# THE EFFECTS OF RESTRICTING THE RESPONSE SPACE AND SELF-EVALUATION ON LETTER QUALITY IN BEGINNING AND EXPERIENCED HANDWRITERS

Julia LePage, B.S.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

December 2001

APPROVED:

Jesus Rosales, Major Professor
Janet Ellis, Committee Member
Sigrid Glenn, Committee Member and, Chair of Graduate Studies in Behavior Analysis
David W. Hartman, Dean of the School of Community Service
C. Neal Tate, Dean of the Robert B. Toulouse School of Graduate Studies

LePage, Julia, The effects of restricting the response space and self-evaluation on letter quality in beginning and experienced handwriters. Master of Science (Behavior Analysis), December 2001, 68 pp., 1 table, 23 illustrations, references, 38 titles.

This study analyzed the effects of restricting the response space and self-evaluation on students' handwriting quality in two beginning handwriters and two experienced handwriters. Students executed letters with and without using a transparent overlay, in a multiple-baseline-across-letters design. The use of the transparent overlay included drawing letters in a space restricted by the transparency; overlaying a model letter on top of the written letter and; evaluating if the two letters matched. Letter quality immediately improved when overlays were used, and handwriting quality maintained when the writing response was not restricted by the overlay transparency. Prompting and feedback were delivered contingent on on-task behavior. Analysis was based on three different measurement systems.

ACKNOWLEDGMENTS

I would like to thank Dr. Jesus Rosales-Ruiz for providing an opportunity to conduct this research as well as his guidance and support regarding this thesis. I owe a special thanks to Dr. Janet Ellis for all of her invaluable contributions during my entire graduate career. She has provided and guided me through numerous hands-one opportunities with endless support. I would also like to thank Dr. Sigrid Glenn for agreeing, more than once, to take the time to serve on my thesis committee. Finally, my deepest gratitude goes to my husband Kyle, for his patience, encouragement, and technical contributions throughout this entire process.

TABLE OF CONTENTS


| | Page |
|---|---|
| ACKNOWLEDGMENTS | ii |
| LIST OF TABLES | iv |
| LIST OF ILLUSTRATIONS | v |
| Chapter | |
| 1. INTRODUCTION | 7 |
| 2. METHODS | 11 |
| 3. RESULTS | 22 |
| 4. DISCUSSION | 29 |
| APPENDIX A | 38 |
| APPENDIX B | 40 |
| REFERENCE LIST | 64 |

LIST OF TABLES

Table Page

1. Reliability Scores....................................................................................................39

LIST OF ILLUSTRATIONS

Figure Page

1. Sample of letters-of-name worksheet........41

2. Digital picture of self-evaluative overlay, containing the letter O........42

3. Digital picture of self-evaluative overlay, containing the letter O, with pre-cut, 3-lined paper strip threaded through........43

4. Sample of sheet with boxes, where $10^{th}$-letter sample submissions were traced........44

5. Digital picture of 1-, 2-, and 3 mm evaluative overlays, containing extended outlines of the letter a........45

6. Sample of data sheet used to record criteria met when using Helwig's measures........46

7. Sample of data sheet used to record criteria met when using all-or-none measures........47

8. Participant 1. Digitally scanned tracings of the $10^{th}$ letter executed during 10-letter sample submissions collected in baseline, overlay training and fading conditions........48

9. Participant 2. Digitally scanned tracings of the $10^{th}$ letter executed during 10-letter sample submissions collected in baseline, overlay training, and fading conditions........49

10. Participant 3. Digitally scanned tracings of the $10^{th}$ letter executed during 10-letter sample submissions collected in baseline, overlay training, and fading conditions........50

11. Participant 4. Digitally scanned tracings of the $10^{th}$ letter executed during 10-letter sample submissions collected in baseline, overlay training, and fading conditions........51

12. Participant 1. Percent criteria met per session, during baseline, overlay training and fading, using Helwig's measures........52

13. Participant 2. Percent criteria met per session, during baseline, overlay training and fading, using Helwig's measures........53

14. Participant 3. Percent criteria met per session, during baseline, overlay training and fading, using Helwig's measures........54

15. Participant 4. Percent criteria met per session, during baseline, overlay training and

# CHAPTER 1

# INTRODUCTION

Some literature suggests that quality handwriting can be classified as a behavioral cusp (Rosales-Ruiz and Baer 1997). That is, "…any behavior change that brings the organism's behavior into contact with new contingencies that have even more far-reaching consequences" (p. 533). It has been suggested that teachers' perceptions of students' academic abilities shift based on subjective analyses of handwriting quality (Graham, Harris, & Fink, 2000). According to Graves (1984, p. 69), "If handwriting has a poor appearance, the writer is judged poorly by our culture." Barbe, Lucas, and Wasylyk (1984, p. 3) highlight some of the detrimental effects of illegible or barely legible handwriting: "It reduces employability, decreases self-esteem and the respect of others, contributes to underachievement in related academic participants, and costs business and government millions of dollars annually." Similarly, Arena (1984) points out that good handwriting increases students' pride and their ability to communicate, while poor handwriting limits academic expression, inhibits spontaneous productivity, and affects communication. Berninger (1999) and Ramming (1968) note that poor writing slows the flow of ideas and the expression of thought.

Other authors find a positive relation between handwriting and spelling and reading (e.g., Milone, Wilhide, & Wasylyk, 1973; Strickling 1973). Hildreth (1963) reported six ways that early writing aids reading: 1) writing is an active motor response that evokes visual word patterns; 2) learning to write acquaints beginners with their ABC's; 3) writing aids phonetic awareness; 4) writing helps students recognize words; 5) writing creates self-constructed reading material; and 6) writing reinforces left-to-right reading orientation.

Given the consequences that good and bad handwriting can afford for the individual's other behavior, it is unfortunate that systematic instruction of handwriting is not commonplace. Groff (1984) cited the frequent exclusion of handwriting training from language arts textbooks. McMenamin and Martin (as cited in, Stempel-Mathey & Wolf, 1999) offered insufficient time as an explanation for the lack of formal handwriting training programs. O'Brien (1984, p. 121) attributed the lack of focus on formal handwriting instruction to budget constraints, "…Corner-cutting school systems pared grade-school penmanship teachers from the payroll in the depressions years of the 1930's and never put them back".

Although handwriting instruction is not formally systematized, several instructional techniques have been investigated when training handwriting. These include: modeling and copying, drilling, tracing, use of specialized instruments and materials, and self-evaluating methods.

Modeling and copying include requiring students to write letters after model letters have been presented. Examples of modeling and copying could include a teacher writing the alphabet on the board and students being instructed to copy the letters at their desk, or providing handwriting worksheets with model letters and spaces available for students to copy them (e.g., Hopkins, Schutte, & Garton, 1971; Salzberg, Wheeler, Devar, & Hopkins, 1971; Trap, Milner-Davis, Joseph, & Cooper, 1978). Modeling and copying is perhaps one of the oldest techniques used to teach handwriting. Richardson (1995) noted that Plato (428 B.C.-348 B.C.) used these methods when instructing his students in handwriting development. Modeling and copying have been noted as the preferred handwriting instructional technique of elementary teachers according to Birch and Lefford, (as cited in Fauke, Burnett, Powers, & Sulzer-Azaroff, 1973).

Drilling (or practicing) is defined as repetitious writing of letters. An example may include students writing each letter of the alphabet 10 times. This handwriting teaching technique is often considered as traditional and necessary (Lamme, 1984; Ramming, 1968). Herrick and Okada (1963) named drilling as the second most preferred (following modeling and copying) instructional method of elementary school teachers.

Tracing includes writing over model letters for practice or improving letter formation (1984). This technique has a history of frequent use (Askov & Greff, 1975; Fauke et al., 1973; Graham & Miller, 1980; Hirsch & Niedermeyer, 1973).

Tool utilization involves using specialized handwriting instruments and materials. Included are various types of pens and pencils (Krzesni, 1971), specific paper preparations (Halpin & Halpin, 1976; Skinner & Krakower, 1968), and individualized manipulatives, including transparent evaluative overlays (Jones, Trap, & Cooper 1979; Robin, Armel, & O'Leary 1975; Trap et al. 1978).

Finally, self-evaluative techniques have been documented (Furner, 1969; Graham & Miller, 1980; Jones et al. 1977). While self-evaluating, students are judging the quality of their own writing responses. An example of this technique might include a student comparing his/her written letter with a model letter. Self-evaluation can be used in conjunction with other teaching methods such as modeling and copying (Robin et al. 1975; Trap et al. 1978).

This study researches the effects of a particular teaching method to improve handwriting quality of novice and experienced handwriters. The method consisted of a 3-component package, including: 1) self-evaluation using model letters imposed on a transparent overlay, 2) restricted response space within the overlay, and 3) delivery of on-task praise and off-task prompts.

# CHAPTER 2

# METHODS

## Participants and Setting

Four male students, enrolled in an after school-tutoring program, served as participants. They included 2 right-handed, elementary school students ages 8 and 9, 1 left-handed pre-school student, age 5, and 1 right-handed kindergarten student, age 5. None of the participants was diagnosed with a developmental or learning disability. The inclusion criterion was teacher recommendation based on quality of penmanship. The 5-year-old students were beginning writers, hereafter referred to as Participants 1 and 2; the elementary students, more experienced writers, are referred to as Participants 3 and 4.

Handwriting samples were collected as part of an after-school tutoring program, conducted in a set aside area in a community housing project, Monday through Thursday, from 3:00-5:00 p.m. This room contained two long tables, computers, and two chairs. A section of one table was cleared for handwriting training. The primary investigator sat alongside each student when samples were collected. Students submitted daily samples; submissions took approximately 15-30 min.

Apparatus

Materials included letters-of-name worksheets, #2 pencils, self-evaluative plastic overlays, and pre-cut 3-lined paper strips. Production descriptions for worksheets and overlays follow.

Letters-of-name worksheets. Worksheets displayed each participant's name-letters aligned on the left side of each 3-lined segment, with ten, 3-lined segments per page. Necessary segments contained models of each letter in the child's name in Arial font, size 36, (e.g. an Owen page contained 4 letters written vertically, on one page; a Richard page contained 7 letters written vertically on one page. See sample in Figure 1.

Overlays. A set of laminated transparent overlays consisted of the following: an overlay for each letter in the participant's name. The model letter on each overlay consisted of a computer-generated letter in the participant's name. Each computer-generated letter was in Arial font, size 36. The first name-letter was capitalized, the rest were lowercased. The number of overlays required for each name depended on the number of letters in each child's name (e.g., 4 overlay cards for Owen; 7 overlay cards for Richard). Each laminated overlay card contained 3 lines with the model letter imposed within the 3 lines. Additionally, each laminated overlay contained a 1 in. vertical slit cut within 1 in. of the horizontal ends of each overlay card. Square holes, approximately 3/8ths in. wide x 3/8ths in. long were cut out approximately ½ in. from the model letter. See Figure 2.

Paper strips. Pre-cut, 1 in. tall, 3-lined paper strip segments were threaded through the 1 in. vertical slits on each transparent overlay card. See example in Figure 3.

## Behavioral Measures and Recording Procedures

Measures of Handwriting Quality

Handwriting quality was defined as the degree of sample-letter form approximation toward model-letter forms. Three different systems evaluated letter quality. They are: 1) letter tracing evaluation, 2) Helwig's (1976) measures, and 3) all-or-none overlay matching. During all phases of the experiment, samples included each name-letter written 10 times, only the 10th submitted letter per 10-letter sample was analyzed. All 3 scoring methods analyzed the same data.

Letter tracing evaluation. Sample letters were traced for presentation purposes in order to separate the 10th letter submitted. Letters were traced on paper lined with boxes 1 in. wide x 1 ¼ in. long. Each page contained 60 boxes, 10 across, and 6 down. See sample in Figure 4, (due to margin constraints, sample is not true to size). Lines defining the bottom of the 3-lined paper segments, where samples were written, were aligned with the bottom of presentation boxes when sample letters were traced. This alignment allowed for representation of sample letter size and spacing. Along with size and spacing, traced letters were evaluated according to edge smoothness, letter element size and placement, and slant consistency.

Helwig's measures. Helwig, Johns, Norman and Cooper's (1976) handwriting quality assessment technique consists of 6 different criteria. An evaluative overlay is necessary to determine correct responses according the first of the 6 criteria. The evaluative overlays used in the present study were computer-produced letter outlines copied onto a transparency and laminated. Each evaluative overlay contained an outline of a particular letter. The outlines extended the interior and exterior borders of all Arial font, size 36 model letters by 1mm, and were imposed on a 3-line segment. See top sample in Figure 5. Three-lined segments on evaluative overlays were aligned with the 3-line segments present on the submitted sample papers. These lines allowed for correct vertical and horizontal placement by aligning the bottom line on both ends of the plastic evaluative overlay with the bottom horizontal line of the sample papers. The evaluative overlay could be moved horizontally across a 3-line segment. This maneuver, while remaining aligned, allows for inclusion or occlusion of the sample letter within the confines of the extended model letter portrayed on the evaluative overlay.

Letter quality was evaluated following Helwig et al.'s 6 criteria: 1) The total stroke must be within the confines of the boundary lines of the evaluative overlay: 2) each stroke that is not a complete circle must begin and end between the bottom line of the 3-lined segment and in the line forming the confines of the letter; 3) all circles in the letters a, b, d, g, o, p, q and the top of the letter e must be closed curves; 4) all strokes must intersect each successive stroke at one point except for the dot above the lower-case i and j; 5) each letter must be complete with all strokes present; 6) the horizontal stroke in the t and f must intersect the other stroke within the confines of the ellipse near the center of the vertical stroke (Helwig et al. 1976; Trap et al. 1978). A recording sheet was developed listing the model letters contained in each student's name. See sample in Figure 6. Each letter included relevant criteria, with blanks to record the strokes meeting criteria. These criteria are differentially applicable to each letter of the alphabet. For example, according to this system a correctly executed letter "m" must meet 3/6 of

Helwig's measures; whereas, a correctly executed letter "e" must meet 4/6 criterion measures. Individual alphabet letters in upper and lowercased forms also must meet different criterion measures to be scored as correct.

All or none. Additional evaluation of letter quality again included the use of extended model letter outlines on evaluative overlays. The model letters were extended 1, 2 and 3 mm. Figure 5 shows 1, 2, and 3 mm evaluative overlays, respectively from top down, for the model letter "a". Sample letters were scored as correct if they fit within the extended model letter border (including the lines creating the border) when the 3-line segments of paper and overlays were aligned. No other criteria were necessary for a correct score. If any part of the letter was outside the boundary of the overlay the letter was scored as incorrect. These measures are referred to as all-or-none. This type of evaluation procedure utilizes transparent overlays and standardized verbal criteria to measure handwriting samples (Fauke et al. 1973; Helwig et al. 1976; Jones et al. 1977; Kau-To Leung, Treblas, Hill, & Cooper 1979; Robin et al. 1975; and Trap et al. 1978). Graham (1986) found that correct-incorrect (or all-or-none) handwriting measurement procedures met two essential criteria: different raters were able to use the same measures reliably, and a single rater successfully used the same measure on different occasions. See Figure 7 for a sample of an all-or-none data sheet. Separate scoring sheets were used for each participant when the 1, 2, and 3 mm extended model letter evaluative overlays included or excluded 10th letter sample submissions.

Additional Measures

Task timing. The amount of time participants took to complete each 10-letter sample submission, per name-letter, was recorded. Students were not made aware of these recordings. After students were instructed, they began writing a name-letter 10 times. Task timing began when students placed their pencils onto their papers and continued until the 10th version of the name-letter was written. A timing example includes the experimenter starting the timer when a student began writing the first of 10 “e”(s). The timer was stopped when the 10th “e” was completed. Thus the number of task timing measures was dependent upon the number of letters in a participants name; if the student’s name contained 5 letters, 5 task timing measures were collected per day. For simplicity of presentation, data were presented in averages, per condition, including the letters that were in training, the letters kept in baseline, as well as letters written during fading procedures. If a student’s name contained 5 letters and 3 name-letters were being trained, while the remaining 2 name-letters were written under baseline conditions, averages were deducted as follows. The task timing (min and s) of the 3 letters in training were added and divided by 3, yielding an average task timing of 10-letter sample submissions made during training sessions for that day. The task timing (min and s) of the 2 letters remaining in baseline were added and divided by 2, yielding an average task timing of 10-letter sample submissions made during baseline sessions for that day. As the experiment progressed, an increasing number of name-letters were included in training, until finally no name-letters remained in baseline.

Prompts and praise delivery. Praise for continued responding and self-evaluation, as well as, “keep-going” prompts delivered by the experimenter for each letter task were recorded. Continued responding praise statements included, “You are a hard worker today;” “Good job finishing the “r”(s).” These statements were contingent on continued on-task responding, and were administered approximately every 30s. For example, the participant continues writing when submitting a 10-letter sample that takes 50s to execute; at least one continued responding praise statement was delivered. These types of statements were also delivered if students requested assistance. Assistance request examples include, “Do I write it like this?” or “Am I doing good?” Overall task completion praise statements were always delivered and not recorded. Accuracy prompts and praise were not delivered. “Keep-going” prompts, such as, “Let’s finish the “h”(s),” were delivered as needed. For example if a student put down his pencil, sat on the floor and repeatedly tied his shoe, a “keep-going” prompt was administered. Prompts to self-evaluate, as well as praise (e.g. “Did you check to see if it matches?” and “I like the way you are checking all of your “e”(s)!”) were also delivered, contingent on self-evaluative overlay manipulations. Such manipulations involved a student adjusting the model letter portrayed on the overlay atop of his written letter and declaring a match or a non-match. If self-evaluations did not occur, prompts were intermittently delivered; if self-evaluations did occur, praise was intermittently delivered. Intermittent delivery of these experimenter statements occurred approximately every 2-3 letters if prompts were delivered, and at least every 30 s if praise was delivered.

Interobserver agreement. Observers were trained to use the evaluative overlays by measuring 100 letters. An independent observer compared results from each sample with the experimenter's original evaluation. Reliability of the scoring procedures was determined by calculating the percent agreement between two scorers. Percentage agreement was calculated by dividing the number of agreed criteria met by the total number of criteria available per sample and multiplied by 100.

Reliability checks for Helwig's measures and the all-or-none scores ranged from 89% to 100% across all 4 participants. See results in Table 1.

Procedures

Baseline. During baseline, participants were seated at the table to the right of the experimenter. Students were handed 1 letters-of-name worksheet, and given the following instructions, "You are going to write each letter 10 times, take your time and do the best you can. Look at the model on the left and try to make yours look like that one." Once initial responding occurred, timing began and continued until the student had written a name-letter 10 times. If continued responding occurred, occasional praise statements were delivered. If pauses in writing occurred, students were prompted to "keep-going." Number of continued responding praise statements and "keep-going" prompts delivered were documented per 10-letter sample, per participant. Accuracy prompts and accuracy praise statements were not delivered.

Overlay training. Participants were given a 3-lined paper strip and the self-evaluative plastic overlay containing the appropriate model letter. The experimenter gave instructions and modeled the use of the overlay, "You are going to write each letter 10 times, take your time and do the best you can. Look at the model and try to make yours look like that one. This is how you use the overlay." Then the experimenter demonstrated writing in the open box and sliding the paper strip through the plastic card, allowing for an overlay of the model letter atop the written letter. Students were directed to begin writing on the specified letter(s). Statements of prompt and praise were delivered under the same contingencies described in baseline.

Next, self-evaluative instructions were given and modeled. Examples included, "What is different about your letter and this letter?" "What is the same?" "How can you make yours fit this one?" Self-evaluation, defined as manipulating the overlay and comparing written letters with model letters and labeling matching and non-matching components was also prompted and praised. Self-evaluation examples include a student adjusting the model letter on the self-evaluative overlay atop of his written letter and commenting, "It is too big," or "The "D" fits!" Upon all occurrences of these self-evaluative statements, the experimenter delivered praise statements, such as, "You're right, and I really like the way you checked that letter," or "Good job checking all of those "e"(s)!" If students failed to self-evaluate after writing approximately 2-3 letters, prompt statements like, "Did you check to see if those matched?" were delivered.

Overlay training began with the first name-letter of each student's name and continued until all name-letters were executed. Only 1 letter was introduced to overlay training at a time. When visible improvement for the first letter was achieved, training began on the second letter, and so on, until all letters of the name were trained with the overlay. Letters that had not yet been trained were practiced by repetition writing, 10 times each, on letters-of-name worksheets (a continuation of baseline procedures). So, every day consisted of writing each name-letter 10 times each, whether under baseline or training conditions.

Overlay fading. During fading procedures, pre-cut, 3-lined paper strips were not used, students submitted all samples on the letters-of-name worksheets. Fading consisted of 3 steps carried out over 3 days. Step 1 involved continuing the use of the restricted response space in the self-evaluative overlay, by placing it atop the 3-line segments of the letters-of-name worksheets rather than sliding pre-cut 3-line strips through the overlay. During step 1, students were instructed to self-evaluate after every letter written. After each 10-letter sample submission, students were asked, “How many matched?” and “Which is your best one?” When completing step 2, students were instructed to write a letter 10 times without the restricted response space of the self-evaluative overlay and were given no specific instructions pertaining to self-evaluation; though the plastic overlay was available for self-evaluation purposes. After each 10-letter sample submission, students were asked, “How many matched?” and “Which is your best one?” The 3rd fading step included instructions for students to write each letter 10 times and to “Pick your best one,” without the use of the self-evaluative overlay. Praise and prompt statements were delivered under the same contingencies described during baseline.

Design

The primary independent variable was the 3-component package, consisting of: 1) self-evaluation using model letters imposed on a transparent overlay, 2) restricted response space within the overlay, and 3) delivery of on-task praise and off-task prompts.

A multiple-baseline across-letters design was used in a single participant format. Each child participated in the following sessions in this order: 1) baseline: letters-of-name worksheets; 2) overlay training and remaining name-letters on letters-of-name worksheets; and 3) overlay fading. Students began training with the first letter of their name and continued on baseline conditions for the remaining letters. When the letter tracing evaluation measures indicated improvement in letter formation for the first overlay-trained letter, training was initiated for the second letter of the child's name, while baseline responding continued for the remaining letters. This continued until all letters had been trained. Graphical depictions of this design are illustrated by result figures.

# CHAPTER 3

# RESULTS

Figures 8-11 show the results rendered by the letter-tracing evaluation measures of handwriting quality across letters for each participant. These figures illustrate tracings of the 10th letter submitted during a sample. If fewer than 10 letters were submitted, the last letter was traced and displayed. Each traced letter has been digitized and reduced through computer scanning to display continuous sample responding.

## Letter Tracing Evaluation

Figure 8 shows the results for Participant 1. During baseline, letter drawings were oversized, with jagged edge strokes, inconsistently slanted, and sometimes having malformed/missing/misplaced letter strokes. During overlay training, letters were adequately sized, with smother edged strokes, consistently slanted, and formed with all relevant letter elements present and correctly placed. During fading procedures, samples submitted began to show a slightly inconsistent slant relative to training conditions, though visual comparison of letters indicate a maintained quality improvement over baseline submissions.

Figure 9 shows the results for Participant 2. Letters written during baseline are over and undersized, relatively smooth edged, but inconsistently slanted with incorrect letter element placement. Examples include letters written too big to be displayed, the letter "i" written with the dot beneath the line, and consistently backward versions of the letter “s”. When training procedures were implemented, immediate improvement occurred across all letters pertaining to size, slant, and edge smoothness. Letter element placement also improved, and the letter “s” was no longer reversed. Improvement continued through fading procedures.

Figure 10 shows the results for Participant 3. Baseline letter submissions across all letters are adequately sized, with smooth edged strokes, although many letters were inconsistently slanted, and while all letter elements are present, they oftentimes are incorrectly placed. An example includes the letter “n” with the half-curve meeting the straight line at the bottom of the line rather than at the top. Visual comparison of letters written during training, show an improvement in slant consistency and letter-element placement. Fading samples continued to maintain the improvement gained during training procedures for Participant 3.

Figure 11 shows the results for Participant 4. Baseline letters are adequately sized, with relatively smooth edged strokes, although pre-training letters were commonly inconsistently slanted, and showed incorrect letter element placement. Examples include the letter “e” with open curves or added strokes, the letter “b” without the bottom portion of the straight line, and the letter “s” with straight lines rather than curves. Letters written during training sessions indicate direct improvement across all letters. These letters remained adequately sized, while improving in edge smoothness, slant consistency, and correct letter element placement. Visual comparison of letters submitted during fading procedures demonstrated maintained improvement across these areas of evaluation.

Helwig's Measures

Figures 12-15 show the results of measuring handwriting samples using Helwig's (1976) criteria checklist. Results are graphed using the percentage of criteria met per letter. Again, only the 10th or last submitted letter per sample was measured (same letters illustrated in Figs. 8-11). High percentages indicate model-letter approximation.

Figure 12 shows Participant 1's results. During baseline, the percentage of criteria met was variable across all letters, except for the letter "u". When training began, letter production increased and maintained improved percentages for the letters "D", "u", and "a". Percentages for the letters "e" and "u" were variable but overall did not increase. Percentages of criteria met for the letter "n" were also variable but ultimately fell below baseline levels. Results of fading procedures yielded consistent responding for the letters "D", "e", and "q", variable responding with improvement for the letter "u", and variable responding with decreased percentages for the letters "a" and "n".

Figure 13 shows results for Participant 2. Baselines percentages varied across all letters with the exception of the letter "s". During training, increased percentages were scored for 3 letters, "l", "i", and "s", however, "U" and "e" did not present increased percentage results above baseline levels. Fading procedures yielded maintained criteria percentages for the letters "l" and "i". Two of 3 fading samples from the letter "U" submissions scored lower percentages than those achieved during training procedures. The "s" and "e" scored decreased percentages of criteria met relative to percentage of criteria met during training conditions.

Figure 14 shows results for Participant 3. Baseline percentages varied across all letters with the exception of the letter “M”. During training, percentages met or exceeded baseline levels for the letters “M”, “a”, “n” and “e”. Percentages continued to vary for the remaining letter trained, “u”. Results of fading procedures demonstrated decreased percentage of criteria met for the letters “M” and “e”, and stable responding for the remaining letters, “a”, “n”, and “u”.

Participant 4’s results are shown in Figure 15. During baseline, percentage of criteria met was variable across all letters, though only one session varied for the letter “b”. Upon training, percentages exceeded overall baseline levels for the letters “s” and “t”. The measures of the remaining letters, “S”, “e”, “b”, and “a”, remained virtually unchanged from baseline percentages. Fading percentages were consistent with training.

All or None

Figures 16-19 show the results using the all-or-none measurement system in the form of cumulative records. Results using the 3 levels of model letter outline extension, 1, 2, and 3 mm, are collectively displayed per participant. 1 mm results are shown vertically on the left portions of the figures. 2 mm results are shown vertically in the center of the figures, and 3 mm results are shown vertically on the right portions.

Figure 16 displays Participant 1’s results. When the 1 mm all-or-none criterion was used, very little response change or improvement occurred, across all phases. However, when the 2 mm all-or-none criterion was used to measure letter quality, responding was stable during baseline, and improved immediately after training procedures were implemented. This improvement continued though fading conditions across all letters with the exception of the letter “e”. With the 3 mm measurement criterion, results were similar to those measured under the 2 mm criterion.

Figure 17 show the all-or-none results of Participant 2. Correct responding correlates with training procedures across all 3-measurement levels. Baseline conditions indicated zero variability when the 1 mm criterion was in place. When training procedures began, correct responding increased, and continued through fading procedures for the letters “l”, and “i”. Similar results were indicated for the 2 mm and the 3 mm criteria levels, although some correct responding occurred during baseline for both these measurement criteria. Correct responding during fading maintained across all letters except for the letter “U”.

Participant 3’s cumulative record results are shown in Figure 18. With the strictest criterion level, 1 mm, baseline samples measured ‘correct’ for at least some submissions of all letters, except for the letter “M”. However, across all letters but, “l”, graph dimensions change, indicating correct responding when training began. Correct responding continued through fading for the letters “a”, “n” and “u”, though not for the letters “M”, “e”, and “l”. The 2 mm measurements for Participant 3 show little response variation regardless of experimental condition. The 3 mm measures scored 100% correct during baseline, training, and fading procedures.

Figure 19 shows Participant 4’s results. The 1 mm criterion level results indicate relatively stable baseline responding across all letters. Correct responding was indicated across all letters after training implementation and maintained through fading for all letter responses except the letter “t”. When using the 2 mm criterion measures, all letter “S” and “e” responses were scored as correct. The remaining letter samples, “b”, “a”, “s”, and “t” all began to score correct when training procedures began, and continued through fading. Similar results are demonstrated with the 3 mm criterion level measures.

<u>Additional Measures</u>

Figures 20-23 show the average 10-letter sample completion times and the average number of prompt and praise statements delivered during baseline, training, and fading procedures. Each participant’s task timing averages and prompt and praise statement delivery averages are depicted together on one page, with the top graphs displaying task timing averages, and the bottom graphs displaying prompt and praise statement delivery averages. Those sessions in which times and/or prompts and praise were not recorded have been omitted.

Figure 20 shows Participant 1’s results. Task time averages ranged from 10 s-32 s when submitting baseline samples, and from 1min 25 s-5 min when submitting samples during overlay training. Number of prompts and praise statements delivered while Participant 1 wrote letters under baseline conditions ranged from averages of 0-2.3 per 10-letter sample. The average range during training procedures was from 1.7-10 prompts and/or praise statements delivered per 10-letter sample.

Participant 2’s results are shown in Figure 21. Baseline task times ranged from averages of 20 s-1 min 1 s. Task times during training sessions ranged from averages of 2 min 8 s-3 min 40 s, while completing 1, 10-letter sample. Prompts and praise statements during baseline writing were delivered on average from 1-2 times, and from 2.2-7 times during training sessions.

Figure 22 shows Participant 3's results. Task timing during baseline writing ranged from 5 s-20 s. Letter submission averages during training sessions ranged from 59 s-2 min 3 s. Prompt and/or praise statement deliveries occurred infrequently during baseline and training sessions. Baseline prompt and praise statements were not delivered, while training prompt and praise statements average deliveries ranged from 0-2.5.

Participant 4's results are shown in Figure 23. The range of task timing averages was 6.5 s-16.5 s for baseline letter writing and 51 s-1 min 11 s for letters written during training sessions. There were no prompts or praise statements delivered during this student's baseline samples. Prompt and/or praise statement delivery during training sessions averaged from 0-3 during completion of 1 10-letter sample.

CHAPTER 4

DISCUSSION

In general the results show that the 3-component handwriting teaching method can have an immediate positive effect on student handwriting quality, for both beginning and experienced handwriters. This improvement, defined as model letter approximation, is evident in the letter tracing evaluation outcomes. Results of letter tracing evaluation also show an immediate correction of letter reversals displayed by one participant.

The baseline procedures of the present study resemble the procedures of modeling and copying, and drilling, except for the absence of external feedback with respect to the accuracy of response. While both modeling and copying, and drilling, have been supported in the literature (Hildreth, 1963; Hirsch & Niedermeyer, 1973; Lamme 1984; Ramming, 1968) and both have been noted as preferred handwriting teaching methods by elementary teachers (Herrick & Okada, 1963), current results do not indicate that these methods aid in the improvement of handwriting quality. Participants' quality of handwriting did not improve during baselines of up to 230 letters written. However, quality did not deteriorate as Arena (1968) and Barbe (1984 p. 245) suggest: "The child who writes a letter fifteen times will only make the letter less well the fifteenth time than he did the first."

Previous research has demonstrated the advantages of using evaluative visual overlays over other methods to deliver feedback with respect to the accuracy of response

(Fauke et al. 1973; Robin et al. 1975; Trap et al. 1978). Fauke et al. showed the improvement of handwriting across 8 sessions from baseline lows of 5% correctly written responses to training highs of 90% correctly written responses, for a single, 6 year-old participant. Robin et al. showed improvements in mean handwriting test scores from 108.8 to 158.9, with a group of 10 students, aged 5-6 years, across 20 sessions. Results from Trap et al.'s overlay research offer similar results. Across 20 training sessions, the percentage of correctly trained letter strokes increased from an average baseline score of 23.95% to final intervention scores of 71.68% for 12 students, aged 6-7 years.

The present study showed dramatic improvements during a 10-trial session. However, we do not know how much improvement is attributable to the self-evaluative overlay procedure. Future research is necessary to determine the relative contributions of the overlay, the response restriction, and feedback and correction procedures.

Comparison of the present overlay procedure with previous ones is difficult. Significant differences between components of past overlay studies and the current overlay study exist. A major component of past overlay research has included experimenter feedback concerning accuracy. Fauke et al. attributed improved handwriting skills to instructional procedures involving self-evaluation with a transparent overlay in addition to edible reinforcement delivery contingent on correct letter formation. Robin et al. and Trap et al. utilized transparent overlays as a tool to provide objective accuracy feedback, from the experimenter. While all of the previous studies utilized model letters on transparent overlays allowing for experimenter evaluation and self-evaluation, the present study used model letters on transparent overlays for self-evaluation purposes only. That is, students judged their own letter quality by declaring matches or mismatches. The experimenter only acknowledged the students' report (or self-evaluation). Interestingly, this study demonstrated that external accuracy feedback is not a necessary training component for handwriting quality to improve when using self-evaluative transparent overlays.

The use of restricted response space, is not prevalent in the handwriting literature. One similar strategy offered by Skinner & Krakower (1968) used special paper that changed ink color when letters exceeded a tolerance model. The findings of the current study, while using a very different kind of response restriction, support the utility of this type of technique. While describing "better ways of solving the 'problem of the first instance,'" Skinner (1968, p. 207) states, "Another solution is to use stimuli which elicit or evoke the response to be reinforced." The restricted response space, included on the self-evaluative overlays, used in this study, appear to coincide with this solution. The available writing space offered by the 3/8ths in. wide x 3/8ths in. long, cut-away box, within the overlay, increases the probability that students will execute letters approximating the size of model letters. This occurrence, in turn, increases the probability that student letter-writing responses will be such that a "match" can be determined upon self-evaluation.

Intermittent delivery of prompt and praise statements were often included in other handwriting studies (Fauke et al. 1973; Robin et al. 1975; Trap et al. 1978). Successful results across experiments indicate this is a valuable component.

While all of the present participants demonstrated increased model letter approximation upon training when letter tracings were evaluated, improvement levels varied dependent upon the type of measurement system used. These results may reveal more about the measurement techniques than they do about the quality of participants' handwriting.

Accurate, valid, and objective handwriting measurement techniques have been debated for many years (Bezzi, 1962; Feldt, 1962; Graham, 1986; Graham, Boyer-Shick, & Tippets 1989; Graham & Miller 1980; Helwig et al. 1976). Donoghue (as cited in Handwriting 1984) states that handwriting scales are often cumbersome and unreliable, thus preventing objective measurement. Findings by Helwig et al. (1976) contradict these arguments by presenting an objective system of measurement for handwriting quality, utilizing six stated criteria that demonstrated high reliability scores. Trap et al. (1978) also used Helwig's method. Graham's work (1986) demonstrated that objective measure could be reliable, by using correct-incorrect measure with evaluative transparent overlays, though these measures were not found to be statistically valid.

The results of this study support Graham's (1986) research. Reliability measure were consistently high across both objective measuring systems utilized, Helwig's measures and all-or-none measures, however, earned criteria scores do not appear to be representative of letter quality when letter tracing evaluations are compared. These objective systems include the 6-criteria list offered by Helwig et al. (1976) and the all-or-none (1-, 2-, and 3 mm correct-incorrect only) criteria (Kau-To Leung et al. 1979; Robin et al. 1975).

Instances of the inconsistencies between letter tracing evaluations and Helwig's measures are apparent across samples for all participants. For example, while Student 1's 6th sample submission of the letter "e" scored ¾ of Helwig's measures, his 20th submitted letter "e" (while clearly an improved "e" when visually compared) met only ¼ of Helwig's criteria. Also, the second letter "e" submitted by Participant 2, only met 1 of the 4 possible criteria, even though it could be described as a high quality letter "e" when visually compared with model letters. Participant 3's 6th sample submission of the letter "a" is hardly legible, though it met ¾ criteria for that letter. Finally, Participant 4 submitted 2 letters that could be described as having poor quality when visually compared, the 6th letter "e", and the 2nd letter "a"; however, according to Helwig's measures both the "e" and the "a" scored as meeting 100% relevant criteria.

Similar inconsistencies were found when utilizing the all-or-none criteria, although they were not as prevalent as the inconsistencies found when using Helwig's measures. Such variances could explain the apparent variable responding in Figures 12-15, even when improved handwriting quality is apparent when letters are compared in Figures 8-11.

These findings suggest that although previously offered objective handwriting assessment tools may offer reliable measures, their use may not be practical. This logic is further explained by Baer (1977) in a reviewer's comment, "Just because it's reliable doesn't mean that you can use it."

Visual Communication includes research into which forms of data best communicate intended messages (1989). Pettersson states, "Realistic pictures can provide reasonably objective documentation of a situation, product or course of events" (p. 143). When examining handwriting quality, actual pictures of samples submitted may provide the best from of objective documentation, compared to quantitative graphical depictions of reliable (albeit invalid) measurement results.

Although letter tracing evaluations indicate that handwriting quality improved across all 4 participants, regardless of experience level, differences between beginning and experienced handwriters were apparent in 3 areas: 1) results based on different measurement criteria levels, 2) task timing, and 3) number of prompts/praise delivered.

When the strictest criteria were applied during 1 mm all-or-none measures, experienced writers' samples appeared sensitive to training conditions. As criteria levels lessened to 2- and 3 mm all-or-none measures, samples met criteria before any training occurred, thus indicating that these students were performing at these levels during baseline. Alternately, one of the inexperienced writers rarely exhibited correct responding when the strictest criteria were used, even after training. Next-level criteria of 2- and 3 mm were necessary before quality of letter changes were detectable, indicating that even after training was introduced, this beginning writing student was not able to meet the strictest criteria. These results address the importance of the validity of a particular measurement system. Although letter change quality may have improved across all participants when letter tracings are evaluated, the quantitative measurement systems may not detect or depict these changes, depending on the criteria used and measures of the initial sample submissions.

A second difference between the novice and experienced writers is demonstrated in the amount of time used during sample submissions. The beginning writers' task timing measures were considerably higher when submitting a 10-letter sample than were the task timing measures of the experienced writers across all experimental conditions. However, all 4 writers task timing results were higher during training conditions relative to their own baseline conditions. So, for experienced and novice writers, the utilization of a self-evaluative overlay with a restricted response space, increased the amount of time necessary to complete a writing task.

Finally, prompt and praise statements were administered differently for the novice and experienced writers. For the experienced writers, there were no prompts and praise delivered during baseline, whereas, baseline delivery of prompt and praise statements ranged from 0-2 for the beginning writers. Prompt and praise delivery during training ranged from 0-3 for the experienced writers and from 1-10 for the inexperienced writers. These results demonstrate that the independent use of a self-evaluative overlay can be extremely effective handwriting improvement tool for older, more experienced writers, and that the tool can be equally as effective in training and sustaining quality performances in beginning handwriters given additional minimal supervision.

Because this study used a 3-component package, it is not possible to determine the contributing features of individual components. Perhaps a component analysis of these features could lead to specific conclusions regarding which are necessary to achieve handwriting quality improvement.

Also, this study did not separate the delivery of continued responding praise statements, "keep-going" prompts, and self-evaluation praise statements. This collective grouping of experimenter statements does not allow for a clear picture of which statements are necessary, or when they might be necessary. For example, if a student finished a 10-letter sample in 15s, no continued responding praise statements or "keep-going" prompts were delivered, because these statements were only delivered if a student was on task for approximately 30s or was off-task. The higher rate of prompt and praise delivery to beginning writers might possibly be correlated with their longer completion times, versus the possibility that longer completion times made increased rates of prompt and praise delivery necessary. In addition, if self-evaluative praise had been measured alone, contingent on occurrences of defined self-evaluative behaviors, specific conclusions concerning the inclusion of the model letter on the transparent overlay might be available. The intended function of the self-evaluative feedback, delivered by the experimenter, was to transform correctly executed letters (those matching model letters) into conditioned reinforcers. Because experimenter statements were not separately measured, their functions cannot be determined.

Finally, maintenance measures were deficient and generalization measures were absent. This study allowed for only 3 fading sample submissions due to the closing of the after-school tutoring program. It remains unclear whether or not improved handwriting results would have maintained an extended period, or whether apparent self-evaluative skills and/or handwriting improvement generalized when students wrote letters outside of experimental conditions, or when writing untrained letters.

Future overlay manipulation studies might focus on the self-evaluative dimensions of the overlay. Students might be taught to pinpoint their own self-evaluative behaviors and specific reinforcement and/or necessary prompting of these behaviors could then be examined. In addition, if self-evaluative skills learned with specific letters are found to generalize to other letters, this could increase the likelihood that educators would implement this tool. Another area of focus for future overlay studies could include researching the common problem of reversed letters. Discrimination of reversed letters might be accelerated using the self-evaluative overlay; because 3-lined paper segments feed into the overlay, if letters were reversed they would also be upside down. Specific dimension discriminations of letter stimuli may be revealed through further examination.

# APPENDIX A

# TABLES

Table 1

RELIABILITY SCORES

| | Student 1 | Student 2 | Student 3 | Student 4 |
|---|---|---|---|---|
| Helwig's Measures | 89% | 97% | 95% | 95% |
| All-or-none 1 mm | 97% | 97% | 89% | 90% |
| All-or-none 2 mm | 99% | 97% | 98% | 99% |
| All-or-none 3mm | 100% | 97% | 100% | 97% |

# APPENDIX B

# FIGURES

Figure 1

R

i

c

h

a

r

d

Figure 2

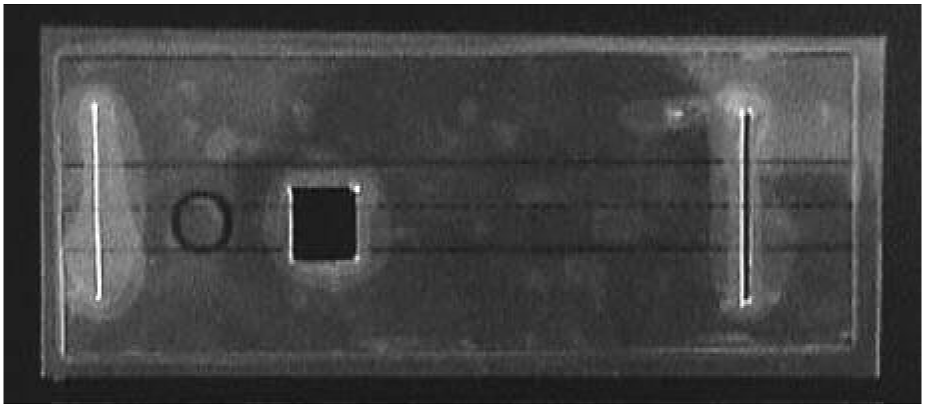


Figure 3

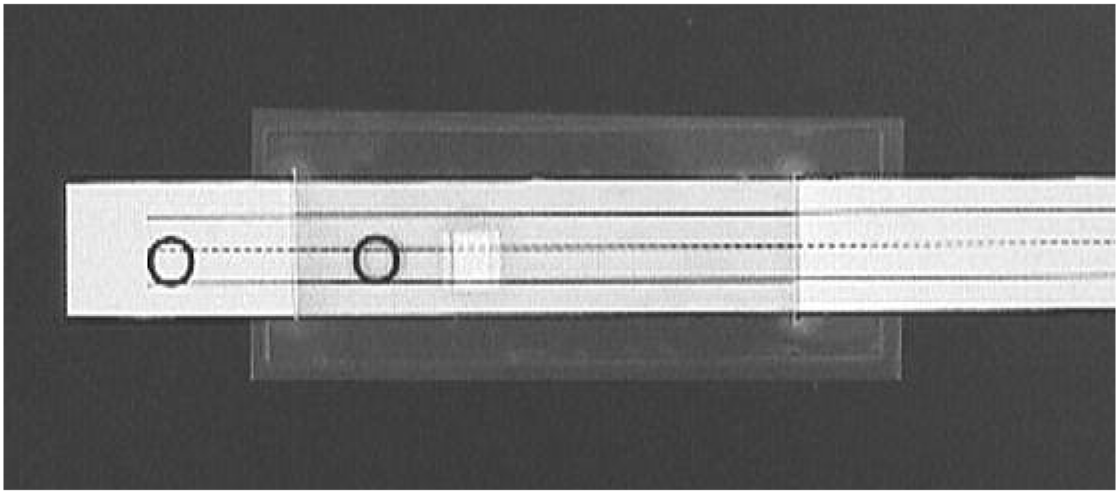


Figure 4

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |

Figure 5

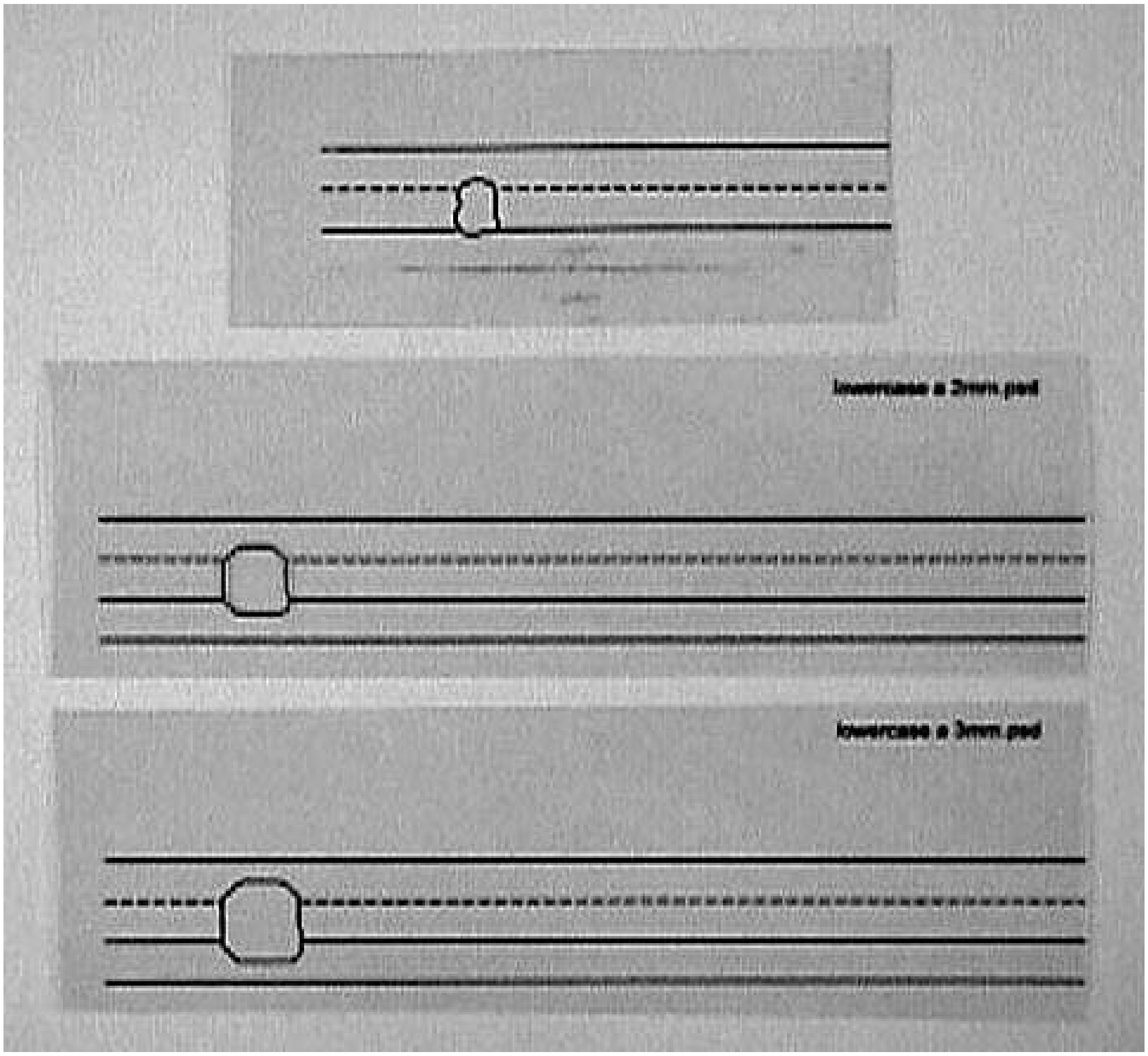


Figure 6

D **LETTER SCORE 4/4**_______

Criterion 1_______

Criterion 2_______

Criterion 4_______

Criterion 5_______

u **LETTER SCORE 4/4**_____

Criterion 1_______

Criterion 2_______

Criterion 4_______

Criterion 5_______

e **LETTER SCORE 4/4**_______

Criterion 1_______

Criterion 3_______

Criterion 4_______

Criterion 5_______

a **LETTER SCORE 4/4**_____

Criterion 1_______

Criterion 3_______

Criterion 4_______

Criterion 5_______

q **LETTER SCORE 4/4**_______

Criterion 1_______

Criterion 3_______

Criterion 4_______

Criterion 5_______

n **LETTER SCORE 4/4** _____

Criterion 1_______

Criterion 2_______

Criterion 4_______

Criterion 5_______

Figure 7

| date | 3/6 | 3/7 | 3/8 | 3/13 | 3/14 | 3/15 | 3/27 | 3/28 | 3/29 | 4/2 | 4/9 | 4/10 | 4/11 | 4/12 | 4/16 | 4/17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| D | | | | | | | | | | | | | | | | |
| e | | | | | | | | | | | | | | | | |
| q | | | | | | | | | | | | | | | | |
| u | | | | | | | | | | | | | | | | |
| a | | | | | | | | | | | | | | | | |
| n | | | | | | | | | | | | | | | | |

ALL = X; NONE = 0

TOLERANCE (I.E. 1-MM)

SCORER_______ RELIABILITY_______

Figure 8

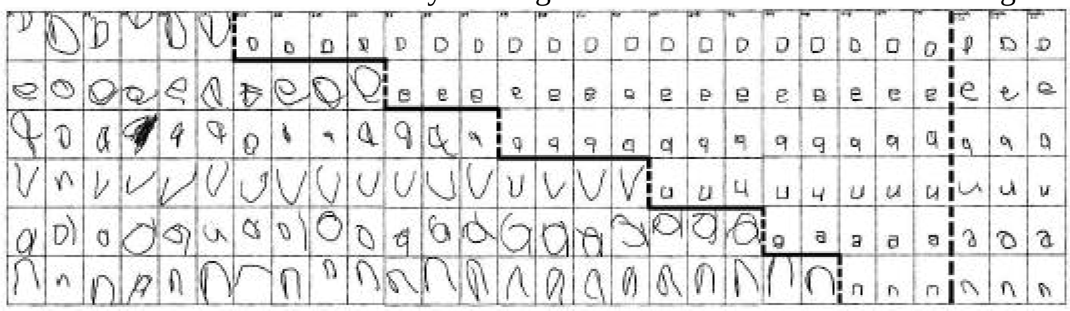
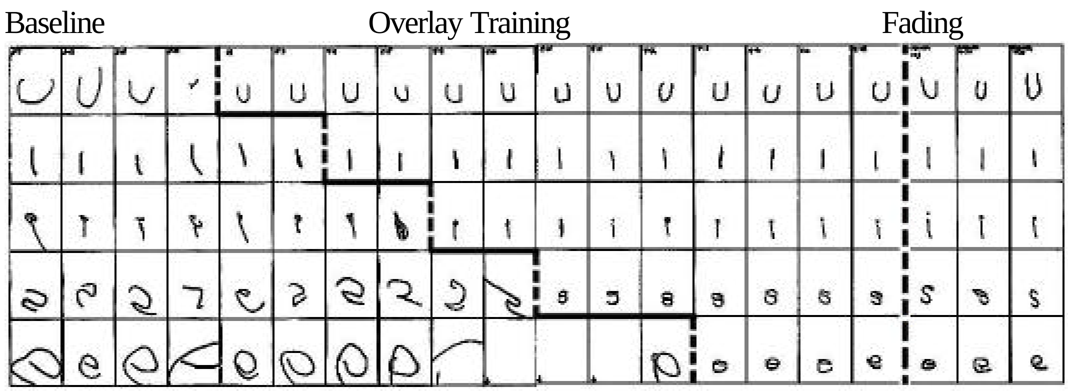


Figure 9



<u>Figure 10</u>

Figure 11

Baseline Overlay Training Fading

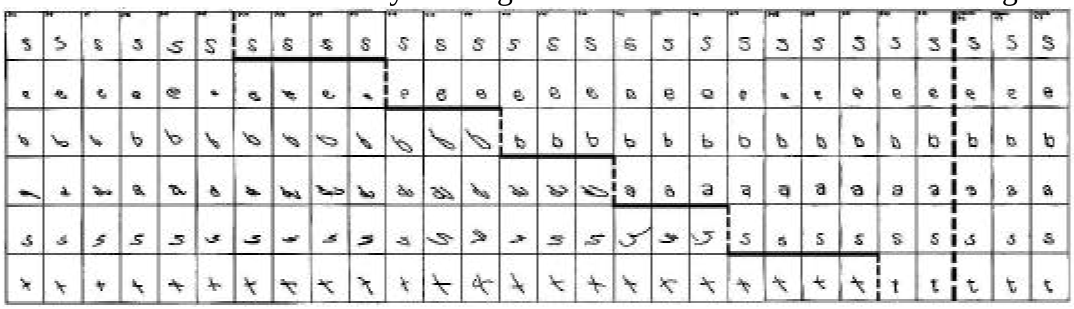


Figure 12

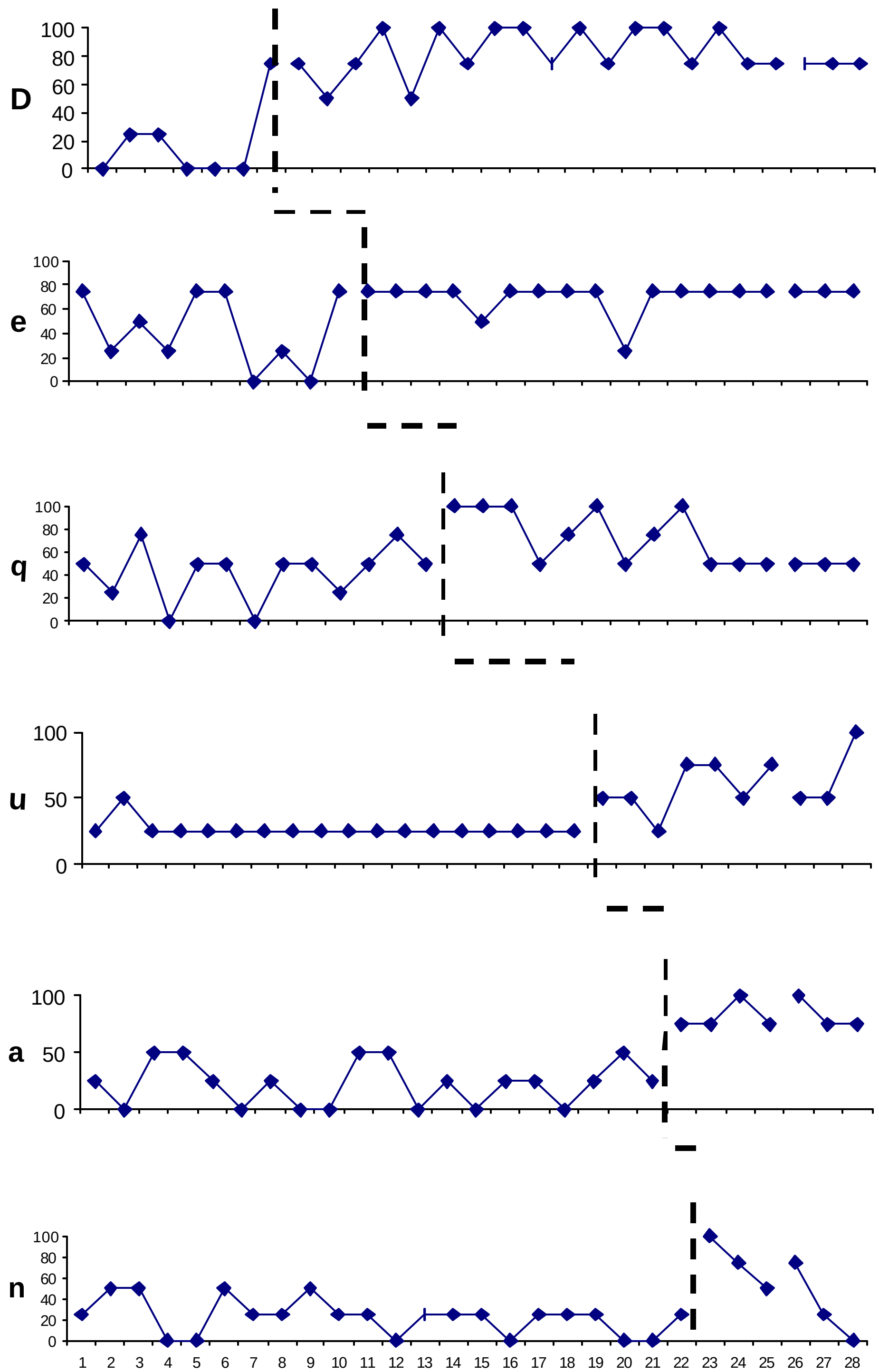

D
e
q
u
a
n
100
80
60
40
20
0
50
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

Figure 13

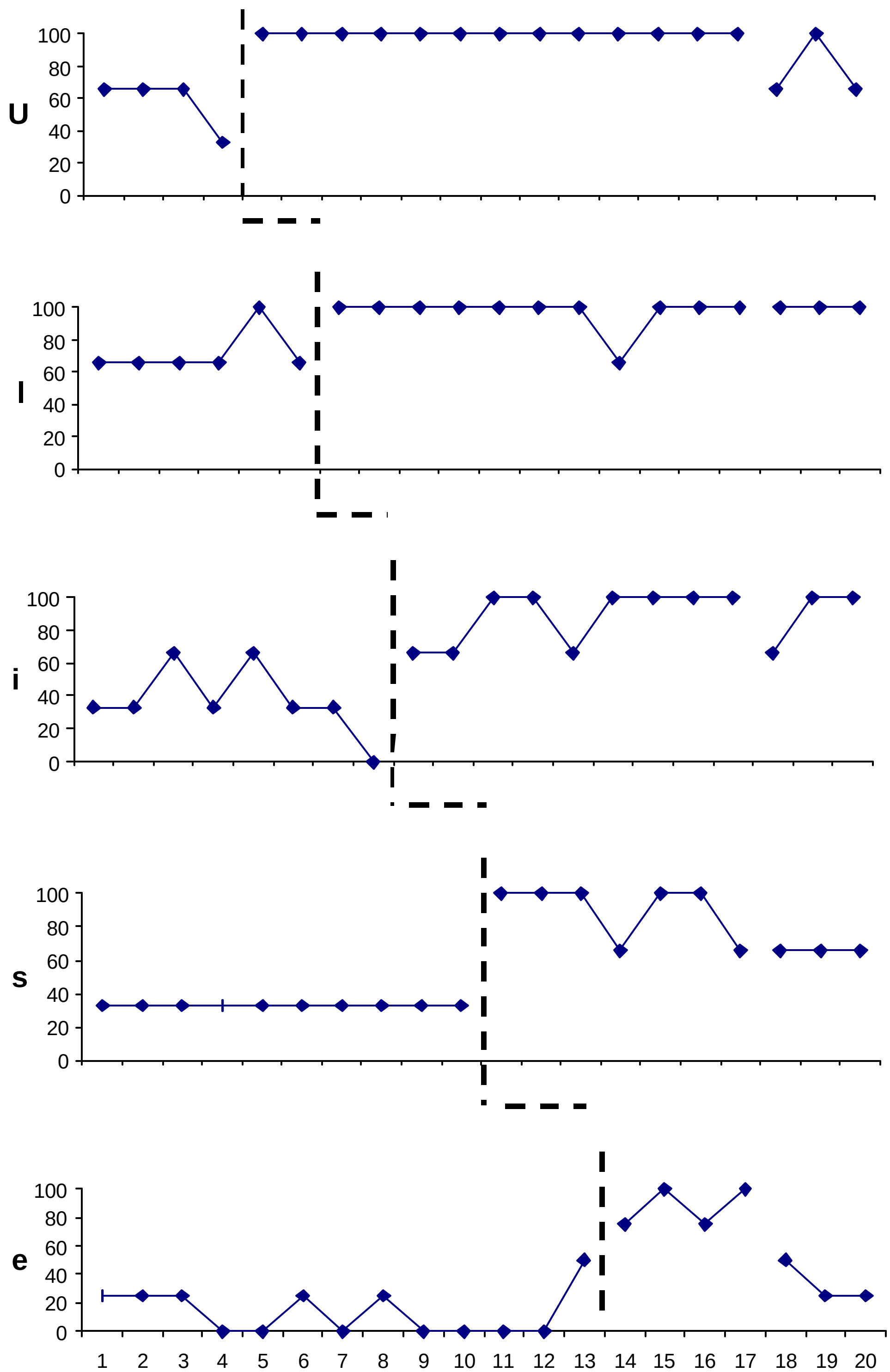
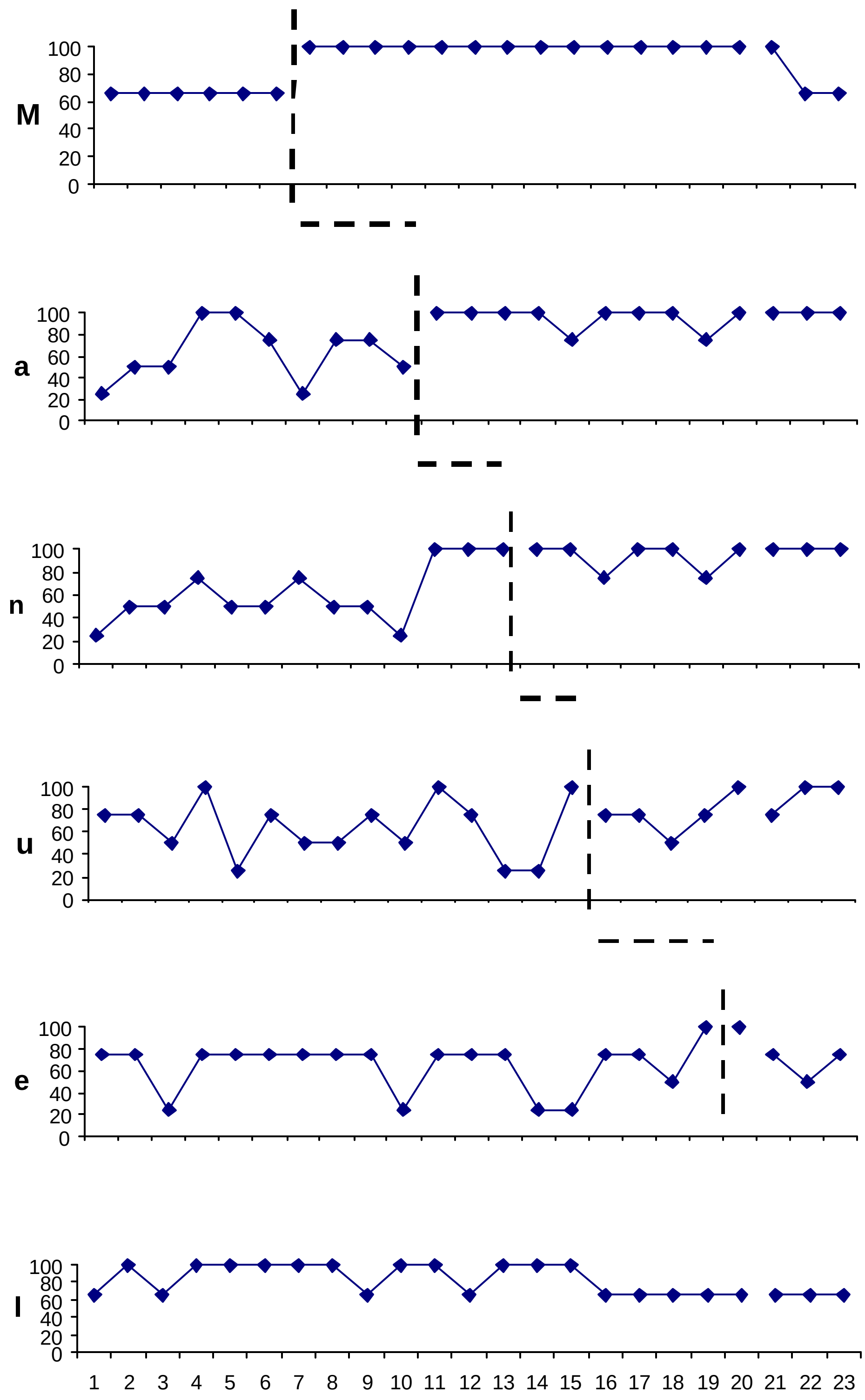
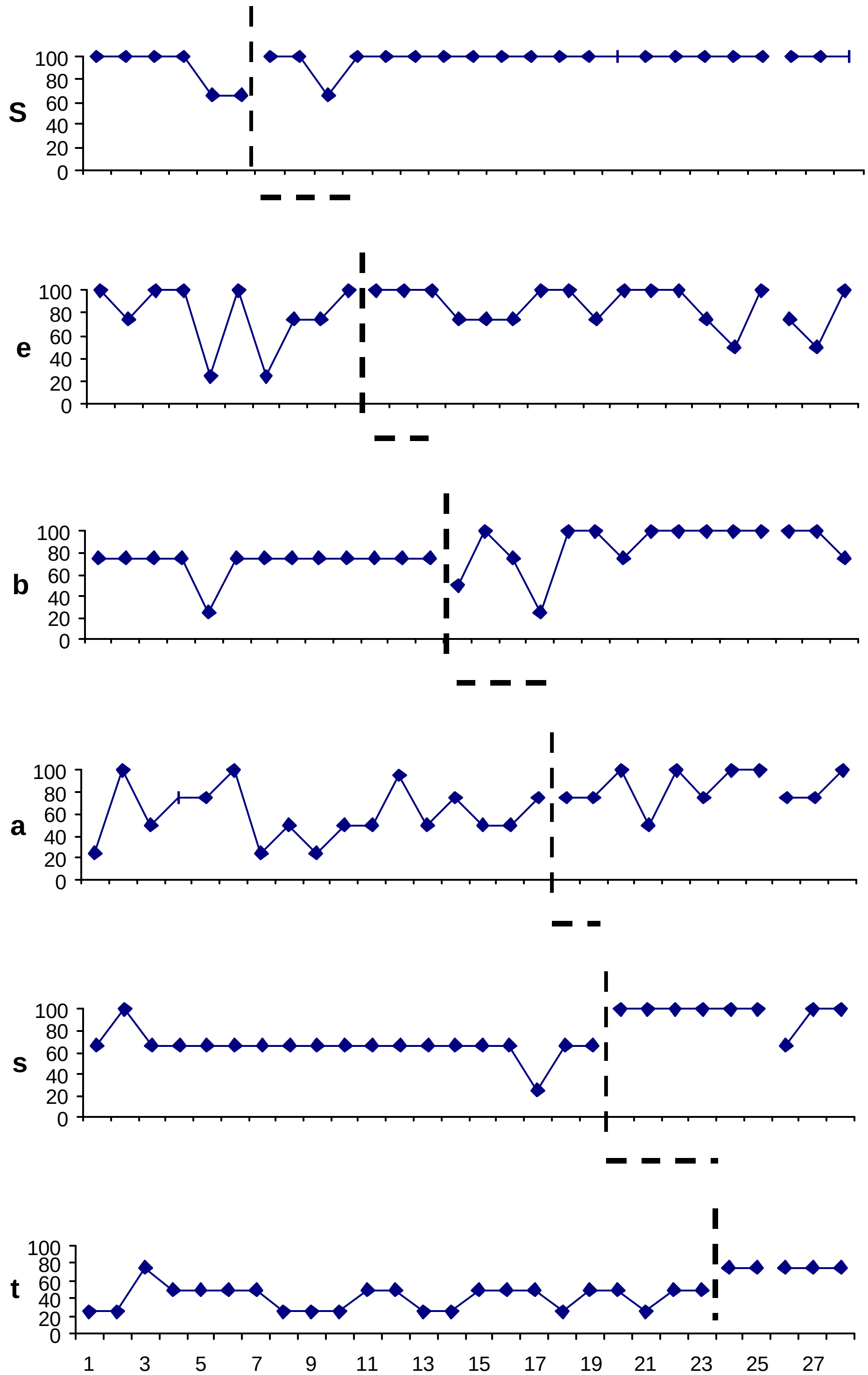


Figure 14



M
a
n
u
e
l
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Figure 15



<u>Figure 16</u>

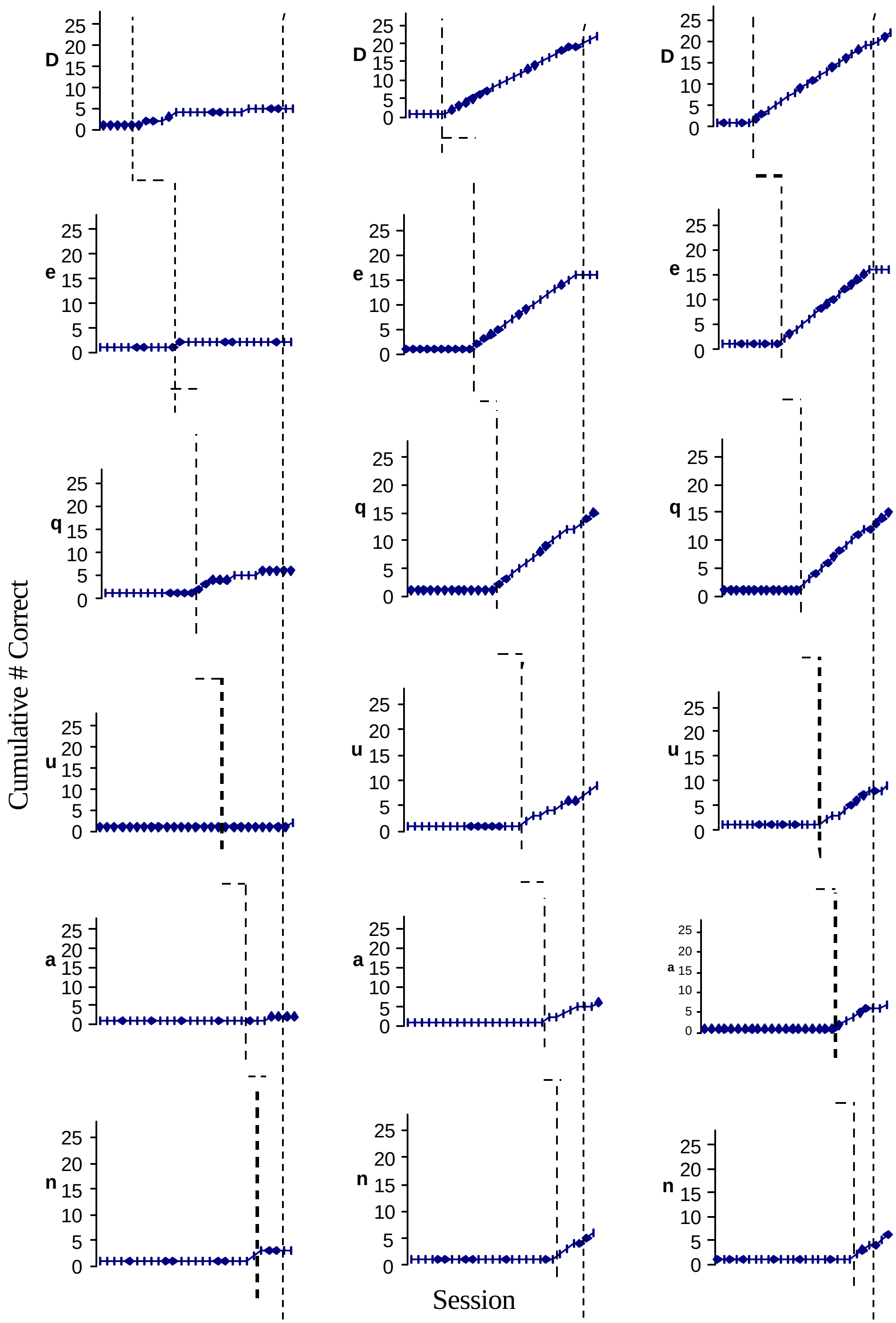


Figure 17

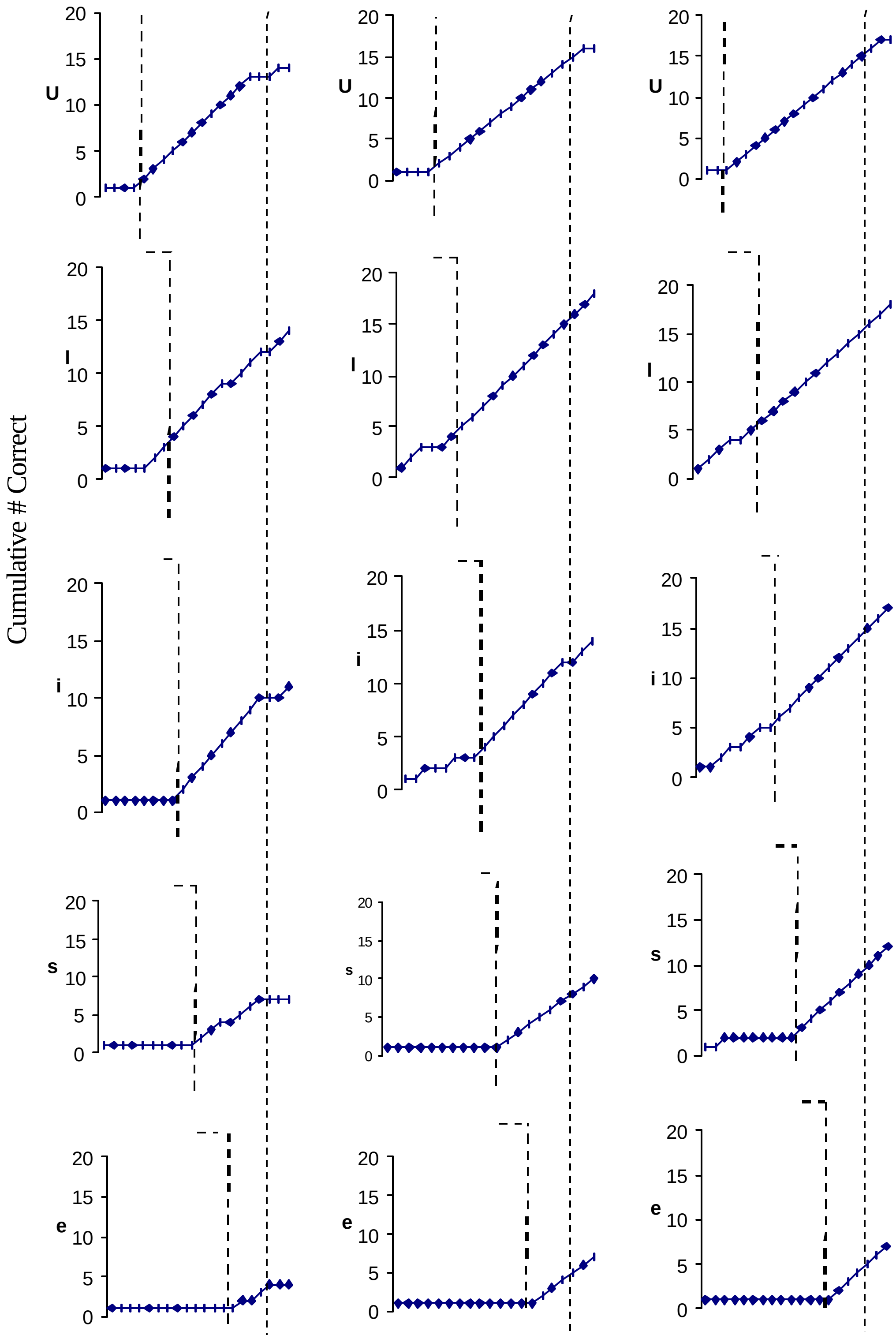


Figure 18

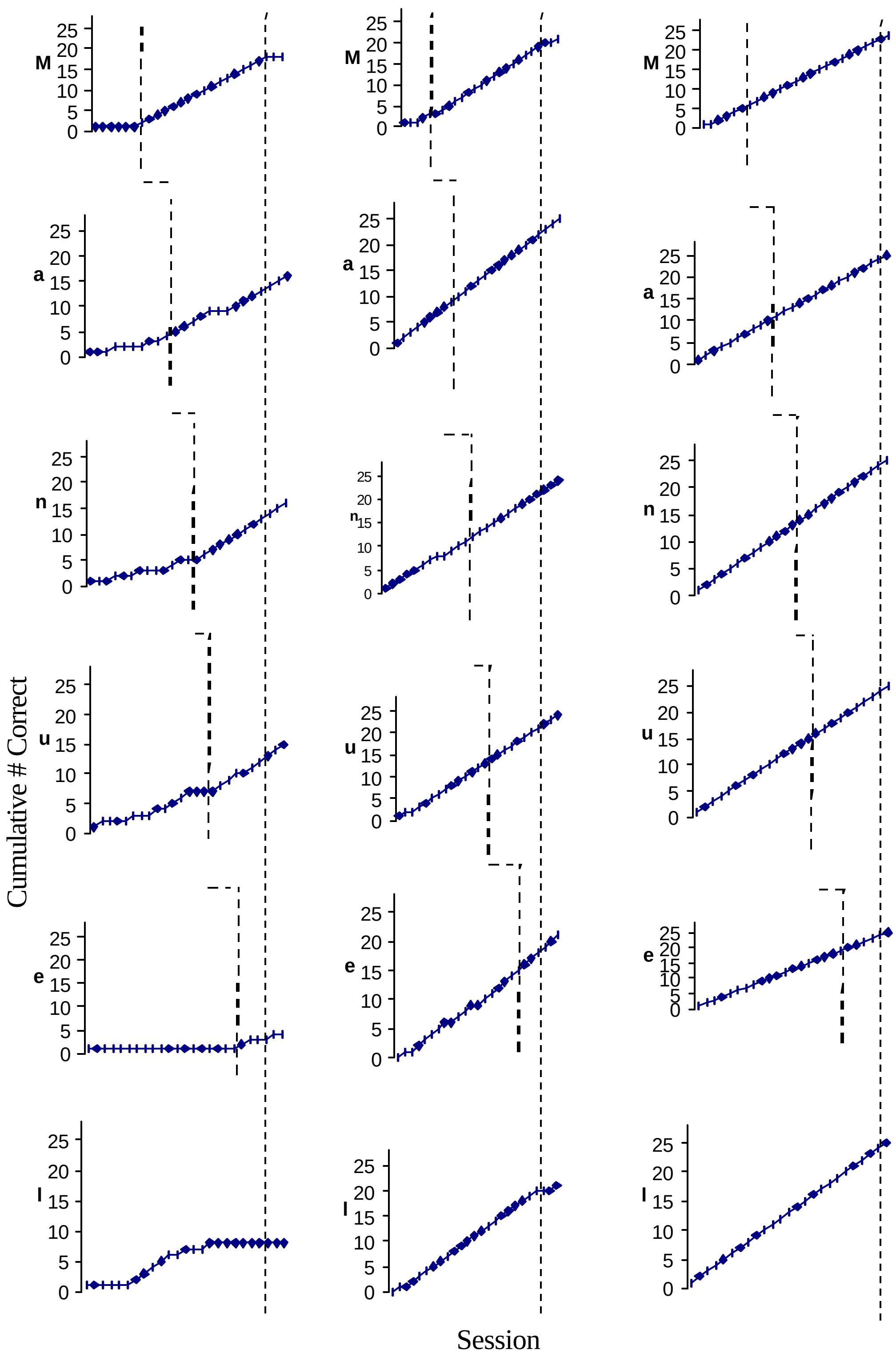


Figure 19

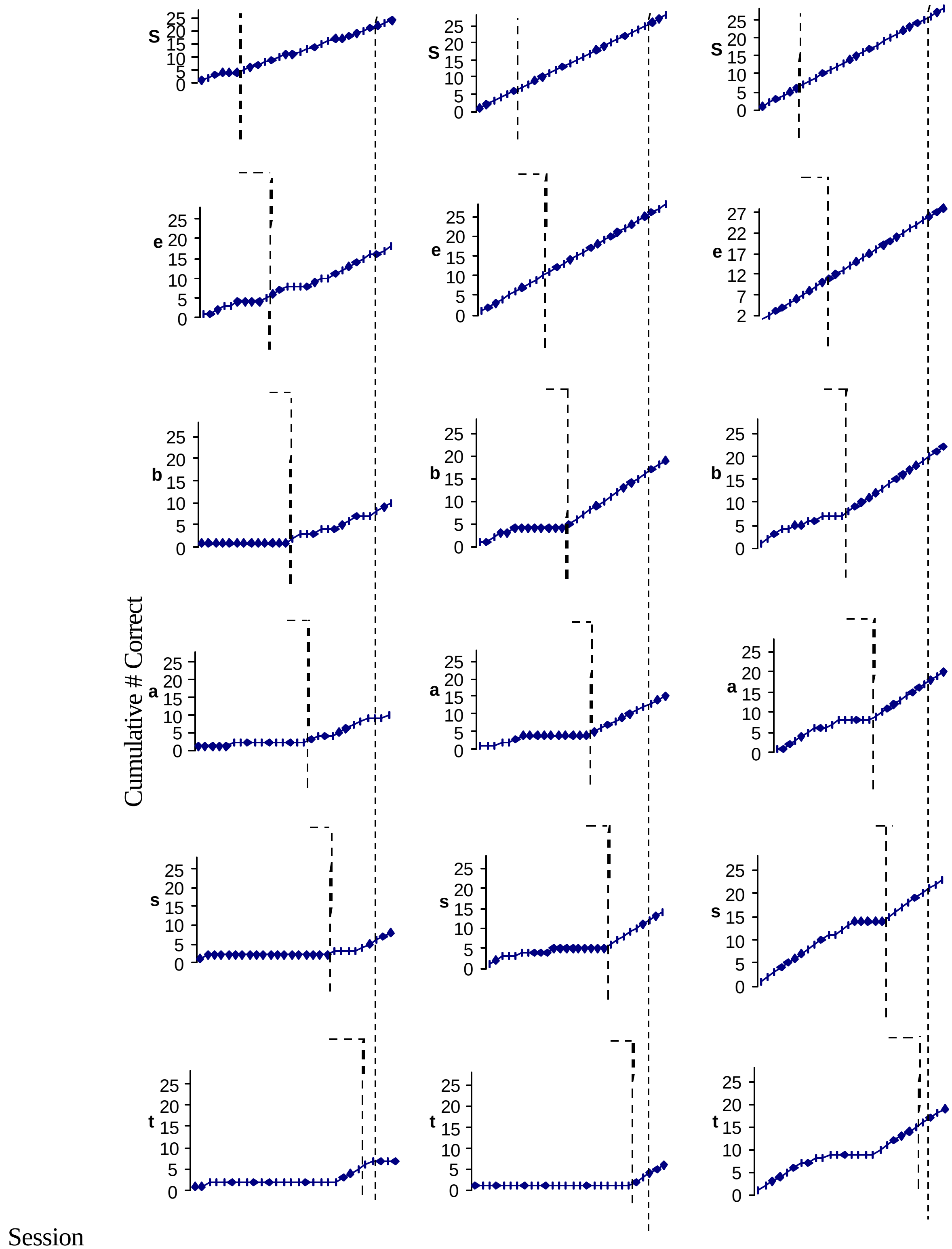
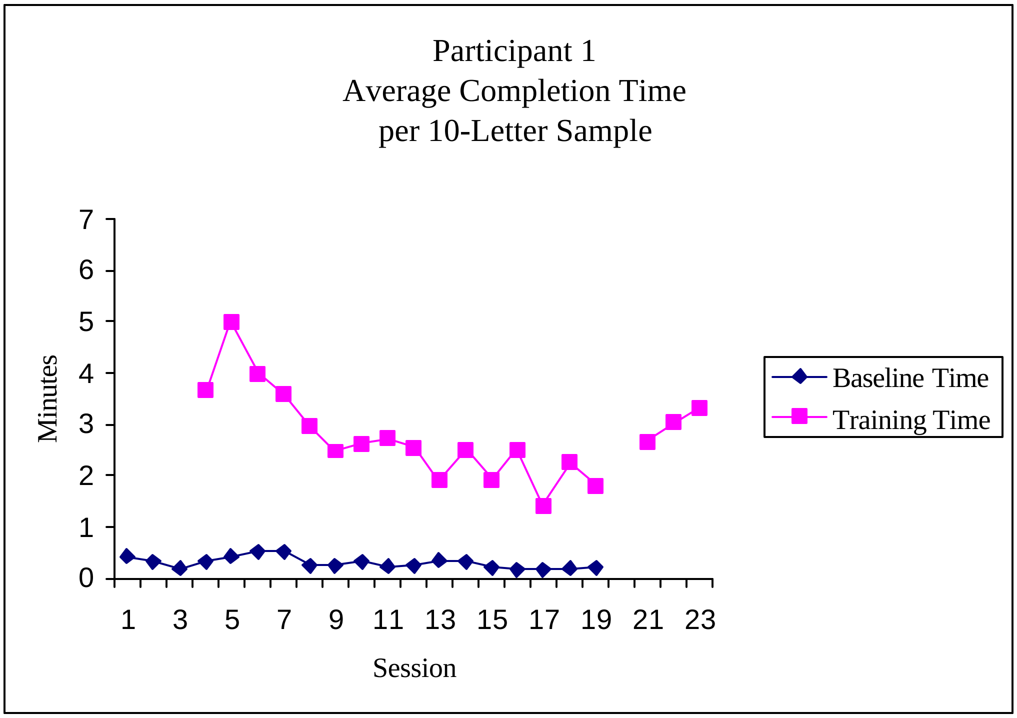
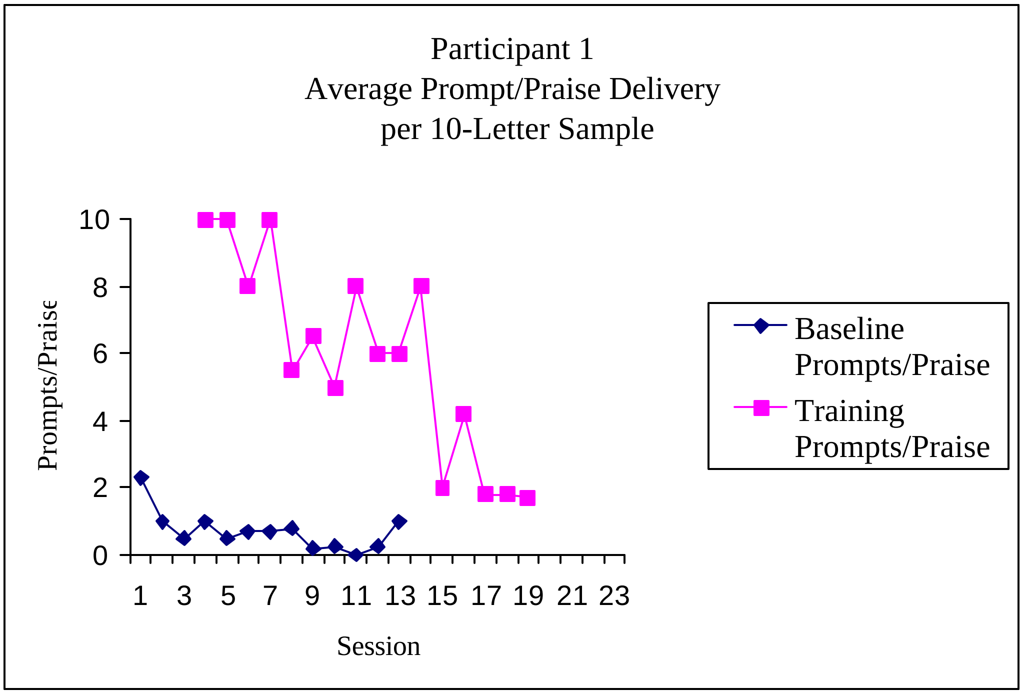


Figure 20

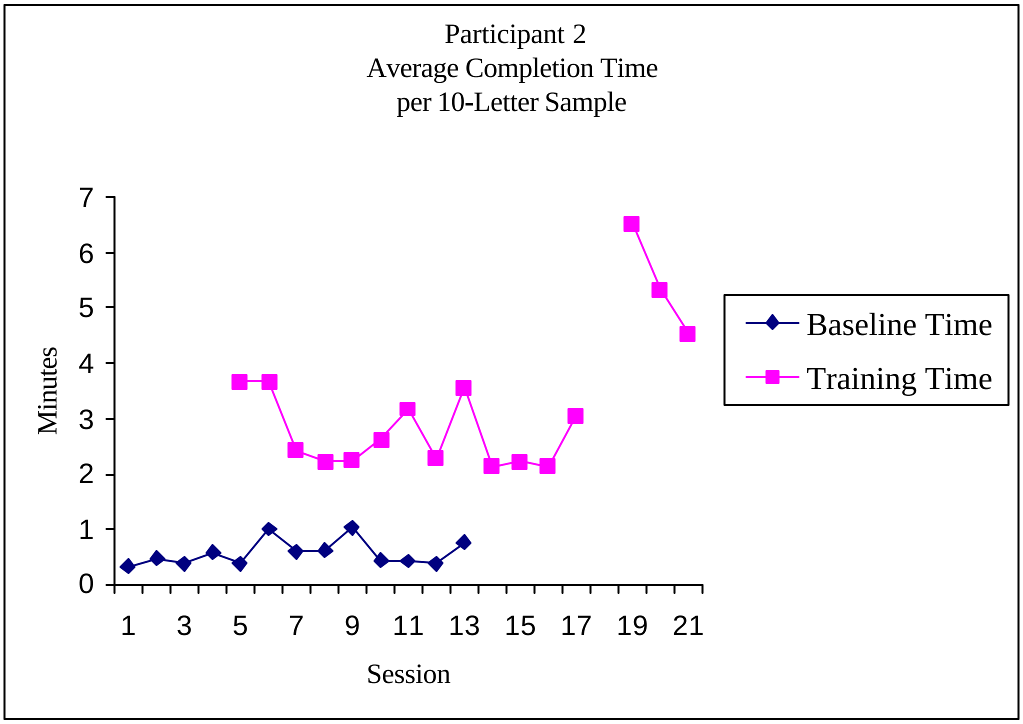


Participant 1
Average Completion Time
per 10-Letter Sample
Minutes
7
6
5
4
3
2
1
0
1 3 5 7 9 11 13 15 17 19 21 23
Session
Baseline Time
Training Time

Participant 1
Average Prompt/Praise Delivery
per 10-Letter Sample
Prompts/Praise
10
8
6
4
2
0
1 3 5 7 9 11 13 15 17 19 21 23
Session
Baseline Prompts/Praise
Training Prompts/Praise

Figure 21


Participant 2
Average Completion Time
per 10-Letter Sample
Minutes
0
1
2
3
4
5
6
7
1
3
5
7
9
11
13
15
17
19
21
Session
Baseline Time
Training Time

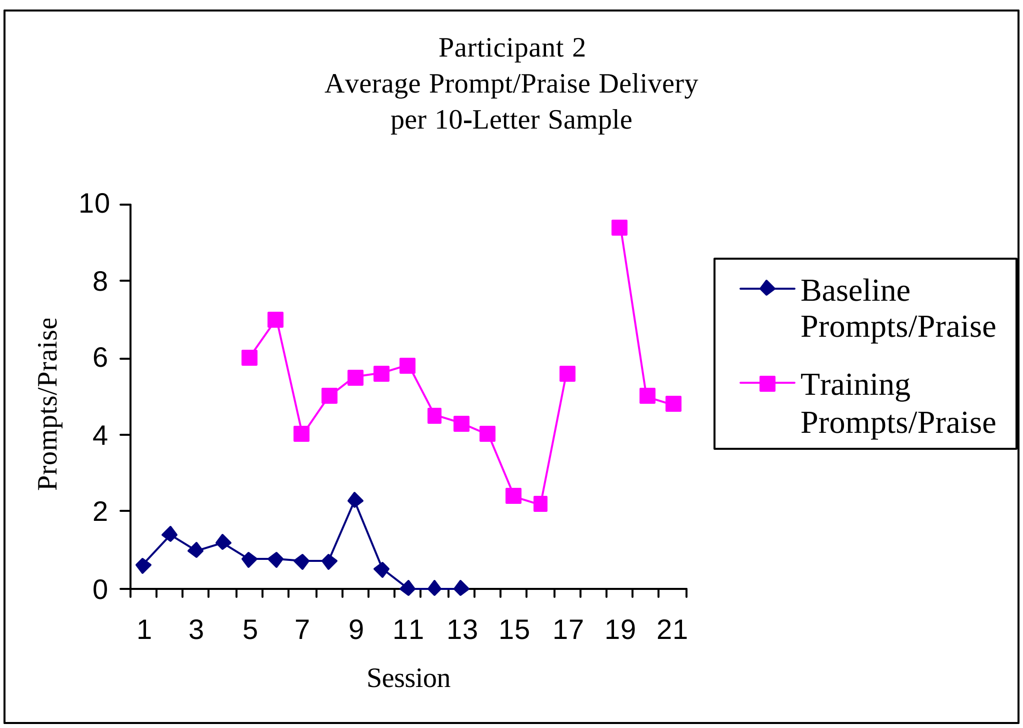
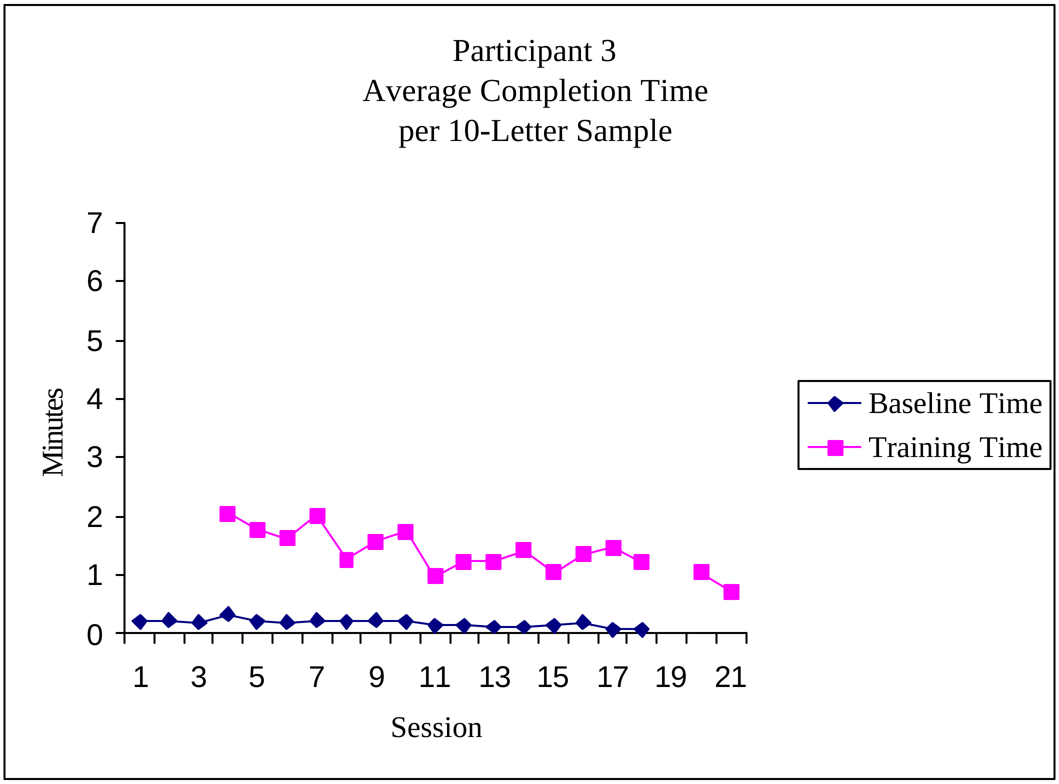

Participant 2
Average Prompt/Praise Delivery
per 10-Letter Sample
Prompts/Praise
0
2
4
6
8
10
1
3
5
7
9
11
13
15
17
19
21
Session
Baseline Prompts/Praise
Training Prompts/Praise

Figure 22


Participant 3
Average Completion Time
per 10-Letter Sample
Minutes
7
6
5
4
3
2
1
0
1 3 5 7 9 11 13 15 17 19 21
Session
Baseline Time
Training Time

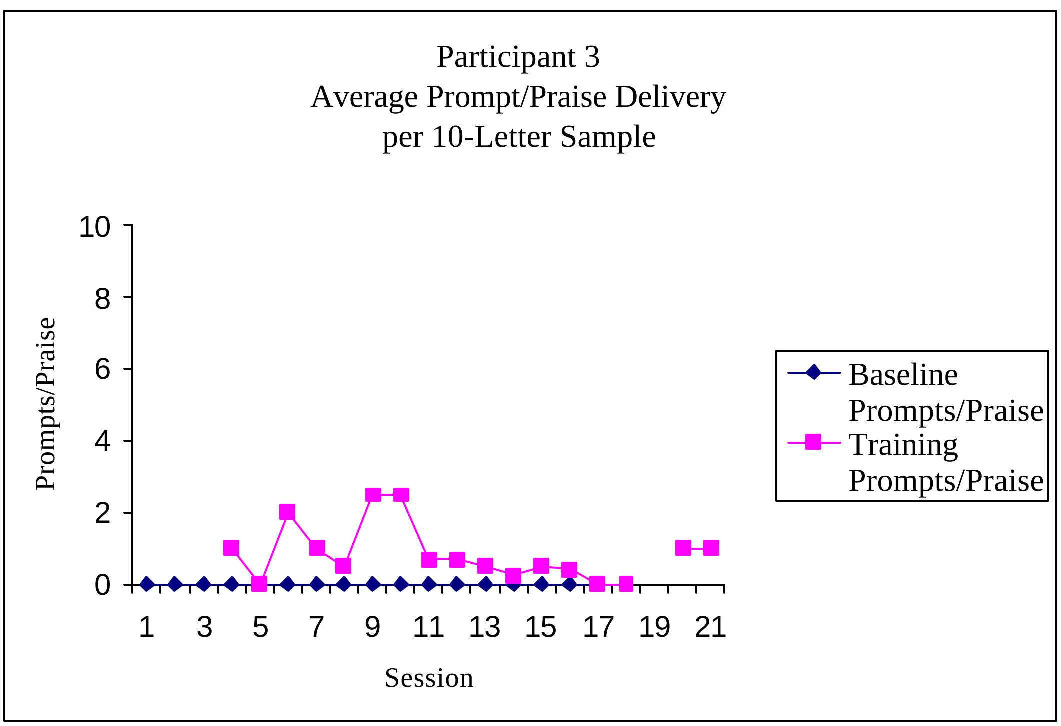
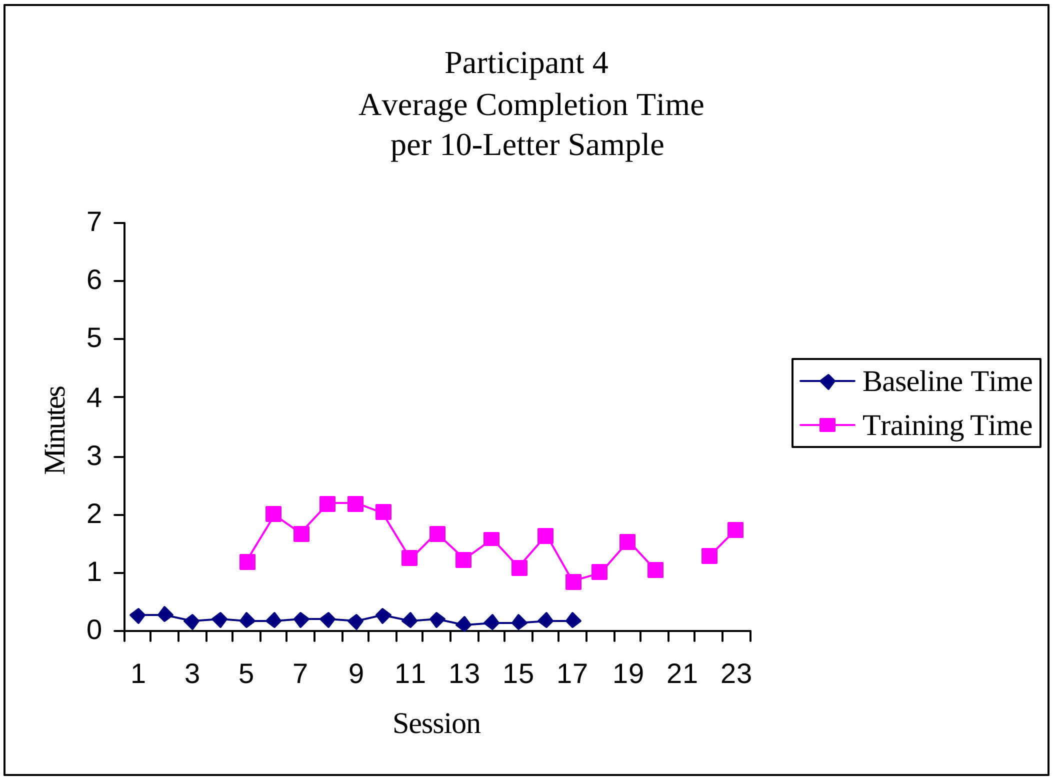
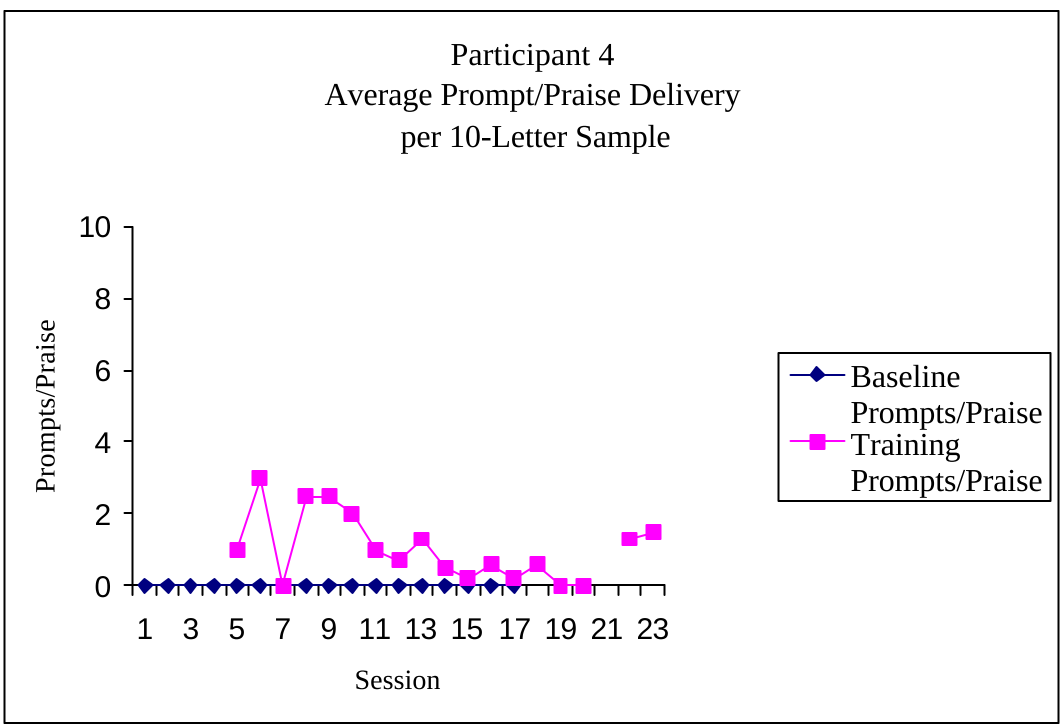

Participant 3
Average Prompt/Praise Delivery
per 10-Letter Sample
Prompts/Praise
10
8
6
4
2
0
1 3 5 7 9 11 13 15 17 19 21
Session
Baseline Prompts/Praise
Training Prompts/Praise

Figure 23



Participant 4
Average Completion Time
per 10-Letter Sample
Minutes
0
1
2
3
4
5
6
7
1
3
5
7
9
11
13
15
17
19
21
23
Session
Baseline Time
Training Time

Participant 4
Average Prompt/Praise Delivery
per 10-Letter Sample
Prompts/Praise
0
2
4
6
8
10
1
3
5
7
9
11
13
15
17
19
21
23
Session
Baseline Prompts/Praise
Training Prompts/Praise

## REFERENCES


Arena, J. (1968). Handwriting: Must it be a learning and instructional enigma? Academic Therapy, 4, 5-6.

Askov, E., & Greff, K. N. (1975). Handwriting copying versus tracing as the most effective type of practice. Journal of Educational Research, 69, 96-98.

Baer, D. M. (1977). Reviewer’s comment: Just because it’s reliable doesn’t mean that you can use it. Journal of Applied Behavior Analysis, 10, 117-119.

Barbe, W. B., Lucas, V. H., & Wasylyk, T. M. (Eds.). (1984). Handwriting. Columbus: Zaner-Bloser, Inc.

Berninger, V. (1999). Coordinating transcription and text generation in working memory during composing: "Automatic and constructive processes.". Learning Disability Quarterly, 22, 99-112.

Bezzi, R. (1962). A standardized manuscript scale for grades 1, 2, and 3. Educational Research, 55, 339-340.

Fauke, J., Burnett, J., Powers, M., & Sulzer-Azaroff, B. (1973). Improvement of handwriting and letter recognition skills: A behavior modification procedure. Journal of Learning Disabilities, 6, 25-29.

Feldt, L. S. (1962). The reliability of measures of handwriting quality. Journal of Educational Psychology, 53, 288-292.

Furner, B. A. (1969). The perceptual-motor nature of learning in handwriting. Elementary English, 46.

Graham, S. (1986). The reliability, validity, and utility of three handwriting measurement procedures. Journal of Educational Research, 79, 373-380.

Graham, S. (1986). A review of handwriting scales and factors that contribute to variability in handwriting scores. Journal of School Psychology, 24, 63-72.

Graham, S., & Miller, L. (1980). Handwriting research and practice: A unified approach. Focus on Exceptional Children, 13(2), 1-16.

Graham, S., Boyer-Shick, K., & Tippets, E. (1989). The validity of the handwriting scale from the test of written language. Journal of Educational Research, 82(3), 166-171.

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of Handwriting Problems in Beginning Writers. Journal of Educational Psychology, 92(4), 620-633.

Halpin, G., & Halpin, G. (1976). Special paper for beginning handwriting: An unjustified practice? Journal of Educational Research, 69, 267-269.

Helwig, J. J., Johns, J. C., Norman, J. E., & Cooper, J. O. (1976). Brief Technical Report The measurement of manuscript letter strokes. Journal of Applied Behavior Analysis, 9, 231-236.

Herrick, V., & Okada, N. (1963). The present scene: Practices in the teaching of handwriting in the U.S.-1960. In V. Herrick (Ed.), New horizons for research in handwriting.

Hildreth, G. (1963). Early writing as an aid to reading. Elementary English, 40, 236-241.

Hirsh, E., & Niedermeyer, F. C. (1973). The effects of tracing prompts and discrimination training on kindergarten handwriting performance. Journal of Educational Research, 67, 81-83.

Hopkins, B. L., Schutte, R. C., & Garton, K. L. (1971). The effects of access to a playroom on the rate and quality of printing and writing of first and second-grade students. Journal of Applied Behavior Analysis, 4, 77-87.

Jones, J. C., Trap, J., & Cooper, J. O. (1977). Technical Report: Students' self-recording of manuscript letter strokes. Journal of Applied Behavior Analysis, 10, 509-514.

Kau-To Leung, E., Treblas, P. V., Hill, D. S., & Cooper, J. O. (1979). Space size and accuracy of first grade students' manuscript writing. Journal of Educational Research, 73, 79-81.

Krzesni, J. S. (48). Effect of different writing tools and paper on performance of the third grader. Elementary English, 48, 821-824.

Lamme, L. L. (1979). Handwriting in an early childhood curriculum. Young Children, 35(1), 20-27.

Lamme, L. L. (1984). Teaching handwriting in an integrated language arts program. In Handwriting (pp. 206-216). Columbus: Zaner-Bloser.

McMenamin, B., & Martin, M. (1980). Right Writing. Spring Valley, CA: Cursive Writing Associates.

Milone, M. N., Wilhide, J. A., & Wasylyk, T. M. (1973). Spelling and handwriting: Is there a relationship? Spelling and Progress Bulletin, 13(2), 246-250.

Pettersson, R. (1989). Visuals for information research and practice. Englewood Cliffs, New Jersey: Educational Technologies Publications.

Ramming, J. (1968). Using the chalkboard to overcome handwriting difficulties. Academic Therapy, 4, 49-51.

Richardson, S. (1995). Specific developmental dyslexia: Retrospective and prospective views. In Clinical studies of multisensory structured language education. Salem, OR: IMSLEC.

Robin, A. L., Armel, S., & O'Leary, K. D. (1975). The effects of self-instruction on writing deficiencies. Behavior Therapy, 6, 178-187.

Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: A developmental and pragmatic concept for behavior analysis. Journal of Applied Behavior Analysis, 30, 533-544.

Salzburg, B. H., Wheeler, A. J., Devar, L., & Hopkins, B. L. (1971). The effect of intermittent feedback and intermittent contingent access to play on printing of kindergarten children. Journal of Applied Behavior Analysis, 4, 163-171.

Skinner, B. F. (1968). R. M. Elliott, G. Lindzey, & K. MacCorquodale (Eds.), The Technology of Teaching. New York: Meredith Corporation.

Skinner, B. F., & Krakower, S. (1968). Handwriting with Write and See. Chicago: Lyons & Carnahan.

Stempel-Mathey, L., & Wolf, B. J. (1999). Teaching Handwriting. In Multi-sensory Teaching of Basic Language Skills (pp. 257-280). Columbia University New York: Teacher's College.

Strickling, C. (1973). The effects of handwriting and related skills upon the spelling score of above average and below average readers in the fifth grade. In Handwriting. Columbus: Zaner-Bloser, Inc.

Trap, J. J., Milner-Davis, P., Joseph, S., & Cooper, J. O. (1978). The effects of feedback and consequences on transitional cursive letter formation. Journal of Applied Behavior Analysis, 11, 382-393.