7</td>
</tr>
<tr>
<td>2</td>
<td>19.7</td>
<td>19.8</td>
<td>21.0</td>
</tr>
<tr>
<td>3</td>
<td>26.4</td>
<td>36.7</td>
<td>35.2</td>
</tr>
<tr>
<td>4</td>
<td>29.4</td>
<td>32.1</td>
<td>23.5</td>
</tr>
<tr>
<td>5</td>
<td>26.0</td>
<td>23.3</td>
<td>29.5</td>
</tr>
<tr>
<td>6</td>
<td>22.5</td>
<td>28.1</td>
<td>30.0</td>
</tr>
<tr>
<td>7</td>
<td>35.9</td>
<td>35.5</td>
<td>35.9</td>
</tr>
<tr>
<td>8</td>
<td>24.3</td>
<td>32.3</td>
<td>32.1</td>
</tr>
<tr>
<td>9</td>
<td>29.8</td>
<td>34.3</td>
<td>31.8</td>
</tr>
<tr>
<td>10</td>
<td>30.0</td>
<td>28.8</td>
<td>30.7</td>
</tr>
<tr>
<td>11</td>
<td>32.6</td>
<td>27.6</td>
<td>30.1</td>
</tr>
<tr>
<td>12</td>
<td>31.4</td>
<td>30.8</td>
<td>35.1</td>
</tr>
<tr>
<td>Average</td>
<td>27.8</td>
<td>30.2</td>
<td>30.5</td>
</tr>
</tbody>
</table>

When one examines the participants as a group, a few trends are apparent. Twenty-five percent of the total participants achieved similar rates of correct responding under all three experimental conditions, suggesting that there was no relationship between the nominal rate required and the rate achieved during training. The accuracy-alone condition was, however, associated with the lowest rate of correct responding for 6 of the 9 participants who achieved different rates of correct responding during training, suggesting that there was a relationship between the
lowest rate required and the lowest rate achieved during training. For the high-and
low-fluency conditions, the relationship is less clear. Although each of these
conditions produced the lowest achieved rate for only one participant, they were both
associated with the highest (or a high) rate of correct responding for five participants.
As such, one can not conclude that either of these conditions consistently produced
the highest rate of correct responding achieved during training.

Table 3

Frequency of Association of the Rate Required Under the Nominal Conditions and
the Rates Achieved Under Those Conditions.

<table>
<thead>
<tr>
<th></th>
<th>Accuracy-Only</th>
<th>Low-Fluency</th>
<th>High-Fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest Rate Achieved</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Moderate Rate Achieved</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Lowest Rate Achieved</td>
<td>6</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

The complexity of these relationships is even more evident when one
evaluates the actual rates achieved by all participants when they are categorized based
upon either the nominal rate required or the actual rates achieved. The average rate of
correct responding achieved for all participants under the accuracy-alone condition
was 27.8 correct responses per minute, while the achieved-rates averaged 30.2 correct
per minute and 30.5 per minute under the low- and high-fluency conditions
respectively. These data suggest that there were no differences between the high- and
low-fluency conditions, and there was no relationship between the rate required and
the rate achieved during training. When the participants are grouped based upon the
rate achieved relative to the other conditions, however, clear differences between the conditions are evident. The average rate of correct responding was 26.5 correct responses per minute for the low-rate condition, 29.4 correct responses per minute for the moderate-rate condition, and 33.0 correct per minute for the high-rate condition. These data suggest that there were 3 distinct condition for comparison.

Advocates of fluency training suggest that the achievement of a high rate of correct responding during training will produce superior retention and generativity. The fact that the relationship between the rate required and the rate achieved during training was not straightforward in this study makes both the analyses of the current results and their comparisons to other research complex. Nonetheless, the differences in retention and generativity revealed when the data are examined based upon either the rate required or the rate achieved has important implications for how research and training is conducted in this area. If it is assumed that the rate achieved was actually the same as the rate that was required and results are examined in terms of the nominal rate required alone, false conclusions about the true effects of the rate achieved during training on retention and generativity may be made. For example, for participant nine, the nominal high-fluency condition produced the highest percent correct observed on all retention probes. If one looks only at the nominal rate required, one would conclude that the imposition of criteria for both accuracy and a high rate of correct responding produces superior retention. If one examines the actual rate achieved, however, this condition resulted in a moderate achieved-rate of correct responding during training and the nominal low-fluency condition produced the
highest rate achieved. This would suggest that it was not critical that this participant learned the tasks to a very high rate of correct responding. Similar differences between the effects of the nominal rate required and the achieved rates were observed for other participants.

It seems likely that the discrepancies between rate required and rate achieved observed during this experiment occurred in other research in this area, however, such discrepancies were not reported in the literature reviewed. Whether the rate required during training actually causes the participant to respond at that precise rate may be a function of the experimental task and parameters employed. For example, during this experiment, participants were required to meet the performance criteria for three consecutive 2-minute sprints, or six minutes. This was done to ensure that responding had stabilized, but also permitted participants to exceed the rate required in the course of meeting the performance requirements. If the experimental task involved briefer intervals, or was terminated as soon as the participant met the performance criteria, the relation between the rate required and achieved may have been stronger than that observed during this experiment.

However, since the limited research in this area has not dealt with this complex issue, a thorough analysis of the current results in terms of both the rate required and achieved is difficult to relate to the current literature. As such, the results of this study will be discussed in terms of the rate achieved, however, graphs of the rate-required data (uncorrected for character set or order – see below) are presented in Appendix i for the reader's review. In addition, the number of practice trials
associated with each condition will be examined.

To facilitate this complex analysis, the reader is urged to examine the various tables and figures while reviewing the written results. Individual results corrected for character set and order in training are presented for each participant in Appendix H. Percent correct and average latencies for each individual are graphed based upon the rate of correct responding achieved during the final three sprints of training (high, moderate and low). The actual number of correct responses per minute (CPM) achieved by that participant and the number of practice trials completed for that condition (PT) are presented next to the line representing each condition. An examination of these figures will significantly enhance the narrative description of the both the individual and overall results. Graphs that present the uncorrected data are presented in Appendix i. The results / discussion section will provide an overview of the results for all participants and a selective analysis of the individual data that either support or refute the efficacy of fluency training.

Analysis of the Effects of Character Sets Independent of the Experimental Conditions

An analysis was completed to examine the effects of the individual character sets on percent correct and average latency. Character sets were balanced across experimental conditions such that each character set was associated with each of the experimental conditions four times. They were not, however balanced for order in training. The analysis was completed based upon the data for all participants and all experimental conditions. Figure 5 presents average percent correct and latency data on
retention probes for all participants with respect to the character sets utilized. Percent correct data on retention probes were mixed, however character set 4 produced accuracy below that observed for the other character sets on 2 of the 4 probes, suggesting that this character set was “harder” for many of the participants. Percent correct data were mixed for character sets 2 and 3 on retention probes, with neither one producing superior performance. Although latency data were similar for all three character sets, character set 4 produced the longest latencies observed during 3 of the 4 retention probes. It appears that the character set associated with a condition produced differential effects on both the accuracy and speed of retention.

Figure 6 presents the same data for generativity probes. Average percent correct was highest for character set 2 on all follow-up probes. Although character sets 3 and 4 produced similar accuracy on the probe completed at 14 days, percent correct for character set 4 was the lowest on 3 of the 4 generativity probes. There were no meaningful differences between the latencies observed on generativity probes with respect to character sets. Appendix J presents the detailed data used for this analysis.

These data suggest that conditions that were associated with character set 4 may have produced results that were a function of the character set rather than the experimental manipulations. These effects may have confounded any differential effects produced by the rate achieved during training. Although character sets were balanced for difficulty, a few participants reported that they thought that the characters in set 4 were less distinct from one another than characters in the other character sets. A more careful examination of the sets in terms of the similarities and
differences between the individual characters may have produced more balanced sets.

Based upon this analysis, a correction factor was applied to the original data to isolate and remove the differential effects of character set 4 independent of the experimental conditions. The factor was established by first calculating the average percent correct for each character set across all participants and all probes to obtain a composite average for that character set. The composite averages for character sets 2 and 3 were then averaged to get an average of the "easier conditions". Finally, the average composite for the "easier conditions" was divided by the composite average for character set 4. The resulting number was the correction factor by which all data involving character set 4 were multiplied. Separate correction factors were established for percent correct and latency on retention and generativity. These correction factors and those employed to correct for order effects (see below) are presented in table 4. Appendix J presents the detailed data used for this analysis.

Analysis of the Effects of Order in Training Independent of the Experimental Condition

Once the data were corrected for character set 4, they were examined to see whether the position that a condition held in the order of training produced differential effects on the percent correct and average latency on probes for retention and generativity. The experimental conditions were balanced for order such that each condition occurred first, second or third in the order of presentation four times. Figure 7 presents average percent correct and average latency data on retention probes for all participants with respect to the order in which conditions were presented in training.
All participants were highly accurate on the retention probe completed immediately after training regardless of where a particular condition fell in the order of training. On follow-up retention probes, however, there appeared to be a clear relation between the position that a condition held in the order in training and percent correct. The highest percent correct was demonstrated for the relationships that participants learned first, moderate accuracy was observed for the condition that was second, and the lowest percent correct was observed for the relationships that participants learned last. Latency data were mixed, but the differences between conditions that were presented first, second or third were no greater than 0.7 seconds on any probe, suggesting that the order in training did not produce any differential effects on latency. The detailed data for this analysis are provided in Appendix K.

Figure 8 presents the same data for generativity probes. Average percent correct on generativity probes was similar to that observed on retention probes. Participants tended to be most accurate on probes for the condition that they trained on first, and least accurate on probes for the condition that they learned last, with moderate average percent correct for the condition that was trained second. Average latencies were similar during generativity probes regardless of the order in which training was presented. Interestingly, average latencies for the condition that was last in the order of training, the condition that tended to produce the lowest percent correct, produced the shortest latencies on all follow-up probes. These data may, however be skewed by the data for participant eight who “guessed.”
Figure 5. Average Percent Correct (upper panel) and Latencies (lower panel) on Retention Probes for All Participants For the Three Character Sets.
Figure 6. Average Percent Correct (upper panels) and Latencies (lower panel) on Generativity Probes for All Participants For the Three Character Sets.
Table 4

Correction Factors for Character Set and Order in Training Employed in the Retention / Generativity Experiment.

<table>
<thead>
<tr>
<th></th>
<th>Retention</th>
<th>Generativity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>For Character Sets:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent Correct</td>
<td>1.213</td>
<td>1.280</td>
</tr>
<tr>
<td>Average Latency</td>
<td>0.978</td>
<td>0.975</td>
</tr>
<tr>
<td><strong>For Order in Training</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent Correct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Order</td>
<td>0.845</td>
<td>0.866</td>
</tr>
<tr>
<td>Third Order</td>
<td>1.096</td>
<td>1.155</td>
</tr>
<tr>
<td>Average Latency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Order</td>
<td>1.137</td>
<td>1.036</td>
</tr>
<tr>
<td>Third Order</td>
<td>1.118</td>
<td>1.143</td>
</tr>
</tbody>
</table>

When taken together, these data suggest that the position that a condition held in the order of training may have affected the accuracy produced by that condition on probes for retention and generativity, but did not differentially affect the rate (as measured by latency, see discussion at the end of the introduction) at which correct responding occurred. These effects may have confounded any differential effects produced by the rate required or achieved during training.

Based upon this analysis, a correction factor was established to isolate and remove the effects of the position that a condition held in the order of training. Like with character sets, the average percent correct for the first, second and third positions...
were calculated for all participants and all probes to obtain a composite average. The average composite for the second order was then divided by the composite average for the first and third orders separately to arrive at the correction factors for that position. This produced separate correction factors for the first and third positions for percent correct and average latency for retention and generativity. The correction factors for are presented in table 4.

It is not clear how the effects of order in training could have been addressed within the experimental parameters employed in this experiment. All sessions were conducted approximately 7 days apart over seven weeks. The training sessions were the first three sessions for all participants, and the order in which a particular condition was presented was balanced across participants, with each of the nominal conditions being presented first, second or third 4 times. Perhaps a different training task would permit training to occur for all three conditions during the same session might have reduced this effect, but the time required to complete training under the current parameters made this impossible.

Overall Results for Rate Achieved During Training

Once data were corrected for both the effects of character set 4 and the order in training, percent correct and average latency data were examined based upon the rate achieved during training. The grouped data only represent those nine participants who achieved different rates of correct responding during training under the different experimental conditions, and each condition was designated as producing a high,
moderate, or low rate of correct responding during training. Individual data for all participants will also be highlighted. Because there were differences in performance on retention and generativity probes, the results will be discussed separately.

Retention

Overall, the results of this experiment do not support the contention that training composite skills until they are both accurate and occur at a high rate will produce superior retention of those skills. Figure 9 presents the average percent correct and latencies for the nine participants included in this analysis on retention probes. Data are organized by the conditions that produced the highest, lowest and moderate rate achieved for individual participants (see Appendix G). There were no meaningful differences in percent correct for the three groups with respect to the rate achieved. With the exception of the 14-day probe, the differences between the groups were no greater than 7% correct. Given the similarity of the data for the high- and low-rate-achieved conditions across all probes, the slight difference for the moderate-rate-achieved condition observed on the 14-day probe can not be interpreted in terms of the rate achieved. These results suggest that for the participants taken as a group, different rates of correct responding achieved during training did not produce significantly different accuracy on retention probes. It is also interesting to note that despite the fact that the lowest-rate-achieved group averaged the fewest number of practice trials, percent correct on retention probes was equivalent to the groups that averaged 60 more practice trials per condition. This suggests that the number of
practice trials did not have a differential effect on performance. Average latencies were also similar for all three groups on all probes, with no meaningful differences in terms of the rate achieved. These data suggest that the rate achieved during training did not differentially affect either the accuracy or speed of retention.

With few exceptions, the individual results mirror those of the group. The data for one participant (5) suggest that the condition that resulted in the highest rate of during training produced the most accurate retention of those skills and the condition that produced the lowest achieved-rate produced the lowest accuracy. However, the high-rate condition also involved the greatest number of practice trials, so one can not conclude that either variable was the critical determinant of performance (see figure 20 in Appendix H). For another participant (1) the condition that resulted in the lowest achieved-rate during training produced the lowest accuracy observed on all follow-up retention probes. This condition, however, involved half as many practice trials as the other conditions, so one can not assume that the percent correct data were a function of the rate achieved during training. In addition, two conditions with higher, but different achieved rates produced 100% correct on all four probes for retention. (see figure 12 in Appendix H). Finally, for one participant (6), the condition that produced the highest achieved-rate during training produced the greatest accuracy on two of the three follow-up probes and accuracy greater than that produced by the low achieved-rate condition on all three follow-up probes (see figure 22 in Appendix H. These data suggest that for one participant the rate achieved produced differential effects on the accuracy of retention independent of the effects of practice trials.
Figure 7.  Average Percent Correct (upper panel) and Average Latencies (lower panel) on Retention Probes for All Participants Based Upon Order in Training.
Figure 8. Average Percent Correct (upper panel) and Average Latencies (lower panel) on Generativity Probes for All Participants Based Upon Order in Training.
Figure 9. Average Percent Correct (upper panel) and Average Latencies (lower panel) for Nine Participants for High, Moderate and Low Rate Achieved During Training on Retention Probes.
For the remainder of the participants, the obtained results do not suggest that the highest achieved rate during training will produce the most accurate retention. For six participants (2, 3, 7, 8, 10, and 12) conditions that produced similar achieved rates during training resulted in different levels of accuracy on probes (see figures 14, 16, 24, 26, 30, and 34 in Appendix H). For example, training under two of the conditions resulted in equally high rates of correct responding for participant ten. On retention probes, however, one of those conditions produced the highest percent correct observed on all three follow-up probes, and the other produced the lowest on two of them (see figure 30 in Appendix H). If the rate achieved during training produces differential performance with respect to accuracy of retention, one would expect similar accuracy for conditions that resulted in similar achieved rates during training.

For two participants (1 and 9), experimental conditions that resulted in different achieved rates during training produced similar accuracy on at least three of the four probes for retention. For example, for participant nine, the conditions associated with the highest and lowest achieved rates during training (34.4 and 29.5 correct per minute) produced similar accuracy on 3 out of 4 probes for retention (see figures 12 and 28 in Appendix H). If the rate of correct responding during training were the critical determinant of accuracy of retention, one would expect to see significant differences in performance on retention trials.

Finally, for several participants, different rates of correct responding during training resulted in different levels of accuracy on probes, however, not in the direction that one would expect if higher achieved-rates during training produce
superior performance. For three participants (4, 8 and 10) a condition that produced a high achieved-rate during training produced the least accurate performance observed on at least two of the follow-up probes (see figures 18, 26 and 30 in Appendix H). For 6 participants (3, 4, 8, 10, 11, and 12), the condition that resulted in the lowest achieved-rate during training produced accuracy equal to or above that produced by one of the higher rate conditions on at least two of the follow-up probes (see figures 16, 18, 26, 20, 32, and 34 in Appendix H). Participant nine demonstrated the greatest accuracy for characters learned under the condition that resulted in a moderate rate of correct responding during training (figure 28 in Appendix H). Finally, for participant eleven, the experimental conditions produced different achieved-rates during training, but percent correct was mixed for all conditions on retention probes with no one condition producing superior accuracy (see figure 32 in Appendix H). These results suggest that although different achieved rates during training may produce different levels of retention, the achievement of a high rate of correct responding will not necessarily produce superior performance.

The results with respect to average latencies also suggest that a high rate of correct responding during training does not produce superior retention. Only one participant (8) achieved the shortest latencies observed for the conditions that resulted in the highest achieved-rate during training, and those results must be discarded as she reported, and the accuracy data confirm, that she guessed on the final retention probes for that condition. The condition that resulted in the lowest achieved-rates during training, however, did produce the longest latencies observed during two of the three
follow-up probes (see figure 26 in Appendix H). For one participant (5), the lowest-rate-achieved condition produced the longest average latencies during two of the three follow-up probes, but the absolute differences between the conditions were small. These data provide very weak support for the contention that the rates achieved during training produce differential effects on retention.

The results for all other participants were mixed. For five participants (1, 2, 3, 7 and 10), conditions that produced similar achieved rates during training produced different latencies on probes. For example, for participant three, although one of the high-rate-achieved conditions produced consistently shorter latencies during probes, a second condition with a slightly higher achieved-rate produced the longest latencies observed on two of the three follow-up probes (see figures 12, 14, 16, 24, and 30 in Appendix H). For six of the participants (1, 4, 5, 6, 11, and 12), conditions that resulted in different achieved-rates during training produced similar average latencies during retention probes (see figures 12, 18, 20, 22, 32 and 34 in Appendix H). These data suggest that different rates of correct responding achieved during training did not differentially affect the speed of retention. In summary, although the results for a few participants suggest that fluency training of component skills produces superior retention, the data for a majority of the participants indicate that some uncontrolled variable other than the achieved-rate of correct responding during training is responsible for the differences observed on probes for retention.
Generativity

Figure 10 presents average percent correct and latencies on generativity probes for the 9 participants examined. Differences between the groups were more apparent on generativity probes, but the differences were not what one would expect if a higher rate of correct responding during training produces superior generativity. The condition that resulted in the highest achieved-rate during training produced accuracy slightly below that produced by the low-rate condition on all three follow-up probes. The moderate-rate-condition produced the highest percent correct on two of the three follow-up probes, but with the exception of the 14-day probe, accuracy was similar for all three conditions. These data suggest that training a skill until it occurs both accurately and at a high rate during training does not produce superior accuracy on retention.

Although average latencies were consistently longer on generativity probes than retention probes, the conditions that produced the highest and lowest achieved-rates during training produced similar average latencies on probes. The moderate-rate conditions produced more variable latencies, but they were longer than those produced by the lowest-achieved-rate conditions during 2 out of 3 follow-up probes. Like with percent correct, it does not appear that differential rates achieved during training produced significant differences in the rate of correct responding observed during probes for generativity.
Figure 10. Average Percent Correct (upper panel) and Average Latencies (lower panel) for Nine Participants for High, Moderate and Low Rate Achieved During Training on Generativity Probes.
The results for individual participants were more mixed than the grouped data suggest, but overall do not support the contention that training component skills until they are both accurate and occur at a high rate during training will produce more accurate generativity. Data for four participants (1, 2, 5, and 6) offer some support for the efficacy of fluency training. For one participant (1), the condition that resulted in the lowest achieved-rate during training produced the lowest percent correct on two of the three follow-up probes, but the moderate- and high-rate conditions produced similar performance on two out of the three follow-up probes (see figure 13 in Appendix H). One participant (5) demonstrated greater accuracy on generativity probes for the skills learned under the condition resulting in the highest rate achieved during training, however, this condition also involved significantly more practice trials than either of the other conditions, and the moderate- and low-rate groups produced similar accuracy on two of the three follow-up probes (see figure 21 in Appendix H). For two participants (2 and 6), training conditions produced both similar achieved rates during training and percent correct on probes, but accuracy was fairly low for all conditions (see figures 15 and 23 in Appendix H). Finally, participant six demonstrated the lowest percent correct for the condition with the lowest achieved-rate on all three follow-up probes, but this condition also involved the fewest practice trials (see figure 23 in Appendix H). These results provide very limited support for the notion that the rate achieved during training determines the accuracy of generativity.

Data for the remainder of the participants do not support this notion. For four
participants (3, 7, 8, and 10) conditions that produced similar achieved-rates during training resulted in very different levels of accuracy on probes. For example, training under two of the conditions resulted in equally high rates of correct responding during training for participant eight. On generativity probes, however, those conditions produced very different levels of accuracy that were below the low-rate condition on two of the three probes (see figure 27 in Appendix H). In contrast, for four participants (1, 9, 11 and 12), experimental conditions that resulted in different achieved rates during training produced similar percent correct on at least 2 of the 3 follow-up probes (see figures 13, 29, 33 and 35 in Appendix H). If the rate of correct responding during training were the critical determinant of the accuracy of generativity, one would expect the percents correct observed during probes to vary with differences in the rates achieved during training.

Finally, for four participants, different rates of correct responding during training resulted in different levels of accuracy on probes, however, not in the direction that one would expect if higher rates produce superior performance. For two participants (4, and 9) the condition with the highest achieved-rate during training produced the least accurate performance observed on probes (see figures 19 and 29 in Appendix H). Two participants (8 and 11) demonstrated the highest degree of accuracy on characters learned under the condition that resulted in the lowest achieved-rate during training on at least two of the three follow-up probes (see figures 27 and 33 in Appendix H), while one participant (9) demonstrated equal accuracy for the conditions that resulted in moderate and low achieved-rates during training (see
figure 29 in Appendix H). Overall, both grouped and individual data do not support
the contention that training component skills until they are both accurate and occur at
a high produces more accurate generativity.

The results with respect to average latencies were also mixed, and more
variable than those observed during probes for retention. For one participant (six) the
condition with the highest achieved-rate during training produced the shortest
latencies and the lowest-rate condition produced the longest latencies on two of the
three follow-up probes. The low-rate condition, however, involved the least practice
(see figure 23 in Appendix H). For one participant (5), the condition with the lowest
achieved-rate during training produced the longest latencies on two of the three
follow-up probes, but the highest-achieved-rate condition produced latencies longer
than the moderate-rate condition on three of the four total probes (see figure 21 in
Appendix H). Finally, for participant eight, the low-rate condition produced the
longest latencies observed on two of the three follow-up probes, but conditions with
equally high achieved-rates produced very different latencies, and this participant
"guessed" in the final probes, making these data uninterpretable. These data represent
the extent of the support for the effects of fluency training on the speed of
generativity.

Data for all other participants were mixed. For five participants (2, 3, 7, 8, and
10), conditions that resulted in essentially equal achieved rates during training
produced different and mixed average latencies during probes (see figures 15, 17, 25,
27, and 31 in Appendix H). In contrast, conditions that produced different rates of
correct responding during training produced similar average latencies during probes for two participants (1 and 12 – see figures 13 and 35 in Appendix H).

Finally, for five participants, different rates of correct responding during training resulted in different average latencies on probes, however, not in the direction that one would expect if higher rates produce superior performance. Latencies were shortest during probes for the condition that produced the lowest achieved-rate during training for one participant (11 – see figure 33 in Appendix H)), and the condition that produced moderate correct rates for one participant (4 – figure 19 in Appendix H). The condition that resulted in the highest achieved-rate during training actually produced the longest latencies observed during probes for two participants (4 and 9 – see figures 19 and 29 in Appendix H). Finally, for participant three, two conditions that produced equally high achieved-rates during training produced both the shortest and the longest latencies observed during probes (see figure 17 in Appendix H). These data suggest that the rate achieved during training did not produce differential effects on the average latencies observed during probes for generativity.

The results for participants both individually and as a group suggest that the rate achieved during training did not produce differential effects on retention and generativity independent of the number of practice trials. For a few participants under a few conditions, the rate achieved during training appears to have affected performance during probes, but for the majority of participants and conditions it did not do so. It is difficult to compare these data to previous research as much of the previous research was confounded because the number of practice trials was not
controlled. Furthermore, none of the three studies that did control for the number of practice trials evaluated the relationship between the rate required and the rate achieved during training. As such, although the results of the experiment contradict those obtained by Berquam (1981) and Shirley & Pennypacker (1994), it is impossible to identify the reasons for those contradictions.

Greater differences between the groups might have been observed if the differences between the rates achieved under the experimental conditions had been larger. Pilot data suggested, however, that the highest rate required (30 correct per minute) was as fast as possible for most people. Because the position of the numbers in the circle around the character rotated and the pointer returned to the center of the circle after each trial, a certain amount of time was necessary to scan for the correct number and move the pointer to the correct position. Nonetheless, out of the 36 total training sessions completed under all conditions, participants exceeded the rate of 30 correct per minute during 14 training sessions. In contrast, it was very easy for pilot participants to achieve the low-fluency criteria of 20 correct per minute, and with one exception, every participant exceeded this correct rate under the accuracy-alone condition. However, since 30 correct responses per minute appeared to be the highest rate possible for most pilot participants, the low rate was selected to maximize potential differences between the fluency groups. If the experimental task had permitted a wider range in terms of the rate of correct responding, it is possible that differential effects may have been observed. Similarly, the imposition of criteria that involved larger differences between the rate required for the different conditions may
have produced greater differences in the rates achieved during training and performance on probes than those observed during this study. Future research might replicate the experimental controls employed in this research with a task that permits greater variability in the rate achieved during training.

**Effect of the Number of Practice Trials**

Training under conditions that involve criteria for both accuracy and a high rate of correct responding typically involves more practice than training under criterion for accuracy alone (see Table 1). In a number of the studies reviewed in this area, the number of practice trials completed during training was permitted to vary with the rate requirements of the experimental condition, resulting in significant differences in the amount of exposure that participants had to the learning materials. As a result, it was not possible to discern the effects of either the experimental contingencies or the number of practice trials on the behavior being studied. During this study, the number of 2-minute sprints completed by participants was equated under the experimental conditions using a yoked procedure. During each of the accuracy-alone and low-fluency training session completed after the high-fluency condition, participants were presented with the same number of sprints that they required to meet the performance criteria during the high-fluency condition (see Appendix C). As a result, participants completed a similar number of practice trials under each of the yoked conditions (see Table 1).

Two participants (2 and 4) did not meet the performance criteria under the
high-fluency condition. Data for both were excluded from the numerical analysis, however the data for participant two were unusual enough to warrant a brief discussion. This participant completed 281 practice trials under the accuracy-alone criterion, 692 under the high-fluency condition, and 209 under the low-fluency criteria. Despite these significant differences, training under each condition resulted in essentially equal achieved-rates during training. During retention probes, however, the accuracy-alone condition (moderate number of practice trials) produced the most accurate retention and the conditions associated with the greatest and fewest number of practice trials produced similar, but significantly lower percent correct than the accuracy-alone condition. Average latencies were variable, with no one condition producing consistently shorter latencies on follow-up probes (see figure 8 in Appendix H). Accuracy on probes for generativity was more mixed, but the condition that involved the greatest number of practice trials produced the lowest percent correct observed on 2 of the 3 follow-up probes. Latency data were variable, but the condition that involved the greatest number of practice trials did produce the shortest latencies on two of the three follow-up probes (see figure 9 in Appendix H). These results suggest that for this participant, the number of practice trials completed, independent of the rate required or achieved, did not produce differential effects on the accuracy of retention or generativity, but did affect the speed with which correct responding occurred during generativity probes.

In order to isolate the effects of practice trials alone, the numerical analysis compared only those conditions where a participant achieved similar rates of correct
responding during training. That is, achieved rates that are within 2 correct responses per minute of one another. If there are significant differences between both the rate achieved and the number of practice trials, it is not possible to distinguish which variable might be the functional one. As such, three additional participants (3, 5, and 11) were excluded from the analysis. For the remaining 5 participants (1, 6, 8, 9, and 10) examined, the conditions compared had differences of at least 20 practice trials completed. Comparisons were framed in terms of the nominal rate required.

Overall the data do not support the contention that the number of practice trials associated with a condition produces differential performance on probes for retention and generativity. For two participants (6 and 7), the condition that involved the fewest number of trials produced the lowest percent correct on follow-up probes for retention and generativity (see figures 22, 23, 24 and 25 in Appendix H). In addition, the two conditions with a similar and higher number of practice trials produced higher accuracy on generativity for participant seven (see figures 24 and 25 in Appendix H). For one participant (8), the condition that involved the greatest number of practice trials produced the shortest latencies observed during probes for both retention and generativity, but this participant reported, and accuracy data confirmed, that she guessed on the final probes for this condition (see figures 26 and 27 in Appendix H). These results provide limited support for the differential effects of the number of practice trials on retention and generativity.

For the remainder of the participants, the condition that involved the greatest number of practice trials did not produce better performance on retention or...
generativity probes than the conditions that involved fewer trials for the remaining participants. For two participants (8 and 10), percent correct was greatest for the condition that involved the fewest number of practice trials (see figures 26, 27, 30 and 31 in Appendix H). For one participant (6) percent correct was similar for the compared conditions on probes for generativity, suggesting that the difference in the number of practice trials did not have a differential effect on retention. In probes for retention, the condition that involved fewer practice trials produced superior accuracy on two of the three follow-up probes (see figures 22 and 23 in Appendix H). Finally, participant eight demonstrated the poorest accuracy for characters learned under the condition with the greatest number of practice trials (see figures 26 and 27 in Appendix H). These results suggest that conditions that involved more practice trials did not produce better accuracy of retention or generativity.

With respect to average latencies, the number of practice trials under a condition did not produce differential effects on performance on either retention or generativity probes. For one participant (6) the condition that involved the greatest number of practice trials produced the longest latencies observed on probes for retention and generativity (see figure 22 in Appendix H). For two participants (9 and 10) latency data with respect to the compared conditions were mixed on probes for both retention and generativity (see figures 28, 29, 30 and 31 in Appendix H). When taken together, these results suggest that the number of practice trials alone did not produce differential effects retention or generativity. These data do not support the common sense notion that more practice will result in superior performance. They are
also not consistent with those obtained by Makepeace (1998) who found that the condition that involved the greatest number of practice trials produced the best performance on tests for generalization and generativity.

These unusual results may, in part, be the result of the small actual differences in the number of practice trials completed for the different conditions during this experiment. Although the differences in the number of practice trials completed per condition were substantial for some participants, when one examines the data for all participants, the differences between conditions are fairly small. The average number of trials completed for the high-fluency condition for all participants was 29% higher than the average number completed for the accuracy-alone condition and 16% higher than the average number completed under the low-fluency criteria. If one examines the actual number of trials completed, however, the condition that involved the greatest number of practice trials involved an average of 72 more trials, or approximately 2 ½ additional 2-minute sprints than the condition that involved the fewest number of trials. Perhaps if the actual differences between conditions had been more substantial, the effects of this variable may have been more pronounced. It is also possible that this variable impacted performance in combination with other variables that may or may not have been manipulated during this experiment.

Loss of Performance Over Time

The increased time and cost related to fluency training has practical implications for both educational and business settings. If one invests the additional
time and resources necessary to complete fluency training, it is reasonable to ask how long after training one can expect the positive effects of high-rate fluency training to last, and if there is a point in time after training when those positive effects are lost? In order to address these questions, grouped data were examined with respect to the rate achieved during training.

Figure 10 presents the average percent correct and latencies observed on retention probes for the nine participants whose data were categorized into the lowest-, moderate- and highest-rate-achieved conditions. For the conditions that produced the highest and lowest achieved-rates during training, percent correct declined at essentially the same rate across probes with no meaningful differences between them in terms of rate achieved. The moderate-rate condition showed a slight increase at 14 days and then decreased to within 6% of the other conditions at 28 days. Average latencies were also essentially equal for all three conditions on all four retention probes. These results suggest that a high rate achieved during training neither produced differential effects on the accuracy or speed of retention nor produced performance that was more robust over time than conditions that resulted in a lower rate achieved during training.

Figure 10 presents the same data for generativity probes. Percent correct for the highest-rate-condition decreased at 7 days and then stabilized at approximately 70%. The moderate- and low-rate conditions showed similar trends across probes with the moderate-rate condition producing slightly higher accuracy on two of the three follow-up probes. Percent correct for all three conditions was approximately

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
equal at 28 days. Although percent correct was most stable for the high-rate condition across follow-up probes, this condition showed the largest decreases from training on all three follow-up probes. These data suggest that although different rates achieved during training may predict differences in how rapidly accuracy declines with respect to generativity, the differences are not what one would expect if fluency-training produces superior performance.

Average latency data also suggest that the rate achieved during training did not differentially affect the rate at which latency increased across time. Average latencies for the conditions that produced the highest and lowest achieved-rates during training produced essentially equal latencies on all four probes for generativity. Data for the moderate-rate condition were more variable, but the absolute differences between the conditions were small. Like the percent correct data, these data suggest that achieving a high rate of correct responding during training does not produce superior generativity that persists over time.

When taken together, these results suggest that achieving a high rate of correct responding during training did not produce superior performance on probes for retention or generativity. Moreover, the rate achieved during training did not differentially affect the rate at which either accuracy or speed declined over time. These results are consistent with those obtained by Grindle (1996) who found that accuracy remained stable across retention probes regardless of the rate required during training. Similarly, Bucklin (1998) found that although accuracy was higher on retention probes for the group that completed high-fluency training, the rate required
during training did not differentially affect how quickly the rate of correct responding declined over time. Finally, Ivarie (1986) found that the group that was required to respond at a low rate during training demonstrated the greatest gains in rate of correct responding on retention tests. The results described above, however, must be interpreted cautiously because the number of practice trials was uncontrolled and varied with the experimental conditions. The current results suggest that fluency training did not produce any positive and enduring effects on performance in terms or either retention of generativity that might offset the increased time and cost involved in fluency training.

General Discussion

It appears that neither rate achieved during training nor the number of practice trials associated with a condition produced differential effects on retention or generativity. It does, however, appear that the position that a condition held in the order that training was presented may have differentially affected the accuracy with which participants performed on retention and generativity probes. Participants tended to be most accurate on probes for the condition that was presented first, and least accurate on the condition that was presented last. The particular character set associated with a condition also appeared to produce differential effects on the accuracy of performance on retention and generativity probes independent of the rate required or achieved. Character set 4 was consistently associated with lower accuracy on probes than the other character sets. Given the similarity of the data when it was
grouped with respect to each of the variables evaluated, however, it seems unlikely that any one of the variables examined produced substantial effects alone. Rather, it appears that some combination of the variables examined, or other variables that were not evaluated during this experiment were responsible for any differences in performance that were observed for individual participants. It does not appear that the data that may have been lost for individual trials for several participants altered the overall results obtained for those participants or all the participants as a whole.

As was noted above, it is also possible that the relatively small differences between the conditions in terms of the rate achieved during training may have masked differences that would have been more evident if the differences between the rates required and/or achieved had been greater. It is also possible that the number of practice trials would have been shown to have a greater effect if the differences between the conditions had been greater with respect to this variable.

Procedural variables may also have impacted the rate of correct responding achieved during training and ultimately performance on probes independent of the effects of the variables that were controlled during this study. Due to concerns about the way that the program was recording data, participants were asked to record the correct rate of responding achieved after each 2-minute sprint on a piece of paper. The same paper contained this information for all three training sessions and was present during all three sessions. Several participants reported that regardless of the rate requirement imposed for a given training condition, they tried to beat their best performance on previous conditions. For these individuals, it appears that seeing their
scores during previous sessions exerted greater control over the rate achieved than the
criteria imposed by the experimenter.

The probes for retention and generativity were also quite short, consisting of
only 24 trials. It is possible that differences may have emerged if these probes had
involved a larger number of trials that might have permitted participants to “warm
up”. It is also possible that the results might have been different if the actual
experimental interactions and responses required were different. Much of the research
in the fluency area has involved face-to-face interactions between the experimenter
and the participant with paper-and-pencil type materials. Participants were then asked
to write or speak. During this experiment, the participant was alone in a room and
simply had to click the mouse. The only interactions with the experimenter occurred
separate from the actual training or probe. It is possible that these procedural
differences were at least in part responsible for the different results obtained.

It is also possible that individual differences affected the obtained results. The
purpose of the experiments was not explained to participants before they completed
the study. Several participants, however, developed their own theories about what was
being studied and altered their performance based upon those theories. For example,
participant four reported that she knew that the experiment was a test of her self-
esteeem, and the instructions that i gave her about rate under the low-fluency condition
were false. This participant achieved the highest rate observed during training under
this condition. Other participants may have developed their own rules about how to
correctly respond during the experiment. The participants also presented with
differing levels of expertise in working with the computer mouse. Several participants worked on the computers all day and were able to manipulate the mouse easily and effectively. Others struggled with the mechanics of the task. Although these variables should not have affected the comparisons on the level of the individual, they may have affected the overall grouped comparisons.

Finally, all of the participants in these experiments had a personal relationship with the experimenter, either as colleagues or friends, and all of them refused the compensation that was offered for participation. It is possible that participants who did not know the experimenter may have responded differently under the experimental contingencies.
CHAPTER III

EXPERIMENT II: EFFECTS OF FLUENCY TRAINING ON DISTRACTIBILITY

Participants

Participants were recruited through in-person solicitations and public notices posted at Living Ways, Inc. Sixteen adult participants completed the distractibility experiment, and five of them completed both experiments. In addition, 7 individuals began this experiment, but were unable to meet the specified performance criteria. As such, these participants were excluded from the study. Each participant was between 25 and 65 years old, was a high school graduate, and had taken at least some college courses. All participants could read, write, and manipulate the computer mouse.

Participants were offered $5.00 for each sessions that they completed and bonuses based upon their performance during the testing phase, however, all 16 declined compensation. Participants completed an informed consent form prior to starting either experiment. A copy of the informed consent form is included in Appendix A, and the letter validating the approval of the human subjects institutional review board is included in Appendix B.

Setting

The distractibility study was completed either in an office at Living Ways, Inc. or in a quiet room in the participant’s residence. The office is approximately eight feet
by ten feet and contains one large window. During the first part of training, the experimenter was in the room with the participant, but was positioned such that she could not observe the participant’s performance on the computer. After one or two sprints, the participant was left alone in the room. All participants were alone in the room during the testing phase.

Apparatus

All sessions were completed on a Dell latitude CPT laptop computer with an external mouse. A variation of the Synwork1 - a PC-based synthetic work environment (version 2.13) software (Elsmore, 1996) program was utilized for this study.

Experimental Design

This study employed a between-subjects design. When originally designed, five participants were to be randomly assigned to the accuracy-alone group and each of the two fluency groups (8 and 4 correct responses per minute). After running several participants, however, it was evident that the original performance criteria for the high-fluency condition were too high and needed to be revised. Based upon these participants’ performance, the criteria were lowered and ultimately reduced to 3 consecutive sprints during which the participant made no more than one error and maintained a correct response rate of at least 5 per minute. Given the low number of correct responses required for this condition, it was not possible to have both an accuracy-alone and an additional low fluency condition. As such, this experiment involved two groups: a fluency group wherein participants were required to meet the
performance criteria described above (8 participants); and an accuracy-alone group wherein participants were required to complete three consecutive sprints during which they made no more than one error (8 participants).

Participants were assigned to a particular group based upon when their experimental session occurred. Individual sessions were scheduled at the participant’s convenience. The first eight participants were then assigned to the fluency group, and the final eight participants were assigned to the accuracy-alone group. In order to control for the possible effects of the number of practice trials, the number of trials completed by participants in the accuracy-alone group was based upon the number of trials required to meet criteria for participants in the fluency group. Specifically, 3 participants were required to complete at least 99 practice trials (the average of the fluency group). The other five participants were matched to individuals in the fluency group as precisely as possible such that accuracy-alone participants completed as many or more practice trials than the corresponding participant in the fluency group. The actual number of practice trials provided to each participant is presented in Table 5.

Procedure

A variation of the Synwork1 (version 2.13) software program was utilized (Elsmore, 1996). This PC-based synthetic work involves four different tasks, a primary arithmetic task and three different “distracting” tasks. The tasks will be described in detail below. At the beginning of the session, each participant was
presented with a brief “pre-training” exercise wherein they received an explanation of each task and completed a 2-minute trial of each task individually and all four tasks together. Following a short break, participants worked on the arithmetic task alone until they met the performance criteria for three consecutive sprints (training phase). After each 2-minute sprint participants received the following information about their performance: percent correct; number of correct responses, number of errors; total score; and rate of correct responses per minute. The experimenter provided the final measure. All other measures were provided via a pre-programmed screen.

Table 5

Number of Practice Trials Completed for All Participants in the Fluency and Accuracy-Alone Groups.

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>Fluency Group</th>
<th>Accuracy-Alone Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Trials to Meet Criterion</td>
<td>Number of Trials Required to Complete (Yoked)</td>
</tr>
<tr>
<td>1</td>
<td>160</td>
<td>111</td>
</tr>
<tr>
<td>2</td>
<td>151</td>
<td>108</td>
</tr>
<tr>
<td>3</td>
<td>102</td>
<td>124</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>75</td>
</tr>
<tr>
<td>5</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>6</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>7</td>
<td>77</td>
<td>78</td>
</tr>
<tr>
<td>8</td>
<td>98</td>
<td>98</td>
</tr>
</tbody>
</table>

Once they mastered the arithmetic task at the specified performance level,
participants were presented with the arithmetic task and three additional “distracting”
tasks simultaneously (testing phase). The “distracting” tasks included a “memory
task”, a visual monitoring task, and an auditory monitoring task. Participants were
required to respond to all tasks presented for 15 minutes. The screen template for the
testing phase of this study is presented in figure 11.

Arithmetic Task.

A single arithmetic task involving addition problems with random numbers
between 0 and 999 was presented in the upper right quadrant of the computer screen.
The numbers were presented vertically so that one number appears below the other.
Participants were required to complete the math problems by clicking on the “+” and
“-” boxes below each column until the correct total for that column appears, and then
clicking on the “done” box when they complete the entire equation. All participants
received the same equations. Incorrect responses resulted in a low tone and the loss
of 10 points, while correct responses earned 10 points. This task was completely self-
paced and there was no time limit for participants to respond to a given equation. The
next equation was presented immediately after participants clicked on the “done” box.

Data were collected for percent correct, the number of correct and incorrect
responses per minute, and the total score achieved per minute. The total score per
minute represents the overall performance of a participant on the arithmetic task
during training and testing divided by the number of minutes of exposure to the task
during each phase. Subtracting the points lost due to errors from the total score
produced by correct responses and dividing that number by the duration of exposure produced this score.

![Screen Template for Testing Phase of Distractibility Study](image)

**Figure 11.** Screen Template for Testing Phase of Distractibility Study.

**Memory Task**

A list of six randomly chosen letters (the positive list) from the alphabet with the exceptions of the letters c, d, m, q, and v, was presented in the upper left quadrant of the computer screen. A negative list containing six other randomly selected letters from the alphabet was also established. At the beginning of the testing phase, the
positive list was presented for five seconds and then covered. Every 20 seconds participants were presented with a randomly selected letter from either the positive or negative list for five seconds. Participants were required to click the mouse on the “yes” button if the letter is a member of the positive list, and the “no” button if the letter was part of the negative list to earn points. Incorrect responses resulted in a low tone and the loss of 10 points, while correct responses earned 10 points. If an individual failed to respond during a trial it was considered an incorrect response. Participants could review the original list at anytime, but lost ten points each time that they did so.

Data were recorded for percent correct, the number of correct and incorrect responses per minute, the total score achieved on this task, the number of lapses, and the number of “peeks”. The number of lapses represents the number of times that a trial was presented, but the participant did not respond. The number of “peeks” refers to the number of times that the participant uncovered the list of letters during the testing phase.

Visual Monitoring Task

Participants were presented with a “fuel gauge” with a pointer that moved from the “full” position to the “empty” position at a fixed rate (200 msec per pixel) in the bottom left quadrant of the computer screen. The participants’ task was to prevent the pointer from reaching “empty” by clicking the mouse on the black half circle at the bottom of the gauge. This action returned the pointer to “full”. Points were
awarded for each reset, with the number of points being proportional to the distance of the pointer from the “full” end at the time of reset. Participants could earn up to a maximum of ten points for resets in the most distant 10% of the scale. For example, clicking the mouse when the pointer was near “full” produced 2 points, while clicking the mouse when the pointer was very near “empty” earned ten points. If the participant permitted the pointer to go all the way to “empty”, an “uh oh” sound was heard and ten points were deducted for each second that the pointer remained there. Performance was measured using the total score on the task, the number of times that the participant reset the gauge, the number of lapses when the pointer got to “empty”, and the average amount of time that elapsed before the participant reset the gauge.

Auditory Monitoring Task.

During this task, a brief tone was sounded every five seconds. The tone was either of two frequencies, low (1000 hz) or high (2000 hz). Participants were required to click the mouse on the “alert” button in the bottom right quadrant of the computer screen following a high tone. The probability of high tones was 0.2. Responses were considered correct if they occurred following a high tone, but prior to the next scheduled tone. All other responses were considered incorrect. Ten points were awarded for each correct response, and deducted for each error. If participants failed to respond after a high tone before the next scheduled tone was presented, it was considered an error.

The participants’ performance was evaluated in terms of how they responded
after a high- or low-tone was presented. The percent of “hits” represents the percentage of times that the participant clicked on the “alert” button after a high-tone was sounded. The percent of “false alarms” represents the percentage of times that the participant clicked on the “alert” button after a low-tone was sounded. The percent of “quiets” represents the percentage of times that the participant did not hit the “alert” button after a low-tone sounded, and the percent of “misses” represents the percentage of times that the participant did not click on the “alert” button after a high-tone sounded.

Results / Discussion

A brief analysis of the relationship between the rate required and achieved during training revealed individual differences, however the average rates achieved for all participants in the accuracy-alone and fluency group were essentially equal. Although the experimental contingencies did not require a high rate of correct responding during training for the accuracy-alone group, seven of the eight participants in that group exceeded five correct responses per minute, the performance requirement for the fluency group. Nonetheless, since the achieved rates were equal for the two groups, the only difference between the groups was the rate required during training. As such, the overall data will be discussed in terms of the nominal rate required during training and significant individual results will be noted. Table 6 presents averaged data on the arithmetic task for the fluency and accuracy-alone groups during training (average of last three sprints) and testing.
Total Score Per Minute

The total score per minute represents the overall performance of a participant on the arithmetic task during training and testing divided by the number of minutes of exposure to the task during each phase. This score is produced by subtracting the points lost due to errors from the total score produced by correct responses and dividing that number by the duration of exposure to the task. Participants earned 10 points for each correct response and lost 10 points for each error.

The two groups achieved approximately equal average scores during the final 3 sprints of training. The fluency group earned an average of 19.4 points per minute, with a range of 16.1 to 23.9. Accuracy-alone participants earned an average 20.1, with a range of 12.8 to 26.7. Individual data revealed that three participants in the accuracy-alone group exceeded the average score earned by the fluency group during the final three sprints of training. During the testing phase both groups demonstrated increases in the number of points earned per minute. Participants in the fluency group increased to an average of 24.4 points per minute with a range of 16 to 36 per minute, while the accuracy-alone group averaged 28.3 points per minute, with a range of 16 to 59.3. Four of the eight participants in the accuracy-alone group exceeded the average score earned by the fluency group. The percent increase in average performance from training to testing was 25.6% and 40.3% for the fluency and accuracy-alone groups respectively. These results suggest that requiring a high rate of correct responding during training did not produce superior performance overall on this task.
Table 6
Comparison of Average Training and Testing Scores on the Arithmetic Task for Fluency and Accuracy-Alone Groups.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>FINAL TRAINING Fluency</th>
<th>FINAL TRAINING Accuracy-Alone</th>
<th>TESTING Fluency</th>
<th>TESTING Accuracy-Alone</th>
<th>% CHANGE Fluency</th>
<th>% CHANGE Accuracy-Alone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Score per Minute</td>
<td>19.4</td>
<td>20.1</td>
<td>24.4</td>
<td>28.3</td>
<td>25.6%</td>
<td>40.3%</td>
</tr>
<tr>
<td>Percent Correct</td>
<td>96.6%</td>
<td>96.6%</td>
<td>86.1%</td>
<td>84.8%</td>
<td>-10.9%</td>
<td>-12.2%</td>
</tr>
<tr>
<td>Correct Responses per Minute</td>
<td>6.1</td>
<td>6.3</td>
<td>3.2</td>
<td>3.4</td>
<td>-47.2%</td>
<td>-45.6%</td>
</tr>
<tr>
<td>Incorrect Responses per Minute</td>
<td>0.2</td>
<td>0.2</td>
<td>0.5</td>
<td>0.6</td>
<td>-144.0%</td>
<td>-176.0%</td>
</tr>
</tbody>
</table>

Thirteen of the 16 participants demonstrated increases in the total score achieved on the arithmetic task during the testing phase. Individual comparisons of participants matched for the number of practice trials revealed mixed results. In four cases, participants in the accuracy-alone group exhibited greater increases in their total scores, while in two cases the fluency participant demonstrated larger increases. In one case, the fluency participant exhibited an increase in his total score while the matched accuracy-alone participant demonstrated a slight decline. Finally, one comparison was omitted because although the fluency participant demonstrated greater increases in his total score, this participant received 29% more practice trials than the accuracy-alone participant. Although overall performance on the arithmetic task improved for both groups, the performance of the accuracy-alone group during testing appears to be superior to that of the fluency group.
Both groups averaged 96.6% correct during the final three sprints of training. Given the performance requirement of no more than one error for three consecutive sprints, there was very little variability in individual performance. During testing, average percent correct declined to 86.1% for the fluency group with a range of 72.1% to 98.3%, and 84.8% for the accuracy-alone group with a range of 75.9% to 90.8%. Four of the eight participants in the accuracy-alone group achieved accuracy equal to or greater than the average of the fluency group. The percent decrease in accuracy from training to testing was -10.9% and -12.2% for the fluency and accuracy-alone groups respectively. These data suggest that the imposition of criteria for both percent correct and a high-rate of correct responding during training did not produce superior accuracy in the presence of distracting tasks. The disruption in performance with respect to percent correct was essentially equal for both groups.

Individual comparisons of participants who were matched in terms of the number of practice trials revealed mixed results. In 3 of the 8 comparisons, the participant from the accuracy-alone group demonstrated greater declines in percent correct, however, in 3 other comparisons the participant from the fluency group demonstrated larger declines. In one case, the declines were equal. One comparison was omitted because although the fluency participant achieved greater accuracy during testing, this participant received 29% more practice trials than the accuracy-alone participant.

Overall, it appears that the accuracy of participant’s performance on the
arithmetic task declined equally for both groups. This would suggest that the imposition of criteria for both accuracy and rate of correct responding did not produce performance that was less susceptible to disruption by distracting tasks than the imposition of an accuracy-alone criterion.

**Correct Responses Per Minute**

The number of correct responses per minute was essentially equal for the two groups during the last 3 sprints of training. The fluency group averaged 6.1 correct responses per minute with a range of 5.0 to 7.3, while the accuracy-alone group averaged 6.3 correct with a range of 4.0 to 8.2. During testing, the average number of correct responses per minute declined to 3.2 for the fluency group with a range of 2.3 to 4.0, and 3.4 correct responses per minute for the accuracy-alone group, with a range of 2.0 to 6.7. This represented a decline of 47.2% and 45.6% for the fluency and accuracy-alone groups respectively. These data do not support the contention that requiring a high rate of correct responding during training of component skills will result in performance that is less susceptible to disruption in the face of distractions. The fact that similar rates achieved during training resulted in similar rates of correct responding during testing may seem to support the notion that the rate achieved during training produces differential effects on performance. However, both groups averaged approximately the same number of practice trials, 98.5 and 94.3 for the fluency and accuracy-alone groups respectively, so it is impossible to determine which variable might have been the functional one.
Individual comparisons of participants who were matched in terms of the number of practice trials revealed mixed results. In 4 of the 8 comparisons, the participant from the accuracy-alone group demonstrated larger decreases in the number of correct responses per minute, however, in 3 other comparisons the participant from the fluency group demonstrated larger declines. One comparison was omitted because although the fluency participant achieved greater accuracy during testing, this participant received 29% more practice trials than the accuracy-alone participant. As with accuracy, it appears that the number of correct responses per minute on the arithmetic task declined at equal rates for both groups in the presence of the distracting tasks. This would suggest that the imposition of a rate criterion did not produce performance that was less susceptible to disruption by distracting tasks than the imposition of an accuracy-alone criterion. Again, the high response rates achieved by the accuracy-alone group at the end of training may diminish the functional differences between the experimental contingencies to which the two groups were exposed.

Incorrect Responses Per Minute

Both the fluency and accuracy-alone groups averaged 0.2 incorrect responses per minute with a range of 0 to 0.5 during the final three sprints of training. During testing, the averages increased to 0.5 incorrect responses per minute for the fluency group and 0.6 for the accuracy-alone group. Individual comparisons of participants who were matched in terms of the number of practice trials revealed similar results.
With respect to the number of incorrect responses per minute, there were no meaningful differences between the groups. If the rate required during training produces differential effects on performance, one would expect to see a larger increase in the number of incorrect responses per minute for participants in the accuracy-only group. Again, these results do not support the contention that imposing criteria for both accuracy and a high-rate of correct responding produces superior performance in the face of distracting stimuli. The similar number of practice trials completed by these groups makes it difficult to form any conclusions about the effects of the rate achieved during training on performance.

Performance on Distracting Tasks

Table 7 presents the average final test scores on each of the three distracting tasks presented during the testing phase for the fluency and accuracy-alone groups. Overall, the fluency group averaged 23% more points on all of the four tasks combined during the testing phase. These data suggest that this group performed better overall in the presence of all 4 tasks. The relative results on the individual distracting tasks are difficult to interpret. Although differences in performance may have been a function of the criteria imposed during training, many other uncontrolled variables could have also affected those scores. For example, two participants reported that they completely ignored the “memory task”. One participant failed to look at the positive list at the beginning of the testing phase and forgot that he was able to access the list by “peeking”. The second participant, however, indicated that
she made a “strategic” decision. That is, this participant was exceptionally accurate and fast on the math task (the highest achieved rate observed during training). She indicated that she did not want to “worry” about watching for the letter presentation and knew that she could make up for lost points by improving her score on the math task.

When interviewed, at least half of the participants indicated that they used a strategy that involved focusing on 2 of the tasks, and ignoring, or sporadically responding to the other two. Most participants indicated that they focused on the math task and almost always responded to the auditory monitoring task. However, three participants indicated that at least early in the testing phase, they confused the sound heard after an error is made with the tones produced as part of the auditory monitoring task, resulting in errors on that task. These individual differences clearly affected the results obtained separate from any effects of the independent variable.

Overall, the results of this study do not support the contention that the imposition of criteria for both accuracy and speed in training produces performance that is more robust and less susceptible to disruption in the face of distractions than performance that is trained under conditions that only require accuracy. When both the number of practice trials completed and the rate of correct responding achieved during training were equal, the performance of both groups was similar. Both accuracy (percent correct) and the number of correct responses per minute declined approximately equally for both the fluency and accuracy-alone groups in the presence of the distracting tasks. Given that both groups completed approximately the same
number of practice trials during training, it is not possible to evaluate the possible effects of equal rates achieved during training.

It is difficult to compare the results of this experiment with previous research in the area of fluency training and distractibility. None of the studies completed in this area (Binder 1979a, 1985, and 1987) involved sufficient experimental controls to determine the effects of the independent variable. The number of practice trials associated with the experimental conditions was uncontrolled and varied by participant. In addition, none of these studies involved a control group or condition with which to compare the results of the experimental manipulation. Each of these studies involved a small number of participants who were exposed to the same experimental contingencies. Given these limitations, it is impossible to make any conclusion about the relative effects of the independent variables examined. As such, the discussion will be limited to the results obtained in this experiment.

Like the results of the retention/generativity study, the results for this experiment suggest that the criteria imposed by the experimenter did not determine the maximum rate of correct responding achieved by participants during training. In addition, the nominal required rate did not produce differential performance in the presence of distracting stimuli. It is not possible to evaluate the effects of either the rate achieved or the number of practice trials on performance as these were essentially equal for the two groups. It appears that either one of these variables, or a variable that was not controlled or measured during this experiment produced the observed results.
The specific nature of the experimental task may have impacted the results obtained. This task required participants to add three digit numbers without the benefit of the visual stimuli often associated with “carrying the one”. A number of the pilot subjects indicated that they were unable to “add in their heads”. Many of these participants were unable to meet the performance criteria and were, therefore, excluded from the study. Other participants reported that they found the visual and mechanical aspects of the task awkward. The vertical columns of numbers to be added did not line up horizontally with the buttons that the participant clicked to alter the answer to the equation. This meant that it was easy to click the wrong button. In addition, some participants reported that they rarely used a mouse, so clicking and moving the mouse was awkward.

These characteristics may have limited the maximum possible rate achieved such that there were not large enough differences between the rates required or achieved for the groups. Even after thousands of practice trials, the experimenter was never able to exceed an average of 12 correct responses per minute. A different task that permitted a higher rate of correct responding might have resulted in greater differences in the rates required and achieved by the groups. These larger differences may then have resulted in differential effects on performance.

The specific nature of the distracting tasks may also have affected the results obtained. Each of these tasks was distinctly different from the primary math task. If the distracting tasks had been more similar to the primary task, perhaps involving counting or numbers, more disruption of performance on the math task may have
been observed. Finally, individual differences like those described above may have affected the individual performance observed during testing.

Table 7
Comparison of Average Final Test Scores on Distracting Tasks for Fluency and Accuracy-Alone Groups

<table>
<thead>
<tr>
<th>Task</th>
<th>Fluency</th>
<th>Accuracy-Alone</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Score for All 4 Tasks</td>
<td>1176.4</td>
<td>960.4</td>
<td>22.5%</td>
</tr>
<tr>
<td>&quot;Memory&quot; Task</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Score</td>
<td>245.7</td>
<td>142.9</td>
<td>72.0%</td>
</tr>
<tr>
<td>Percent Correct</td>
<td>84.5%</td>
<td>74.6%</td>
<td>13.4%</td>
</tr>
<tr>
<td>Number Correct / Minute</td>
<td>2.3</td>
<td>2.0</td>
<td>15.0%</td>
</tr>
<tr>
<td>Number Incorrect / Minute</td>
<td>0.4</td>
<td>0.7</td>
<td>-36.6%</td>
</tr>
<tr>
<td>Number of Lapses</td>
<td>2.4</td>
<td>4.9</td>
<td>-50.0%</td>
</tr>
<tr>
<td>Number of Peeks</td>
<td>1.6</td>
<td>0.6</td>
<td>175.0%</td>
</tr>
<tr>
<td>Visual Monitoring Task</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Score</td>
<td>313.9</td>
<td>170.6</td>
<td>84.0%</td>
</tr>
<tr>
<td>Number of Resets</td>
<td>50.8</td>
<td>42.8</td>
<td>18.7%</td>
</tr>
<tr>
<td>Number of Lapses</td>
<td>1.3</td>
<td>2.9</td>
<td>-56.5%</td>
</tr>
<tr>
<td>Reset Time</td>
<td>17.3</td>
<td>19.2</td>
<td>-10.1%</td>
</tr>
<tr>
<td>Auditory Monitoring Task</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Score</td>
<td>297.5</td>
<td>243.8</td>
<td>22.1%</td>
</tr>
<tr>
<td>Percent of Hits</td>
<td>88.1%</td>
<td>78.5%</td>
<td>12.2%</td>
</tr>
<tr>
<td>Percent of False Alarms</td>
<td>1.0%</td>
<td>1.3%</td>
<td>-20.1%</td>
</tr>
<tr>
<td>Percent Quiets</td>
<td>98.4%</td>
<td>98.4%</td>
<td>0.1%</td>
</tr>
<tr>
<td>Percent Misses</td>
<td>11.3%</td>
<td>20.8%</td>
<td>-45.7%</td>
</tr>
</tbody>
</table>

Additional research that imposes rigid experimental controls will permit a more thorough understanding of the effects of various types of training on
performance in the presence of distracting stimuli. In addition, research with experimental tasks that permit a wider range of variability in performance may reveal differential effects that were not evident during this experiment. Finally, research that examines the effects of different types of distracting stimuli on performance may provide additional information.
CHAPTER IV

OVERALL DISCUSSION

When taken together, the results of these experiments suggest that neither the rate required nor the rate achieved during training produced differential effects on either the accuracy or speed of retention, distractibility, or generativity. Furthermore, neither of these variables produced differential effects on the rate at which accuracy and the rate of correct responding decreased over repeated probes for retention and generativity. With respect to the specific participants and experimental conditions evaluated, the imposition or achievement of a high rate of correct responding during training did not produce superior retention or generativity, and did not produce performance that was less susceptible to disruption in the presence of distracting stimuli. The differences between the experimental controls imposed during these experiments and other research in this area make it difficult to compare these results to those obtained through previous research. On the surface, however, they appear to contradict the positive effects of fluency training that have been widely described.

A variety of procedural differences may account for the differences in the results obtained. It is possible that greater differences might have been observed between effects of the different training criteria if the differences between the rate required and achieved for the conditions had been greater. It is also possible that procedural differences including the nature of the experimental tasks, the order in
which the different conditions were presented, the specific character sets involved in training, the specific learning materials presented to the participant, and the nature of the interactions between the experimenter and the participant might have affected the results obtained.

The discovery that the rate required did not determine the maximum rate achieved by participants during training in either experiment might also explain some of the differences between the results obtained here and in previous research. It also has important implications for both research and application in the area of fluency training. Until empirical research in this area identifies the critical variables that produce differential effects on retention and generativity, it is important that both researchers and practitioners examine and report both rates. If one assumes that these two rates are equal, it is possible to draw false conclusions about the differential effects of one or both of these variables. For example, if the nominal high-fluency condition produces the highest percent correct on retention, but actually did not produce the highest achieved-rate during training, one could falsely assume that a high achieved-rate during training produces superior retention and generativity. If however, one only examines the rate achieved, it is not possible to examine the results obtained in terms of the existing literature. As such, it is important that researchers and practitioners examine and report on both the rate required and rate achieved during training.

Empirical research that examines the relationship between the rate required and the rate achieved during training and their effects on performance may help
clarify this complex issue. Additional research that attempts to identify the variables that produced the positive effects of fluency training described by other researchers through rigorous scientific methods would also enhance the limited body of data available in this area. Research tasks that permit greater variability in terms of the rate of correct responding required and achieved may also reveal effects that were not observed under the current experimental parameters. Finally, experimental parameters that control both the minimum and the maximum rate that can be achieved by participants may permit more rigorous comparisons of the rates required and achieved.
Appendix A

Informed Consent Completed by All Participants
An Examination of the Effects of Fluency Training on Retention, Distractibility, and Generativity

Principal Investigator: Jack Michael, Ph.D.
Student Investigator: Victoria Pelletiere, MA

I have been invited to participate in a research program entitled: An Examination of the Effects of Fluency Training on Retention, Distractibility, and Generativity. The purpose of this research is to examine the effects of different training methods on learning and performance. Once completed, this research will partially fulfill the requirements for Victoria Pelletiere to complete her Ph.D. in Psychology.

My consent to participate in this project indicates that I will participate in up to 9 experimental sessions lasting between 15 minutes and one hour and 15 minutes each. I will be paid $5.00 at the end of each session that I complete. In addition, if I participate in the retention / generativity study, and complete all seven sessions, I will receive a $15.00 bonus at the end of the 7th session. If I participate in the distractibility study, I will earn one cent for each point received during the final phase of the study. If I complete both studies, and perform well during the distractibility study, I may earn up to $75.00. I accept that the research will be conducted at the offices of Living Ways, Inc. in Kalamazoo, MI, and that there will be no benefits other than financial for myself as a result of my participation in this study.

I will be asked to complete a variety of tasks on the computer using the mouse including a matching task, arithmetic tasks, a memory task, and visual and auditory monitoring tasks. The researcher will be measuring both my accuracy and speed on these tasks. I am able to read, write, manipulate a computer mouse and keyboard, and have sufficient time to complete the proposed studies.

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 additional treatment will be made available to the subject except as otherwise specified in this consent form. I may become bored, frustrated, fatigued, or experience eyestrain as a result of looking at the computer screen during each session. I may, however, terminate this experiment at any time, or extend the breaks that are pre-programmed during the study. I may also inform the researcher of my discomfort and the session will be interrupted for a break or terminated at my discretion.

All information collected from me will be confidential. That means that my name will not appear on any information that identifies my performance during this research. A randomly assigned number that will be used for all data recording and reporting will identify me. A Research Associate will maintain a coded master list that identifies subjects' names and numbers in a locked file. The Student Investigator will never have knowledge of assigned numbers. Once all of the data have been collected, the master list will be destroyed. If this research is published, my name will not appear in the published report.
I may refuse to participate or quit at any time without prejudice or penalty. If I have any questions or concerns about this research, I may contact Jack Michael at 372-3075 or Vicky Pelletier at 343-6355. The participant may also contact the Chair, Human Subjects Institutional Review Board at 387-8293 or Vice President for Research at 387-8298 if questions or problems arise during the course of the study. My signature below indicates that I understand the purpose, procedures, and requirements of this study and agree to participate.

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

Participant Signature
Date

Witness Signature
Date
Appendix B

Protocol for Clearance From the Human Subjects Institutional Review Board
Date: May 2, 2001

To: Jack Michael, Principal Investigator
   Victoria Pellettiere, Student Investigator for dissertation

From: Michael S. Pritchard, Interim Chair

Re: HSIRB Project Number: 01-03-11

This letter will serve as confirmation that your research project entitled “An Examination of the Effects of Fluency Training on Retention, Distractibility, and Generativity” has been approved under the expedited 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: May 2, 2002
Appendix C

Presentation of Experimental Conditions for Each Participant in the Retention / Generativity Experiment
Presentation of Experimental Conditions for Each Participant in the Retention / Generativity Experiment

<table>
<thead>
<tr>
<th>Participant</th>
<th>Order of Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First</td>
</tr>
<tr>
<td>1</td>
<td>Accuracy-only</td>
</tr>
<tr>
<td>2</td>
<td>Accuracy-only</td>
</tr>
<tr>
<td>3</td>
<td>High-fluency</td>
</tr>
<tr>
<td>4</td>
<td>High-fluency</td>
</tr>
<tr>
<td>5</td>
<td>Low-fluency</td>
</tr>
<tr>
<td>6</td>
<td>Low-fluency</td>
</tr>
<tr>
<td>7</td>
<td>Accuracy-only</td>
</tr>
<tr>
<td>8</td>
<td>Accuracy-only</td>
</tr>
<tr>
<td>9</td>
<td>High-fluency</td>
</tr>
<tr>
<td>10</td>
<td>High-fluency</td>
</tr>
<tr>
<td>11</td>
<td>Low-fluency</td>
</tr>
<tr>
<td>12</td>
<td>Low-fluency</td>
</tr>
</tbody>
</table>

Note: During yoked conditions, the number of practice trials presented was equal to the number completed by that participant under the high-fluency condition.
Appendix D

Character Sets Utilized During Training in the Retention / Generativity Experiment
Character Sets Utilized During Training in the Retention / Generativity Experiment

<table>
<thead>
<tr>
<th>Number</th>
<th>Character Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>123456789</td>
</tr>
<tr>
<td>2</td>
<td>西海日中自月圆旦由</td>
</tr>
<tr>
<td>3</td>
<td>来年単于节上生乐市</td>
</tr>
<tr>
<td>4</td>
<td>夫宴祖学小夏喜花参</td>
</tr>
</tbody>
</table>
Appendix E

Overview of Individual Sessions in the Retention / Generativity Experiment
Overview of Individual Sessions of the Retention / Generativity Experiment

Session 1:
1. Pre-training: Single presentation of each procedure (group presentation of relationships, individual presentation of relationships, single sprint, retention probe (six trials), generativity probe (six trials)).
2. Break and questions
3. Training with one of the two fluency criteria or accuracy-alone criterion
4. Within-session retention and generativity probes

Session 2:
1. Seven-day retention probe for relationships learned during Session 1
2. Seven-day generativity probe for relationships learned during Session 1
3. Training with one of the two fluency criteria or accuracy-alone criterion
4. Within-session retention and generativity probes

Session 3:
1. Fourteen-day retention probe for relationships learned during Session 1
2. Fourteen-day generativity probe for relationships learned during Session 1
3. Seven-day retention probe for relationships learned during Session 2
4. Seven-day generativity probe for relationships learned during Session 2
5. Training with one of the two fluency criteria or accuracy-alone criterion
6. Within-session retention and generativity probes
Overview of Individual Sessions of the Retention / Generativity Experiment
(Continued)

Session 4:
1. Fourteen-day retention probe for relationships learned during Session 2
2. Fourteen-day generativity probe for relationships learned during Session 2
3. Seven-day retention probe for relationships learned during Session 3
4. Seven-day generativity probe for relationships learned during Session 3

Session 5:
1. Twenty-eight-day retention probe for relationships learned during Session 1
2. Twenty-eight day generativity probe for relationships learned during Session 1
3. Fourteen-day retention probe for relationships learned during Session 3
4. Fourteen-day generativity probe for relationships learned during Session 3

Session 6:
1. Twenty-eight day retention probe for relationships learned during Session 2
2. Twenty-eight day generativity probe for relationships learned during Session 2

Session 7:
1. Twenty-eight day retention probe for relationships learned during Session 3
2. Twenty-eight day generativity probe for relationships learned during Session 3

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Appendix F

Mathematical Equations Used During Generativity Probes in the Retention / Generativity Study
Mathematical Equations Used During Generativity Probes in the Retention / Generativity Study

<table>
<thead>
<tr>
<th>Within-Session Probe</th>
<th>7-Day Probe</th>
<th>14-Day Probe</th>
<th>28-Day Probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 + 2</td>
<td>7 + 8</td>
<td>9 + 3</td>
<td>5 + 6</td>
</tr>
<tr>
<td>6 - 4</td>
<td>5 - 2</td>
<td>8 - 1</td>
<td>9 - 5</td>
</tr>
<tr>
<td>3 x 4</td>
<td>1 x 7</td>
<td>2 x 1</td>
<td>3 x 1</td>
</tr>
<tr>
<td>6 / 2</td>
<td>8 / 4</td>
<td>6 / 3</td>
<td>4 / 2</td>
</tr>
<tr>
<td>9 + 4</td>
<td>8 + 9</td>
<td>5 + 7</td>
<td>5 + 2</td>
</tr>
<tr>
<td>3 - 2</td>
<td>7 - 4</td>
<td>9 - 2</td>
<td>9 - 3</td>
</tr>
<tr>
<td>6 x 1</td>
<td>1 x 4</td>
<td>3 x 2</td>
<td>3 x 6</td>
</tr>
<tr>
<td>8 / 2</td>
<td>6 / 1</td>
<td>9 / 1</td>
<td>2 / 1</td>
</tr>
<tr>
<td>4 + 2</td>
<td>8 + 7</td>
<td>2 + 8</td>
<td>8 + 6</td>
</tr>
<tr>
<td>6 - 2</td>
<td>9 - 1</td>
<td>8 - 2</td>
<td>9 - 7</td>
</tr>
<tr>
<td>3 x 5</td>
<td>2 x 3</td>
<td>1 x 2</td>
<td>1 x 6</td>
</tr>
<tr>
<td>8 / 1</td>
<td>3 / 1</td>
<td>6 / 2</td>
<td>8 / 4</td>
</tr>
<tr>
<td>5 + 4</td>
<td>6 + 5</td>
<td>2 + 9</td>
<td>4 + 8</td>
</tr>
<tr>
<td>8 - 4</td>
<td>5 - 3</td>
<td>8 - 6</td>
<td>4 - 2</td>
</tr>
<tr>
<td>1 x 3</td>
<td>7 x 1</td>
<td>7 x 2</td>
<td>6 x 3</td>
</tr>
<tr>
<td>5 / 1</td>
<td>9 / 3</td>
<td>8 / 2</td>
<td>6 / 1</td>
</tr>
<tr>
<td>2 + 6</td>
<td>3 + 5</td>
<td>1 + 4</td>
<td>1 + 8</td>
</tr>
<tr>
<td>9 - 8</td>
<td>9 - 6</td>
<td>8 - 3</td>
<td>4 - 3</td>
</tr>
<tr>
<td>1 x 5</td>
<td>1 x 8</td>
<td>8 x 1</td>
<td>2 x 9</td>
</tr>
<tr>
<td>4 / 2</td>
<td>4 / 1</td>
<td>8 / 1</td>
<td>3 / 1</td>
</tr>
<tr>
<td>1 + 3</td>
<td>7 + 5</td>
<td>4 + 7</td>
<td>7 + 8</td>
</tr>
<tr>
<td>5 - 4</td>
<td>7 - 1</td>
<td>6 - 5</td>
<td>2 - 1</td>
</tr>
<tr>
<td>1 x 9</td>
<td>4 x 2</td>
<td>5 x 2</td>
<td>2 x 7</td>
</tr>
<tr>
<td>2 / 1</td>
<td>7 / 1</td>
<td>5 / 1</td>
<td>9 / 3</td>
</tr>
</tbody>
</table>
Appendix G

Designation of High-, Moderate- and Low-Rate Groups Based Upon Average Rate of Correct Responding Achieved During the Final Three Sprints of Training
Designation of High-, Moderate- and Low-Rate Groups Based Upon Average Rate of Correct Responding Achieved During the Final 3 Sprints of Training

<table>
<thead>
<tr>
<th>Subject</th>
<th>Accuracy-Only</th>
<th>Designation in Terms of Rate Achieved</th>
<th>Average Correct Responses per Minute</th>
<th>Designation in Terms of Rate Achieved</th>
<th>Designation in Terms of Rate Achieved</th>
<th>Designation in Terms of Rate Achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>29.4</td>
<td>High</td>
<td>23.5</td>
<td>Low</td>
</tr>
<tr>
<td>1</td>
<td>25.4</td>
<td>Low</td>
<td>32.8</td>
<td>High</td>
<td>30.7</td>
<td>Moderate</td>
</tr>
<tr>
<td>2</td>
<td>19.7</td>
<td>NA</td>
<td>19.8</td>
<td>NA</td>
<td>22.0</td>
<td>NA</td>
</tr>
<tr>
<td>3</td>
<td>26.4</td>
<td>Low</td>
<td>36.7</td>
<td>High</td>
<td>35.2</td>
<td>High</td>
</tr>
<tr>
<td>4</td>
<td>29.4</td>
<td>Moderate</td>
<td>32.1</td>
<td>High</td>
<td>23.5</td>
<td>Low</td>
</tr>
<tr>
<td>5</td>
<td>26.0</td>
<td>Moderate</td>
<td>23.3</td>
<td>Low</td>
<td>29.5</td>
<td>High</td>
</tr>
<tr>
<td>6</td>
<td>22.5</td>
<td>Low</td>
<td>28.1</td>
<td>Moderate</td>
<td>30.0</td>
<td>High</td>
</tr>
<tr>
<td>7</td>
<td>35.9</td>
<td>NA</td>
<td>35.5</td>
<td>NA</td>
<td>35.9</td>
<td>NA</td>
</tr>
<tr>
<td>8</td>
<td>24.3</td>
<td>Low</td>
<td>32.3</td>
<td>High</td>
<td>32.1</td>
<td>High</td>
</tr>
<tr>
<td>9</td>
<td>29.8</td>
<td>Low</td>
<td>34.3</td>
<td>High</td>
<td>31.8</td>
<td>Moderate</td>
</tr>
<tr>
<td>10</td>
<td>30.0</td>
<td>NA</td>
<td>28.8</td>
<td>NA</td>
<td>30.7</td>
<td>NA</td>
</tr>
<tr>
<td>11</td>
<td>32.6</td>
<td>High</td>
<td>27.6</td>
<td>Low</td>
<td>30.1</td>
<td>Moderate</td>
</tr>
<tr>
<td>12</td>
<td>31.4</td>
<td>Low</td>
<td>30.8</td>
<td>Low</td>
<td>35.1</td>
<td>High</td>
</tr>
<tr>
<td>Average</td>
<td>27.8</td>
<td>30.2</td>
<td>30.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix H

Individual Results for Retention / Generativity Experiment Corrected with Respect to Character Set and Order in Training
Results for Individual Participants in the Retention / Generativity Experiment

Percent correct and average latencies for each individual are graphed based upon the average rate of correct responding achieved during the final three sprints of training (highest, moderate, and lowest). Data are presented for probes of retention and generativity completed immediately after training and at approximately 7, 14, and 28 days after the training session. The average number of correct responses per minute (CPM) achieved during the final three sprints of training and the number of practice trials completed for that condition (PT) are presented next to the line representing each condition. The data and graphs presented in Appendix H have been corrected for both the effects of character set 4 and order in training. Graphs for the original, uncorrected data are presented in Appendix I for the reader's review.

Participant One

Participant One was exposed to the experimental conditions in the following order: accuracy-alone; high fluency; and low-fluency. Under a yoked condition, the number of sprints required under the low-fluency condition was equated to the number of sprints required to meet the performance criterion under the high-fluency condition. As a result, this participant completed 103 practice trials during accuracy-alone training, and 202 and 213 trials under the high- and low-fluency conditions respectively. Due to a data recording error, it appears that a single trial was not recorded during the first sprint under the accuracy-alone training condition. If one assumes that this trial was incorrect, this participant would have been required to
complete one additional sprint, or approximately 26 additional trials under this condition.

An analysis of the average rate of correct responding during the final three sprints of training revealed that this participant actually achieved the highest rate of correct responding under the low-fluency condition (32.8), but it was only slightly higher than the rate of 30.7 correct per minute achieved during the high-fluency condition. She averaged 25.4 correct responses per minute during the final three sprints completed under the accuracy-alone criterion.

Figure 12 presents Participant One's performance on the retention probes.

**Percent Correct:** There appeared to be a relation between percent correct and the rate achieved during training. The condition associated with the lowest rate achieved produced the least accurate performance. This condition also involved the fewest number of practice trials. The conditions that produced the highest and moderate rates during training produced 100% correct on probes. There also appeared to be a relation to the number of practice trials. The conditions that involved a similar and high number of trials produced accuracy above that produced by the condition that involved the fewest practice trials.

**Average Latencies:** There was no apparent relation between average latency and the achieved-rate. The highest- and lowest-rate-achieved conditions produced similar, but variable latencies. The long latencies observed for the 7-day and 14-day probes for the moderate condition are uninterpretable in terms of the rate achieved. There is no apparent relation to the number of practice trials.
Figure 12. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 1 on Retention Probes.
Figure 13 presents Participant One’s performance on the generativity probes.

**Percent Correct:** There appeared to be a relation between percent correct and the highest achieved rate during training. Percent correct was 100% for the highest-achieved-rate condition across all probes. The lowest-rate-achieved condition produced accuracy below that of the other conditions on two of the three follow-up probes. The drop for the moderate-rate condition at 28-days is difficult to interpret in light of this participant’s overall performance. With respect to the number of practice trials, the condition that involved the greatest number produced the highest percent correct, and the condition that involved the fewest produced the lowest, suggesting a relation.

**Average Latencies:** The moderate-rate condition showed slightly longer latencies at 7 and 28 days. Otherwise there is no relation with the rate achieved or the number of practice trials.

The data for this participant appear to support the contention that training skills until they are both accurate and occur at a high rate produces more accurate retention and generativity. The similarity of the results for the moderate- and highest-rate conditions in terms of the actual rate achieved and the percent correct produced, does no permit any conclusions about the impact of a moderate rate of correct responding during training. The rate achieved during training did not, however, have differential effects on the rate of correct responding during probes as measured by average latency.
Figure 13. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 1 on Generativity Probes.
Participant Two

Participant Two was exposed to the experimental conditions in the following order: accuracy-alone; low-fluency; and high-fluency. The number of practice trials was uncontrolled and varied by experimental condition such that this participant completed 281 practice trials under the accuracy-alone condition, 692 under the high-fluency criteria, and 209 under the low-fluency criteria. Due to data recording errors it appears that a single trial was not recorded during the final sprint of the low-fluency condition, and two trials were omitted during the 18th of 21 sprints under the accuracy-alone condition. If one assumes that these trials were incorrect, this participant would have been required to complete one additional sprint in each condition, or approximately 13 trials under the accuracy-alone condition and 21 trials under the low-fluency performance criterion.

Participant Two did not meet the performance requirement during training under the high-fluency condition. After completing 33 sprints and 692 practice trials, the session was terminated. An examination of the last 3 sprints of training under each condition revealed that average correct response rates were almost equivalent for all three conditions, with rates of 21 correct per minute under the high-fluency condition, 19.8 per minute under the low-fluency condition, and 19.7 correct responses per minute under the accuracy-alone condition. Despite these similarities, the significant differences in performance on the retention probes for the three conditions, and the large differences in the number of practice trials completed under the three different conditions make an examination of this participant’s performance valuable.
Figure 14 presents Participant Two’s performance on the retention probes.

**Percent Correct:** There was no apparent relation for either the rate achieved or the number of practice trials. Training under the three nominal conditions produced essentially equal correct rates during training. Percent correct dropped sharply and with similar trends on follow-up retention probes for all three conditions, but there were differences between them. For two conditions, accuracy decreased sharply on the probe completed at 7 days and remained below 35% correct on all subsequent probes. For the third condition, however, percent correct was higher than the other two conditions on all follow-up retention probes. With respect to practice trials, the condition that involved a moderate number of practice trials produced more accurate responding than either other condition. Given that this participant completed 692 practice trials under the high-rate condition, more than twice the number completed for the accuracy-alone condition, and more than three times the number completed for the low-fluency condition, these accuracy data are surprising.

**Average Latency:** Average latencies were mixed. For the high- and moderate-rate conditions latency varied, but remained below 5.1 seconds on all probes. For lowest-rate-achieved condition, average latencies spiked on the probe completed at 7 days, then decreased to the levels observed for the other conditions.

Figure 15 presents Participant Two’s performance on the generativity probes.

**Percent Correct:** There does appear to a relationship between the rate achieved and percent correct on generativity. All three conditions produced both similar rates achieved and poor accuracy on follow-up probes. However, there were differences
between the conditions, with the highest-rate condition producing the lowest accuracy on two of the three follow-up probes. There was no relation to the number of practice trials, as the condition that involved the greatest number of trials produced percent correct below that of the other conditions on two of the three follow-up probes.

**Average Latency:** There is no apparent relation to either the rate achieved or the number of practice trials. Average latencies for all three conditions were variable and mixed. With respect to practice trials, the condition that involved a moderate number of trials produced longer latencies on two of the three follow-up probes.

Overall, the results for this participant provide mixed support for the contention that the rate of correct responding during training determines future performance on probes for retention and generativity. Despite the fact that these conditions produced similar rates of correct responding during training, there were differences in accuracy during probes for retention, and in average latencies during probes for generativity. However, similar rates achieved produced similar accuracy under probes for generativity and similar latencies under probes for retention. These results also suggest that the number of practice trials completed in a given condition alone does not determine future performance of skills. The condition that involved the greatest number of practice trials produced the poorest accuracy on two out of three probes. Given the result with respect to number of practice trials, it is unclear what effect additional practice trials under the low-fluency (19.8 correct per minute) or accuracy-alone (19.7 correct per minute) criteria would have on performance on retention probes.
Figure 14. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 2 on Retention Probes.
Figure 15. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 2 on Generativity Probes.
Participant Three

Participant Three was exposed to the experimental conditions in the following order: high-fluency; accuracy-alone; and low-fluency. The number of sprints required under the accuracy-alone and low-fluency conditions was equated to the number of sprints required to meet the performance criteria under the high-fluency condition. As a result, this participant completed 435 practice trials during accuracy-alone training and 454 and 453 trials under the high- and low-fluency conditions respectively. An analysis of the rate of correct responding during the final three sprints of training revealed that the terminal correct response rates for the high-fluency and low-fluency conditions were essentially equal, with average correct response rates of 35.1 per minute for the high-fluency condition and 36.7 per minute for the low-fluency condition. Under the accuracy-alone condition this participant averaged 26.4 correct responses per minute during the final 3 sprints of training.

Figure 16 presents Participant Three’s performance on the retention probes.

Percent Correct: This participant obtained 100% correct on the probe completed immediately after training under all three conditions. There appeared to be a relation between the rate achieved during training and percent correct, but not in the direction that one would expect if higher achieved rates produce superior performance. One of the two conditions that produced resulted in equally high rates of correct responding during training produced accuracy below that of the other conditions on two of the three follow-up probes. The other high-rate condition
produced 100% correct on the first two probes, then accuracy dropped to near the
other high-rate condition on days 14 and 28. Finally, the low achieved-rate condition
produced 100% correct on probes completed immediately after training, and at 7 and
14 days, but accuracy dropped to just below the other conditions at 28 days. With
respect to the number of practice trials, there is no relation, as the condition that
involved the fewest number of trials produced superior retention on two of the three
follow-up probes.

Average Latency: There was no apparent relation between average latencies
and either the rate achieved or the number of practice trials. The conditions that
produced equal and high achieved-rates produced both the shortest and the longest
latencies observed during retention probes. The low-rate condition produced moderate
latencies on two of the three follow-up probes. Similarly, the two conditions that
involved equal numbers of practice trials produced the longest and shortest latencies
observed.

Figure 17 presents Participant Three’s performance on generativity probes.

Percent Correct: Accuracy data were variable and mixed for all three
conditions across all 3 follow-up probes, with no apparent relation to the rate
achieved or the number of practice trials. Furthermore, the two conditions with equal
achieved-rates produced different levels of accuracy on two of the three follow-up
probes. There was no apparent relation to the number of practice trials.

Average Latency: Latency data were more stable and similar to, but higher
than those observed during retention probes. Like with retention, there was no
Figure 16. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 3 on Retention Probes.
Figure 17. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 3 on Generativity Probes.
apparent relation between average latencies and either the rate achieved or the number of practice trials. The conditions that produced equal and high achieved-rates, and involved an equal number of practice trials, produced both the shortest and the longest latencies observed during retention probes. The low-rate condition produced moderate latencies on all three follow-up probes.

The results for the participant do not support the contention that training skills until they are both accurate and occur at a high rate produces superior retention and generativity. The two conditions that produced equally high achieved-rates during training produced lower percent correct on retention than the low-rate condition, and mixed data with respect to generativity. These conditions also produced both the shortest and longest latencies observed on probes for retention and generativity. The results also suggest that the number of practice trials associated with a condition does not produce differential effects on retention or generativity.

**Participant Four**

Participant Four was exposed to the experimental conditions in the following order: high-fluency; low-fluency; and accuracy-alone. The number of sprints required under the low-fluency and accuracy-alone conditions was equated to the number of sprints required to meet the performance criterion under the high-fluency condition. As a result, this participant completed 539 practice trials during accuracy-alone training and 467 and 487 trials under the high- and low-fluency conditions
respectively.

Participant Four did not meet the response under the high-fluency training condition, but exceeded this requirement during training for the low-fluency condition. After 18 sprints and 467 practice trials, she chose to terminate the session. An analysis of the final 3 sprints of each condition revealed that Participant Four averaged 23.5 correct responses per minute for the high-fluency condition, 32.1 correct for the low-fluency condition, and 29.4 for the accuracy-alone condition. When the experimenter reminded her of the correct response requirement during the low-fluency condition, she commented that she thought that the experiment was about self-esteem and that in fact, I actually wanted her to go faster. This participant also rescheduled two sessions, resulting in slightly different follow-up intervals for the different conditions. Follow-up probes occurred at the following intervals for each condition: high-fluency – 7, 14, and 29 days; low-fluency – 7, 15, and 35 days; and accuracy-alone – 8, 15, and 28

Figure 18 presents Participant Four’s performance on the retention probes. Percent Correct: This participant achieved 100% accuracy on the probes completed immediately after training. There did appear to be a relation between the rate achieved during training and percent correct on retention, but not in the direction that one would expect if fluency training produces superior performance. The condition that resulted in the highest achieved-rate during training produced the lowest accuracy on all follow-up probes. The moderate-rate condition produced accuracy above the low-rate condition on two of the three follow-up probes, but these
data are uninterpretable given her performance for the high-rate condition. Since this Participant did not meet the performance criteria for the high-fluency condition, it is not possible to evaluate the potential effects of the number of practice trials.

**Average Latency:** There was no apparent relation between average latencies and either the rate achieved or the number of practice trials. With the exception of the 7-day probe, average latencies were similar for all three conditions, with no meaningful differences between them. Even on the 7-day probe, the difference between the conditions was no greater than 1.2 seconds.

Figure 19 presents Participant Four’s performance on the generativity probes.

**Percent Correct:** This participant achieved 100% correct on the probe immediately after training for all three conditions. Like with retention, the condition that resulted in the highest rate achieved during training produced much poorer accuracy on probes than the other two conditions. The moderate-rate condition, however, produced accuracy above the low-rate condition on all three follow-up probes. Again, these differences are uninterpretable.

**Average Latency:** The relation between average latencies and rate achieved was not what one would expect if fluency training produces superior generativity. The high-rate condition produced the longest latencies on all three follow-up probes while the moderate-rate condition produced the shortest latencies on two of the three follow-up probes. The absolute differences between the moderate- and low-rate conditions were, however, very small.
Figure 18. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 4 on Retention Probes.
Figure 19. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 4 on Generativity Probes.
When taken together, the data for this participant suggest achieving a high rate of correct responding during training does not produce superior retention or generativity as measured by percent correct and latency.

Participant Five

Participant Five was exposed to the experimental conditions in the following order: low-fluency; accuracy-alone; and high-fluency. The number of practice trials was not controlled and this participant completed 131 trials during training under the accuracy-alone condition, 274 trials under the high-fluency condition, and 149 trials under the low-fluency criterion. Due to data recording errors, it appears that data for a single trial were omitted from 2 final sprints for the high- and low-fluency conditions. If one assumes that those trials were incorrect, this participant would have been required to complete two additional sprints under each criterion, or approximately 54 and 50 additional practice trials for the high-fluency and low-fluency criterion respectively. An analysis of the average rates of correct responding during the final 3 sprints of training indicated that this participant achieved the highest rate of correct responding under the high-fluency condition, with an average of 29.5 per minute. However, she achieved higher rates of correct responding under the accuracy-alone condition (26 per minute) than under the low-fluency condition (23.3 per minute).

Figure 20 presents Participant Five’s performance on the retention probes.

Percent Correct: For this participant, there appeared to be a clear relation between the rate achieved during training and percent correct on retention. The
condition that produced the highest achieved-rate during training and involved the
greatest number of practice trials produced 100% correct on all four probes, while the
condition that produced the lowest achieved-rate during training produced the lowest
percent correct on three of the four total probes. With respect to the number of
practice trials, the condition that involved significantly more trials produced superior
accuracy, but the condition that involved the fewest trials produced accuracy above
the condition with a moderate number. Given that the condition that produced the
greatest accuracy involved both the highest achieved-rate and the greatest amount of
practice, it is impossible to disentangle the effects of these two variables.

Average Latency: There also appeared to be some relation between the rate
achieved and latency. The condition that produced the lowest achieved-rate during
training produced the longest latencies on two out of the three follow-up probes with
an increasing trend. Data for the high-and moderate-rate conditions were mixed, with
no one appearing superior to the other. The absolute differences between the
conditions, however, were small, with no difference greater than one second. There
was no apparent relation to the number of practice trials.

Figure 21 presents Participant Five's performance on the generativity probes.

Percent Correct: For this participant, there appeared to be a relation between
percent correct and a high-rate of correct responding during training. The condition
that produced the highest achieved-rate during training and involved the greatest
number of practice trials produced clearly superior accuracy on all three follow-up
generativity probes. Percent correct for the moderate- and low-rate conditions were,
however, very similar on three of the four total probes. With respect to the number of practice trials, the condition that involved the greatest number of trials produced the highest percent correct, but the conditions that involved a moderate and low number produced similar accuracy on three of the four total probes. Since the condition that produced the greatest accuracy both produced the highest rate achieved during training and involved the greatest number of practice trials, it is impossible to isolate the effects of either of these variables.

**Average Latency:** There was no apparent relation between average latencies and either the rate achieved or the number of practice trials on generativity probes. The lowest achieved-rate condition produced longer latencies on two of the three follow-up probes, suggesting a relation. The highest achieved-rate condition, however, produced latencies longer than the moderate-rate condition on all three follow-up probes. Similarly, the condition that involved the fewest number of practice trials produced the longest latencies on two of the three follow-up probes, but the condition that involved a moderate number produced the shortest.

Overall, the data for Participant Five suggest that the condition that involved either the greatest number of practice trials or the highest correct rate achieved during training produced superior retention and generativity as measured by accuracy. The data with respect to the moderate- and low-rate conditions and number of practice trials were mixed. The data with respect to average latencies suggest that a low rate achieved during training produces longer latencies, but the moderate and high achieved rates produced similar and mixed latencies.
Figure 20. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 5 on Retention Probes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 21. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 5 on Generativity Probes.
Participant Six

Participant Six was exposed to the experimental conditions in the following order: low-fluency; high-fluency; and accuracy-alone. The number of sprints required under the accuracy-alone condition was equated to the number of sprints required to meet the performance criterion under the high-fluency condition such that this participant completed 150 practice trials during accuracy-alone training and 177 and 218 trials under the high- and low-fluency conditions respectively. An examination of average rates of correct responding during the final 3 sprints of training indicated that training under the high-fluency criterion resulted in average correct response rates (30 per minute) slightly higher than those achieved under the low-fluency condition (28.1 per minute), and significantly higher than those achieved under the accuracy-alone condition (22.5 correct per minute). For this participant, it appears that the terminal rate of correct responding achieved under each condition was similar to the criteria imposed by the experimenter.

Figure 20 presents Participant Six’s performance on the retention probes.

Percent Correct: This participant achieved 100% accuracy on the probes completed immediately after training for all conditions. On follow-up probes, the relation between percent correct and rate achieved was mixed. The condition that produced the highest achieved-rate during training produced more accurate performance than lowest achieved-rate condition on all 3 follow-up probes, and produced accuracy superior to the moderate-rate condition on the two of the three follow-up probes. Given that the high-and moderate-rate conditions had very similar
rates of correct responding during training, and this participant completed 40 more practice trials under the moderate-rate condition, the differences in performance between these conditions is surprising, and do not suggest that the rate of correct responding during training is predictive of future performance. There appeared to be no relation between the number of practice trials and percent correct. Although the condition that involved the fewest trials produced the lowest accuracy on two of the three follow-up probes, the condition that involved the greatest number of trials produced accuracy below the moderate condition on two of the three follow-up probes.

**Average Latency:** Despite differences in the correct rate achieved during training, latency data were similar for all conditions with no meaningful differences between them in terms of the rate achieved or the number of practice trials.

Figure 21 presents Participant Six's performance on generativity probes.

**Percent Correct:** For this participant, there did appear to be a relation between the rate achieved during training and percent correct. For the high-and moderate achieved-rate conditions, conditions with very similar achieved rates during training, accuracy decreased similarly across probes. Accuracy rates for the low-rate condition decreased sharply after training and remained low on subsequent probes. The data with respect to the number of practice trials are mixed. Although the condition that involved the fewest trials produced the lowest percent correct on all three follow-up probes, the conditions that involved both the greatest number and a moderate number of trials produced similar accuracy on all probes for generativity.
Figure 22. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 6 on Retention Probes.
Figure 23. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 6 on Generativity Probes.
Average Latency: For this participant, there was a relation between average latency and the rate achieved during training. The condition that resulted in the lowest rate achieved during training produced the longest latencies on all three of the follow-up probes, while the condition that produced the highest achieved-rate produced the shortest latencies on two of the three follow-up probes. There was no apparent relation to the number of practice trials.

Overall, the data for this participant offer some support for the contention that independent of the number of practice trials, a higher rate achieved during training produces superior retention and generativity. Conditions that produced higher and similar rates during training produced higher accuracy on probes than the condition that produced lower rates. However, on retention probes, there were differences between these conditions. With respect to latencies, the data are mixed. While all three conditions produced similar average latencies during retention probes, the high-rate condition produced the shortest latencies during two of the three follow-up probes for generativity. These data do not support the contention that more practice produces superior performance.

Participant Seven

Participant Seven was exposed to the experimental conditions in the following order: accuracy-alone; high-fluency; and low-fluency. The number of sprints required under the low-fluency condition was equated to the number of sprints required to meet the performance criterion under the high-fluency condition such that this
participant completed 130 practice trials during accuracy-alone training and 162 and 166 trials under the high- and low-fluency conditions respectively. An examination of the rate of correct responding achieved during the final 3 sprints of training indicated that this participant achieved essentially equal rates under all three experimental conditions. She achieved 35.9 correct responses per minute under the high-fluency condition, 35.5 correct per minute under the low-fluency criterion, and 35.9 per minute under the accuracy-alone criterion. Given that she completed approximately the same number of practice trials under the high-and low-fluency conditions, the only difference between the 3 experimental conditions was the experimenter’s instructions.

Figure 24 presents Participant Seven’s performance on the retention probes.

Percent Correct: This participant achieved 100% accuracy on the probes completely immediately after training for all conditions and accuracy remained at 100% for one of the conditions. These data suggest that there was no relation between percent correct and the rate achieved during training. Despite identical achieved-rates during training, percent correct was different for the three conditions. It does appear that there was a relation for number of practice trials, but the small absolute differences in the number of practice trials between the conditions would seem to bar any strong conclusions. Nonetheless, the condition that involved 130 trials produced the lowest accuracy and the condition that involved 166 trials produced the highest, with the condition that involved 162 trials in between them.
Figure 24. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 7 on Retention Probes.
Average Latency: Despite essentially identical achieved-rates during training, average latencies varied during retention probes, with the condition associated with the fewest number of practice trials producing slightly higher average latencies on all 4 probes. These data suggest that achieved rates do not produce differential effects on latency of retention. There was no apparent relation to the number of practice trials.

Figure 25 presents Participant Seven’s performance on the generativity probes.

Percent Correct: With respect to accuracy, the data are very similar to those observed during the retention probes. Despite identical achieved-rates during training, percent correct was different for the three conditions. It also appears that there was a relation for number of practice trials, with the condition that involved the most practice producing the highest accuracy and the condition that involved the least producing the lowest accuracy.

Average Latency: Latency data for generativity probes were mixed for all conditions with no one condition producing consistently shorter latencies. These data do not support the contention that achieved-rates during training produce differential effects on retention and generativity. There were no meaningful differences in terms of the number of practice trials.

Taken together, the data for this participant do not support the contention that the rate of correct responding during training produces differential effects on the accuracy or speed of retention or generativity. Essentially identical achieved rates produced different accuracy and latencies on probes for retention and generativity. These data do, however, provide limited support for the contention that a greater
Figure 25. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 7 on Generativity Probes.
number of practice trials during training will produce superior performance. The small differences between the conditions in terms of number of practice trials, however, make any strong conclusions impossible.

**Participant Eight**

Participant Eight was exposed to the experimental conditions in the following order: accuracy-alone; low-fluency; and high-fluency. The number of practice trials was not controlled, so this participant completed 139 practice trials during accuracy-alone training and 201 and 139 trials under the high- and low-fluency conditions respectively. Due to data recording errors, it appears that data for a single trial were omitted from 1 final sprint during the high-fluency and accuracy-alone training conditions. If one assumes that those trials were incorrect, this participant would have been required to complete one additional sprint under each condition, or approximately 29 and 23 additional practice trials for the high-fluency and low-fluency criterion respectively. An analysis of the average rate of correct responding during the final 3 sprints of training indicated that this participant achieved essentially equal rates under the high-fluency (32.1 per minute) and low-fluency (32.3 per minute) conditions. Under the accuracy-alone condition, this participant averaged 24.3 correct responses per minute.

Figure 26 presents Participant Eight’s performance on retention probes.

**Percent Correct:** This participant achieved 100% accuracy on the probes completed immediately after training for all 3 conditions. Data during follow-up
probes suggest that the relation between percent correct and the rate achieved during training was not what one would expect if fluency training produces more accurate retention. Two conditions had essentially equal achieved rates during training. For one of these conditions, accuracy dropped sharply after training and remained below those produced by the other conditions on all three follow-up probes. For the other, accuracy rates were variable, but well above the other high-rate condition. The lowest achieved-rate condition produced accuracy above that of the two high-rate conditions on two of the three follow-up probes. There was no relation to the number of practice trials, with the condition that involved the greatest number of trials producing the lowest accuracy observed, and the conditions with equal numbers of trials producing mixed, but different accuracy.

**Average Latency:** For this participant, there appeared to be a relation between the rate achieved and average latencies. The condition that resulted in the lowest achieved-rate during training produced the longest latencies during two of the three follow-up probes, and demonstrated an increasing trend. With the exception of the 7-day probe, the two conditions with high and equal achieved rates produced similar latencies that were below the low-rate condition. The data with respect to practice trials were mixed. Although the condition that involved the greatest number of trials produced the shortest latencies during all four probes, the two conditions that involved the same number of trials produced different latencies on two of the three follow-up probes. It should be noted, however, that this participant indicated that she had a hard time with the character set associated with the condition that produced the
shortest observed latencies and involved the greatest numbness of practice trials, so she often just "guessed" during the later retention probes. Accuracy data for the final two probes support those statements. As such, one must be cautious when interpreting these results.

Figure 27 presents Participant Eight’s performance on the generativity probes.

**Percent Correct:** There was no apparent relation between percent correct and either achieved rates or the number of practice trials. The condition that resulted in the lowest achieved-rate produced the highest accuracy on two of the three follow-up probes. The other two conditions that had equal achieved rates produced very different accuracy on all three follow-up probes. These data suggest that the rate achieved during training did not differentially affect accuracy on generativity probes. With respect to practice trials, the condition that involved the most practice produced the poorest accuracy and the two conditions with the same number of practice trials produced different accuracy on two of the three follow-up probes.

**Average Latency:** Latency data were more variable for all three conditions during generativity probes and were mixed with respect to the achieved-rates. The condition that produced the lowest achieved-rate during training produced the longest latencies on two of the three follow-up probes. The two conditions with equal achieved-rates, however, produced very different average latencies on all three follow-up probes. These differences suggest that the rate achieved did not produce differential effects on the rate of correct responding during probes for generativity. The data are equally mixed with respect to the effects of the number of practice trials.
The condition that involved the most practice produced the shortest latencies observed on all three follow-up probes, but since this participant indicated that she “guessed” on probes for this condition, any conclusions about this condition would be suspect. The two conditions that involved the same number of trials produced very different latencies, suggesting that this variable did not produce differential effects.

When taken together, the results for this participant do not support the contention that a higher rate of correct responding during training will produce more accurate retention and generativity. The condition with the lowest achieved rate produced higher accuracy on both retention and generativity probes. In addition, the two conditions with equal achieved rates produced very different performance. The data with respect to latencies suggest that a low achieved-rate during training produces longer latencies, but the differences between the conditions with equal rates during generativity probes seem contradictory. In addition, the integrity of the latency data is suspect, and therefore, one can not form any concrete conclusions about the efficacy of fluency training. Similarly, the percent correct data do not provide any evidence of the importance of the number of practice trials. While the latency data suggest an effect, this is the condition that involved “guessing” so the data are suspect. Although additional practice trials under the condition that produced the highest achieved-rate may have enhanced her performance, one can assume that additional trials completed under the condition that resulted in the lowest achieved-rate would do the same, thus maintaining the differences observed during this experiment.
Figure 26. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 8 on Retention Probes.
Figure 27. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 8 on Generativity Probes.
Participant Nine was exposed to the experimental conditions in the following order: high-fluency; accuracy-alone; and low-fluency. Under the experimental design, the number of sprints for the accuracy-alone and low-fluency conditions was to be equated to the number that this participant required to meet criterion during the high-fluency condition. Due to an experimenter error, this participant had one additional sprint under the high-fluency condition resulting in the following number of practice trials per condition: accuracy-alone – 453 trials; high-fluency – 496 trials; and low-fluency – 506 trials. Due to data recording errors, it appears that data for a single trial were omitted from 1 final sprint during the high-fluency and accuracy-alone training conditions. If one assumes that those trials were incorrect, this participant would have been required to complete one additional sprint under each condition or approximately 29 and 28 additional practice trials for the high-fluency and low-fluency criterion respectively. An analysis of the final 3 sprints of training for each condition revealed that this participant actually averaged a higher rate of correct responding, 34.3 per minute, during the low-fluency condition than she did during the high-fluency condition (31.8 correct responses per minute). She also averaged 29.8 correct responses per minute during training under the accuracy-alone criterion.

Figure 28 presents Participant Nine’s performance on the retention probes.

Percent Correct: There was no apparent relations between percent correct and the rate achieved during training. All three conditions produced similar accuracy on the 7-day probe. For the moderate-rate-achieved condition, percent correct remained
stable across the remaining probes and was above that of the other two conditions on
the 28-day probe. The high- and low-rate-achieved conditions produced similar
accuracy on the 14-day probe, and then decreased sharply at 28 days with the highest-
rate-achieved condition producing the lowest accuracy. There was also no apparent
relation between the number of practice trials and percent correct. Accuracy was
similar for the conditions that involved the greatest and least number of practice trials.

Average Latency: Latency data were mixed, but do not suggest that a higher
achieved-rate during training will produce superior retention. The condition with the
lowest achieved-rate produced the shortest latencies on two of the three follow-up
probes, and the highest achieved-rate condition produced the longest latencies
observed on two of the three follow-up probes. With respect to the number of practice
trials, the condition that involved the most practice produced the longest latencies on
two of the three follow-up probes and the condition that involved the least practice
produced the shortest latencies on two of the three follow-up probes.

Figure 29 presents Participant Nine’s performance on the generativity probes.

Percent Correct: The relation between percent correct and the rate achieved for
this participant was not what one would expect if fluency training produces superior
accuracy of generativity. With the exception of the 7-day probe, the conditions that
produced the moderate and lowest rates during training produced essentially equal
accuracy on generativity probes. The highest-achieved-rate condition produced the
lowest accuracy on two of the four total probes, and percent correct for that condition
was below that produced by the moderate-rate condition on three of the four total
probes. The relation between percent correct and the number of practice trials was also contrary. The condition that involved the greatest amount of practice produced the lowest percent correct on three of the four total probes, while the conditions that involved the least and a moderate amount of practice produced essentially equal accuracy on three of the four total probes.

**Average Latency:** The relation between average latency and the achieved rate was also different than one would expect if fluency training produces superior performance. The condition with the highest achieved-rate produced the longest latencies on all three follow-up probes, while the data for the low-and moderate conditions were mixed, but similar. With respect to practice trials, the condition that involved the greatest number of trials produced the longest latencies observed on all three follow-up probes, and the conditions with a moderate or low number of practice trials produced variable, but similar average latencies.

Taken together, the results for this participant suggest that achieving a high rate of correct responding during training does not produce superior retention or generativity. They also suggest that providing more practice trials for a condition will not produce superior performance on either retention or generativity. Given that greater practice did not appear to produce superior performance for this participant, it seems unlikely that additional practice under the moderate- and low-rate conditions would enhance performance. If, however, any improvement was observed, this would only increase the observed differences between these conditions and the condition that produced the highest achieved-rate during training.
Figure 28. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 9 on Retention Probes.
Figure 29. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 9 on Generativity Probes.
Participant Ten

Participant Ten was exposed to the experimental conditions in the following order: high-fluency; low-fluency; and accuracy-alone. The number of 2-minute sprints completed for the accuracy-alone and low-fluency conditions was equated to the number that this participant required to meet criterion during the high-fluency condition such that this participant completed 671 practice trials under the accuracy-alone criterion, 631 trials under the high-fluency criterion, and 750 trials under the low-fluency criterion. Due to illness, this participant re-scheduled two sessions, resulting in slightly different follow-up intervals for the different conditions. Follow-up probes occurred at the following intervals for each condition: high-fluency – 7, 14, and 28 days; low-fluency – 7, 19, and 28 days; and accuracy-alone – 12, 14, and 29 days. An analysis of the final 3 sprints of training under each condition revealed that this participant achieved approximately equal rates of correct responding under all three conditions, with a slightly lower rate under the low-fluency condition. The actual rates of correct responding achieved were 30.7 per minute under the high-fluency condition, 28.8 per minute for the low-fluency condition, and 30 per minute under the accuracy-alone condition. The substantial differences in both performance and the number of practice trials completed, however, warrant examination.

Figure 30 presents Participant Ten’s performance on the retention probes.

Percent Correct: There was no apparent relation between percent correct and the rate achieved or the number of practice trials. Despite the fact that all three conditions produced essentially equal rates of correct responding during training,
differences in accuracy were observed. The condition that produced an achieved rate of 30.7 per minute produced the highest accuracy on all three follow-up probes, while the condition that resulted in 30.0 correct per minute produced the lowest accuracy on two of the three follow-up probes. The condition that produced 28.8 correct per minute, the lowest absolute rate achieved during training, produced accuracy between the two higher conditions on two of the three follow-up probes. With respect to practice trials, the condition that involved the fewest number of trials produced the highest percent correct on all three follow-up probes and the condition that involved the greatest number of trials produced accuracy between the other two conditions.

**Average Latency:** Latency data also suggest a mixed relation for rate achieved. With the exception of the 14-day probes, the conditions that produced the highest and lowest rates during training produced similar latencies, suggesting that similar rates in training produced similar performance. The latencies for the absolute moderate-achieved-rate condition, however, were varied and cannot be explained in terms of the rate achieved. The data are equally mixed with respect to practice trials. The conditions with the greatest and least number of practice trials produced essentially equal latencies on three of the four total probes.

Figure 31 presents Participant Ten’s performance on the generativity probes.

**Percent Correct:** Percent correct on generativity probes was more variable than that observed on retention probes for all three conditions, and do not suggest a functional relation between the rate achieved during training and accuracy of retention. Although accuracy was similar on two of the three follow-up probes for the
condition associated with the lowest achieved-rate and one of the high-achieved-rate conditions, the other high achieved-rate condition produced highly variable accuracy that was significantly below that of the other conditions on two of the three follow-up probes. There is no apparent relation to the number of practice trials. The conditions that involved the most and the least practice produced similar accuracy on three of the four total probes.

**Average Latency:** Latency data were mixed, with no one condition producing superior performance across probes. As such, there were no apparent relations to either rate achieved or the number of practice trials.

When taken together, the results for this participant do not support the contention that training skills until they are both accurate and occur at a high rate produces superior retention or generativity. Conditions that produced equally high achieved-rates during training produced very different accuracy on probes for both retention and generativity. Latency data were mixed, but again, the conditions with similar achieved rates during training produced different average latencies across follow-up probes. These data also suggest that the number of practice trials associated with a condition is not a critical determinant of future performance. The condition that involved the fewest number of practice trials produced the most accurate retention and generativity.
Figure 30. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 10 on Retention Probes.
Figure 31. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 10 on Generativity Probes.
Participant Eleven was exposed to the experimental conditions in the following order: low-fluency; accuracy-alone; and high-fluency. The number of 2-minute sprints completed for all conditions were uncontrolled, however, this participant completed a similar number of practice trials under each condition, 170 under the accuracy-alone criterion, 152 under the high-fluency criteria, and 167 under the low-fluency criteria. Due to apparent data recording errors, it appears that data for a single trial were omitted from 1 of the final sprints during training under the low-fluency and accuracy-alone criterion. If one assumes that those trials were incorrect, this participant would have been required to complete one additional sprint under each of these conditions, or approximately 28 practice trials under the low-fluency and 24 trials under the accuracy-alone criterion.

An analysis of the average rate of correct responding during the final 3 sprints of training indicated that this participant achieved the highest average rate of correct responding under the accuracy-alone condition, with a rate of 32.6 per minute. Under the high-fluency condition, this participant achieved a moderate rate of 30.1 correct per minute, and under the low-fluency condition she achieved the lowest rate of 27.6 per minute.

Figure 32 presents Participant Eleven’s performance on the retention probes. Percent Correct: There was no apparent relation between percent correct and either the rate achieved or the number of practice trials. This participant achieved 100% accuracy on the retention probe completed immediately after training for all
three conditions, and data were variable for all three conditions on follow-up probes. Despite differences in the rate achieved, no one condition produced superior accuracy on more than one follow-up probe for retention, and each condition was associated with the lowest accuracy on one probe. The results were equally mixed with respect to the number of practice trials. Conditions with essentially equal numbers of practice trials produced different accuracy on probes for retention.

**Average Latency:** Average latencies were similar on retention probes for all three conditions, with no meaningful differences in terms of either the rate achieved or the number of practice trials.

Figure 33 presents Participant Eleven’s performance on generativity probes.

**Percent Correct:** There was no apparent relation between percent correct and rate achieved. The conditions that resulted in the highest and lowest rates achieved during training produced similar percent correct on three of the four probes, and the low-rate condition was superior on the fourth probe. The moderate achieved-rate condition produced variable data with respect to percent correct, but it was below that produced by the low-rate condition on two of the three follow-up probes. It did, however, appear that there was a relation between the number of practice trials and percent correct. The two conditions that involved almost the same number of trials produced similar accuracy on two of the three follow-up probes and the condition that involved the fewest trials produced the lowest percent correct on two of the three follow-up probes.
Figure 32. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 11 on Retention Probes.
Figure 33. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 11 on Generativity Probes.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
**Average Latency:** The relation observed between average latencies and rate achieved was not what one would expect if fluency training produces superior generativity. The low achieved-rate-condition produced the shortest latencies on all three follow-up probes. The high-rate condition spiked at 14 days, but produced average latencies below the moderate-rate group on two of the three follow-up probes. The data with respect to practice trials were mixed. The condition that involved the least practice produced the longest latencies observed on two of the three follow-up probes, but the conditions with essentially the same number of practice trials produced different average latencies across probes.

When taken together, the results for this participant suggest that achieving a high rate of correct responding during training does not produce superior retention and generativity. Percent correct data were mixed for the three conditions on retention probes, and the highest and lowest-rate conditions produced similar accuracy on probes for generativity. With respect to average latencies, conditions with different achieved rates produced similar average latencies during retention probes, and the lowest-rate-condition produced the shortest latencies observe during generativity probes. The results with respect to the differential effects of the number of practice trials associated with a condition were mixed. Although the results with respect to percent correct on generativity suggest a relation, the data for accuracy on retention and latency on generativity were mixed. Finally, during retention probes, latencies for all conditions were similar despite differences in the number of practice trials. Additional practice trials under the highest- and lowest rate-achieved conditions may
have enhanced performance, thus increasing the similarity between the effects of these conditions and furthering the observed differences between these conditions and the moderate-rate-achieved condition. Given the trends observed for this participant, it seems unlikely that additional training would alter the overall results for this participant significantly.

Participant Twelve

Participant Twelve was exposed to the experimental conditions in the following order: low-fluency; high-fluency; and accuracy-alone. The number of 2-minute sprints completed for the accuracy-alone condition was equated to the number of sprints this participant required to meet the performance criterion under the high-fluency condition such that this participant completed 120 practice trials under the accuracy-alone criterion, 127 trials under the high-fluency criterion, and 120 trials under the low-fluency condition. Due to an apparent data recording error, it appears that data for a single trial were omitted from 1 of the final sprints during training under the low-fluency criterion. If one assumes that this trial was incorrect, this participant would have been required to complete one additional sprint, or approximately 28 practice trials under this condition.

An analysis of the final 3 sprints of training revealed that the low-fluency and accuracy-alone conditions resulted in almost equivalent average rates of correct responding, with 30.8 correct per minute under the low-fluency condition, and 31.4 per minute under the accuracy-alone high-fluency criteria. The high-fluency condition
resulted in the highest average rate of correct responding achieved, 35.1 per minute.

Figure 34 presents Participant Twelve’s performance on retention probes.

**Percent Correct:** This participant achieved 100% correct on the probes completed immediately after training for all three conditions. On follow-up probes, there was no apparent relation between percent correct and either the rate achieved or the number of practice trials. The two conditions with essentially equal achieved rates during training produced different accuracy on probes. One produced 100% accuracy on three of the four total probes while the second produced 87% correct on all three follow-up probes. It is possible, however, that additional practice trials under this condition (30.8 CPM) might produce performance that was more similar to the other low-rate condition. The condition that resulted in the highest achieved rate during training produced variable data, but accuracy was never greater than both the other conditions on any probe. With respect to the number of practice trials, all three conditions involved essentially the same number of trials, but accuracy differed for all three conditions, suggesting that there was no relation.

**Average Latency:** With the exception of the 7-day probe, average latencies were similar for characters learned under all three conditions, with no meaningful differences between them in terms of the rate achieved. These data do not support that contention that a high rate of correct responding during training will produce superior retention. These data do, however, provide some evidence that the number of practice trials is a determinant of future performance. Conditions that involved a similar number of trials produced similar average latencies on three of the four total probes.
Figure 34. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 12 on Retention Probes.
Figure 35 presents Participant Twelve’s performance on generativity probes.

**Percent Correct:** Like on the retention probes, this participant achieved 100% accuracy on the generativity probes completed immediately after training for all three conditions. Data for the follow-up probes do not suggest a relation between percent correct and the rate achieved during training or the number of practice trials. The two conditions that resulted in essentially equal and low achieved rates during training produced different accuracy. For the high-rate and one of the equal, but low rate conditions, accuracy remained at 100% on all follow-up probes. For the other low condition, percent correct was well below the other conditions on all three follow-up probes. Despite the fact that all three conditions involved about the same number of practice trials, differences were observed in accuracy on all three follow-up probes. Additional practice for the condition that produced 30.8 correct responses per minute during training would not have altered the overall results obtained.

**Average Latency:** The data with respect to average latencies were mixed and like the average latencies observed during the retention probes, no particular condition appears to have produced superior performance on generativity probes. These data suggest that although there was no relation between the rate achieved and average latencies, there was some support for the importance of the number of practice trials. All three conditions involved a similar number of trials and produced similar average latencies on three of the four total probes.

When taken together, the data for this participant suggest that achieving a high rate of correct responding does not produce superior retention or generativity in
terms of either percent correct or average latencies. Accuracy for at least one of the low-rate conditions was equal to or higher than that produced by the high rate condition on all probes for retention and generativity. Latency data were similar for all conditions regardless of the rate achieved during training. With respect to the effects of practice trials, percent correct data do not suggest differential effects, but average latency data suggest that conditions that involve the same amount of practice will produce similar latency.
Figure 35. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 12 on Generativity Probes.
Appendix I

Individual Results for Retention / Generativity Experiment Uncorrected With Respect to Character Set or Order in Training
Figure 6. Percent correct (upper panel) and latencies (lower panel achieved by Participant 1 on Retention Probes) Uncorrected for Character Set or Order.
Figure 7. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 1 on Generativity Probes Uncorrected for Character Set or Order.
Figure 8. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 2 on Retention Probes Uncorrected for Character Set or Order.
Figure 9. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 2 on Generativity Probes Uncorrected for Character Set or Order.
Figure 10. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 3 on Retention Probes Uncorrected for Character Set or Order.
Figure 11. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 3 on Generativity Probes Uncorrected for Character Set or Order.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 12. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 4 on Retention Probes Uncorrected for Character Set or Order.
Figure 13. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 4 on Generativity Probes Uncorrected for Character Set or Order.
Figure 14. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 5 on Retention Probes Uncorrected for Character Set or Order.
Figure 15. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 5 on Generativity Probes Uncorrected for Character Set or Order.
Figure 16. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 6 on Retention Probes Uncorrected for Character Set or Order.
Figure 17. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 6 on Generativity Probes Uncorrected for Character Set or Order.
Figure 18. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 7 on Retention Probes Uncorrected for Character Set or Order.
Figure 19. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 7 on Generativity Probes Uncorrected for Character Set or Order.
Figure 20. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 8 on Retention Probes Uncorrected for Character Set or Order.
Figure 21. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 8 on Generativity Probes Uncorrected for Character Set or Order.
Figure 22. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 9 on Retention Probes Uncorrected for Character Set or Order.
Figure 23. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 9 on Generativity Probes Uncorrected for Character Set or Order.
Figure 24. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 10 on Retention Probes Uncorrected for Character Set or Order.
Figure 25. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 10 on Generativity Probes Uncorrected for Character Set or Order.
Figure 26. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 11 on Retention Probes Uncorrected for Character Set or Order.
Figure 27. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 11 on Generativity Probes Uncorrected for Character Set or Order.
Figure 28. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 12 on Retention Probes Uncorrected for Character Set or Order.
Figure 29. Percent correct (upper panel) and latencies (lower panel) achieved by Participant 12 on Generativity Probes Uncorrected for Character Set or Order.
Figure 30. Average Percent correct for all Participants and All Experimental Conditions on Retention (upper panel) Generativity Probes (lower panel).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 31. Average Latencies for all Participants and All Experimental Conditions on Retention (upper panel) Generativity Probes (lower panel).
Figure 32. Average Percent Correct (upper panel) and Average Latencies (lower panel) for Nine Participants for High, Moderate and Low Rate Achieved During Training on Retention Probes.
Figure 33. Average Percent Correct (upper panel) and Average Latencies (lower panel) for Nine Participants for High, Moderate and Low Rate Achieved During Training on Generativity Probes.
Figure 34. Average Percent Correct and Latencies on Retention Probes for All Participants Based Upon Order in Training.
Figure 35. Average Percent Correct and Latencies on Generativity Probes for All Participants Based Upon Order in Training.
Figure 36. Average Percent Correct and Latencies on Retention Probes for All Participants For the Three Character Sets.
Figure 37. Average Percent Correct and Latencies on Generativity Probes for All Participants For the Three Character Sets.
Appendix J

Average Percent Correct and Latencies for All Participants on All Retention and Generativity Probes Based Upon the Character Set Associated With That Condition
Average Percent Correct and Average Latencies for Nine Participants on all Retention and Generativity Probes Based Upon the Character Set Associated with the Condition.

### Average Percent Correct

<table>
<thead>
<tr>
<th>Character Set</th>
<th>Immediate</th>
<th>7-Day</th>
<th>% Change From Training</th>
<th>14-Day</th>
<th>% Change From Training</th>
<th>28-Day</th>
<th>% Change From Training</th>
<th>Immediate</th>
<th>7-Day</th>
<th>% Change From Training</th>
<th>14-Day</th>
<th>% Change From Training</th>
<th>28-Day</th>
<th>% Change From Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>100%</td>
<td>75%</td>
<td>-25.3%</td>
<td>75%</td>
<td>-24.7%</td>
<td>84%</td>
<td>-16.0%</td>
<td>96%</td>
<td>76%</td>
<td>20.1%</td>
<td>85%</td>
<td>10.8%</td>
<td>81%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>99%</td>
<td>74%</td>
<td>-25.7%</td>
<td>82%</td>
<td>-17.0%</td>
<td>75%</td>
<td>-24.7%</td>
<td>95%</td>
<td>72%</td>
<td>22.9%</td>
<td>70%</td>
<td>25.0%</td>
<td>69%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>100%</td>
<td>61%</td>
<td>-38.5%</td>
<td>74%</td>
<td>-25.3%</td>
<td>59%</td>
<td>-40.6%</td>
<td>88%</td>
<td>59%</td>
<td>29.2%</td>
<td>69%</td>
<td>18.4%</td>
<td>49%</td>
<td></td>
</tr>
</tbody>
</table>

### Average Latency (in seconds)

<table>
<thead>
<tr>
<th>Character Set</th>
<th>Immediate</th>
<th>7-Day</th>
<th>% Change From Training</th>
<th>14-Day</th>
<th>% Change From Training</th>
<th>28-Day</th>
<th>% Change From Training</th>
<th>Immediate</th>
<th>7-Day</th>
<th>% Change From Training</th>
<th>14-Day</th>
<th>% Change From Training</th>
<th>28-Day</th>
<th>% Change From Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.0</td>
<td>3.0</td>
<td>-51.6%</td>
<td>3.2</td>
<td>-61.6%</td>
<td>2.9</td>
<td>-47.5%</td>
<td>4.8</td>
<td>6.0</td>
<td>-25.8%</td>
<td>6.1</td>
<td>-28.5%</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.1</td>
<td>3.7</td>
<td>-78.7%</td>
<td>3.1</td>
<td>-46.7%</td>
<td>3.2</td>
<td>-54.3%</td>
<td>5.3</td>
<td>6.5</td>
<td>-24.0%</td>
<td>6.4</td>
<td>-21.1%</td>
<td>5.3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.3</td>
<td>3.7</td>
<td>-60.9%</td>
<td>2.9</td>
<td>-26.1%</td>
<td>3.8</td>
<td>-65.2%</td>
<td>5.3</td>
<td>5.9</td>
<td>-11.3%</td>
<td>6.1</td>
<td>-15.1%</td>
<td>5.6</td>
<td></td>
</tr>
</tbody>
</table>
Appendix K

Average Percent Correct and Latencies for All Participants on All Retention and Generativity Probes Based Upon the Order in Training
Average Percent Correct and Latency for Nine Participants on all Retention and Generativity Probes Based Upon the Order in Training

<table>
<thead>
<tr>
<th>Order in Training</th>
<th>Retention Probes</th>
<th>Generativity Probes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent Change From Training</td>
<td>Percent Change From Training</td>
</tr>
<tr>
<td>First</td>
<td>99% 79% 20.5% 84% 15.3% 83%</td>
<td>89% 75% 14.6% 86% 3.5% 79%</td>
</tr>
<tr>
<td>Second</td>
<td>100% 70% 29.5% 75% 24.6% 73%</td>
<td>95% 74% 20.5% 76% 19.1% 64%</td>
</tr>
<tr>
<td>Third</td>
<td>100% 60% 39.6% 73% 27.1% 61%</td>
<td>95% 58% 37.2% 63% 31.6% 56%</td>
</tr>
</tbody>
</table>

Average Latencies (in seconds)
BIBLIOGRAPHY


Binder, C. (1979b, May). Response rate measurements in a mediated transfer paradigm: teaching severely retarded students to read. Paper presented at the meeting of the Association for Behavior Analysis, Dearborn, MI.


