

January 01, 2010

## Effects of mastery criteria on maintenance skills in students with developmental disabilities

Shawn F. Vieira  
*Northeastern University*

---

### Recommended Citation

Vieira, Shawn F., "Effects of mastery criteria on maintenance skills in students with developmental disabilities" (2010). *Counseling Psychology Master's Theses*. Paper 13. <http://hdl.handle.net/2047/d20000791>

This work is available open access, hosted by Northeastern University.

**Effects of Differing Mastery Criteria on Maintenance Skills in Students  
with Developmental Disabilities**

**A Thesis Presented**

**by**

**Shawn F. Vieira**

The Department of Counseling and Applied Educational Psychology

In partial fulfillment of the requirements

for the degree of

**Master of Science**

**in the field of**

**Applied Behavior Analysis**

Northeastern University

Boston, MA

April 2010

NORTHEASTERN UNIVERSITY

Bouvé College of Health Sciences Graduate School

Thesis Title: Effects of Differing Mastery Criteria on Maintenance Skills in Students  
with Developmental Disabilities

Author: Shawn F. Vieira

Department: Counseling and Applied Educational Psychology

Approved for Thesis Requirements of Master of Science Degree

---

Hanna C. Rue, Ph.D., BCBA-D

---

Date

---

Katherine Gilligan, M.S., BCBA

---

Date

---

Karen E. Gould, Ph.D., BCBA

---

Date

Running head: EFFECTS OF DIFFERING MASTERY CRITERIA

**Effects of Differing Mastery Criteria on Maintenance Skills in Students  
with Developmental Disabilities**

by

Shawn F. Vieira

B.A., University of Massachusetts, Dartmouth

Submitted in partial fulfillment of the requirements for the degree of  
Master of Science in Applied Behavior Analysis  
in the Bouvé College of Health Sciences Graduate School  
of Northeastern University, April 2010

### Acknowledgements

The author would like to thank his thesis committee, Dr. Hanna Rue, Ms. Katherine Gilligan, and Dr. Karen Gould for their assistance during all steps of the manuscript writing process. Additional thanks goes to Dr. Hanna Rue and Ms. Katherine Gilligan for their guidance and support in all aspects of this research.

A special thank you to Ms. Stephanie Barney, Ms. Lara Bollinger, Mr. Noah Borgondy, Mr. Scott DaCosta, and Ms. Jennifer Merry for their assistance during the data collection process.

Also, thank you to his family for their continued encouragement and support.

Effects of Differing Mastery Criteria on Maintenance Skills in Students  
with Developmental Disabilities

Table of Contents

I.	Abstract .....	6
II.	Introduction .....	7
	a. Teacher-defined criteria on course grades and student performance.....	7
	b. Effects of mastery criteria and assignment length .....	9
	c. Effects of mastery criteria on student performance .....	9
	d. Purpose .....	11
III.	Method .....	12
	a. Participants.....	12
	b. Setting and Materials.....	13
	c. IOA and Procedural Integrity.....	14
	d. Procedure .....	14
IV.	Results .....	17
V.	Discussion .....	18
VI.	References .....	24
VII.	Appendixes .....	25
VIII.	Figure Caption .....	29
IX.	Figure .....	30

### Abstract

Skills of students with developmental disabilities are often considered learned once they have achieved a mastery criterion. Teachers often assign an arbitrary criterion to determine whether a student has demonstrated sufficient learning of a skill, however, little research has been done on the maintenance of learned skills of students' with developmental disabilities. The purpose of this study was to compare the effects of a high and low mastery criteria on the maintenance of a receptive identification task. Three adolescents diagnosed with developmental disabilities participated in the study. Participant 1 achieved both the low and high criteria as designed. Participants 2 and 3 demonstrated the skill at the high criteria only. Results indicated that other variables independent of mastery criteria might control how well a student maintains a skill.

## Effects of Differing Mastery Criteria on Maintenance Skills in Students with Developmental Disabilities

The skill acquisition of students with developmental disabilities has historically focused on achieving a set mastery criterion (Luiselli, Russo, Christian, & Wilczynski, 2008). For these students, the annual goal as stated in their Individualized Education Program (IEP) dictates how and if a skill has been mastered. As defined in the Merriam-Webster Online Dictionary (Mastery, 2009), mastery is “the possession or display of great skill or technique.” Skills of students with developmental disabilities are often measured in many ways including percent independence, percent accuracy, by number of prompts to complete a task, duration, etc. When writing goals, educators often assign an arbitrary criterion to determine whether a student has demonstrated sufficient learning to be able to move onto a more complex skill. Little research has been done on the maintenance of learned skills of students’ with developmental disabilities.

Most research looking at the effects of differing mastery criteria of skills has been conducted with typical students enrolled in undergraduate studies. Johnston & O’Neil (1973) investigated the influence of teacher-defined criteria on course grades and student performance. Sixty-five undergraduate students enrolled in an abnormal psychology course participated. All students participated in performance sessions where they were given a pool of item cards that were to be answered orally according to a text or lecture unit as specified by the undergraduate student. Student teachers recorded the duration of each session and the total number of items that were

answered correctly or incorrectly. The student teachers reviewed those items that were answered incorrectly during the session with the undergraduate students.

The students were randomly assigned to one of five conditions. Experiment I contained no teacher-defined criteria. Students were only told that they would be graded on a curve. The purpose of this condition was to compare the effects of having no teacher-defined criteria to having defined criteria as used in the other experiments.

The criteria in Experiments II, III, and IV were sequential variations of three levels of correct responses, i.e. high (greatest number of correct responses), medium, and low (least number of correct responses). Changes in criteria in Experiments II to IV were made individually and at different points for each student. In all three experiments, students received a course grade of "A" as defined by the high, medium, or low criteria. No criteria were announced for other course grades.

Experiment V also contained a high, medium, and low criteria. However, at the start of the experiment, students were told that only those students who met the high criteria would receive an "A." A third of the way into the quarter, the medium criteria was introduced, and the students were informed that meeting the medium criteria would produce a "B" grade. When the low criteria were introduced, students were made aware that these criteria would result in a "C" grade.

Results from Experiment I demonstrated that without a set criteria in place, students performed poorly, and there was more variability among their responses. Results for Experiments II, III, and IV showed that there was a functional relation between the criteria and students performance. The students dropped their performance to the lowest criteria needed to receive a grade of an "A." The results of

Experiment V show that the students usually remained above the “A” level criteria throughout all the phase changes. This study demonstrated that teacher-defined criteria may indeed influence student performance.

Semb (1974) studied the effects of mastery criteria and assignment length on personalized instruction. Eighty-nine undergraduate students were randomly assigned to one of two groups. Using a within-group reversal design, three different manipulations of criteria and assignment length were used. During the high criterion-short assignment manipulation, 100% correct responses based on content and review quizzes were needed to demonstrate mastery. The low criterion-short assignment manipulation required 60% correct or better on content and review quizzes to attain mastery. To achieve mastery in the high criterion-long assignment, 100% correct responses were needed on review quizzes only. Groups were exposed to the experimental conditions in different sequences. The effects of each manipulation were also assessed using a multiple baseline achievement test. The achievement tests were taken on the first day of class and at the completion of each of the four course parts. Results show that a high criterion produced higher percentages of correct answers when compared to the lower criterion. They did not, however, assess the maintenance effects of the different study manipulations.

Reiser, Driscoll, Farland, Vergara, and Tessmer (1986) conducted a study to determine the effects of mastery criteria on student performance. 121 undergraduate educational psychology students were randomly assigned to one of three conditions: The ascending criteria group were informed that the criterion for mastering a unit quiz was 70% correct on the first six unit quizzes, 80% correct on the next three

quizzes, and 90% correct on the last six. The descending criteria group were told that they had to achieve a score of 90% correct on the first six unit quizzes, 80% correct on the following three, and 70% on the last six. Those in the third and final group, the fixed criteria group, were told that they needed to score 80% on each of the fifteen quizzes.

The course consisted of 15 self-instructional units. Participants met with either the instructor or one of two graduate students on an individual basis rather than attend lectures. Multiple-choice and short-answer quizzes were used to determine if the participants met the prescribed criteria. Participants had to pass one unit before moving onto the next one. Mastering a unit required that the participant score at or above the mastery level that their group was assigned. Participants could re-take alternate versions of the quizzes until they met mastery criteria. At the end of the 15-week semester, all participants took a comprehensive final examination, regardless of the number of unit-quizzes they had mastered.

Results from the unit quizzes for the ascending group show that as the criteria increased, so did participants performance. During the 70% correct criteria, participants averaged 90.4% correct. During the 80% correct criteria, the mean quiz average was 91.5%, and during the 90% correct criteria participants averaged 94.2% correct. The constant criteria group's responding remained stable at 90.8%, 90.7%, and 90.6% correct. The results of the quiz scores from the descending group showed that as the criteria decreased so did participants' responding. During the 90%, 80%, and 70% criteria, students responded with 93.9%, 92.7%, and 89.2% correct, respectively.

Results from the comprehensive final examination demonstrate that the mastery criteria had little effect on the final examination performance. Participants in the ascending criteria attained an overall performance of 78.2% correct. Participants in the constant criteria group attained 76.5% correct, while the participants in the descending criteria had an overall performance score of 78.7% correct. Reiser et al. (1986) demonstrated that setting a mastery criterion for unit quizzes had little effect on how the participants would perform on a comprehensive final examination, however, a higher mastery criteria produced a higher quiz performance.

A limitation of this study is that although participants were expected to score at the mastery level of their group, some scored above their group's mastery level (e.g., score of 80% correct during the 70% correct mastery session). This skews the maintenance effects of any condition. Another limitation is indicated by the requirement that all participants take the cumulative final examination, even if they had not yet completed all units and passed all unit quizzes. This situation may have resulted in low scores on the final examination, which suggest that these low scores would be a function of lack of exposure to a unit rather than a function of the differing mastery criteria.

The purpose of the current study was to extend the findings of Reiser et al. (1986) to students with developmental disabilities and to compare the maintenance effects of two different mastery criteria. The maintenance effects of two different mastery criteria were evaluated using two groups of stimuli within the same response class. Each stimulus group was trained until the mastery criteria were reached (75% or 95% correct responding). Once mastery was achieved, training ceased, and the

stimuli were not presented to the student until maintenance sessions had commenced. Maintenance sessions took place approximately one-month following the mastery of criteria of a stimulus group.

## METHOD

### *Participants*

Three adolescents with a diagnosis of developmental disabilities participated in the study. The participants were 3 of an original pool of 12 participants. Twelve participants were initially included based on the assumption that some would be dropped from the study for either learning the target stimulus too quickly or too slowly. Teachers of all potential participants were interviewed as a means to determine if the students had the necessary prerequisite skills to participate. To qualify as a participant, students were required to have attending skills and be able to identify common objects. Students were excluded if they were able to identify states in the United States of America or had been exposed to the stimuli in specific academic programs.

Participant 1 was an 18-year-old male diagnosed with Pervasive Developmental Disorder-NOS, and he had some experience using a time delay prompt during discrete trial sessions. Participant 2 was a 15-year-old male diagnosed with autism. He had little experience using a time delay prompt during discrete trial training and was more accustomed to errorless learning. Participant 3 was an 18-year-old male diagnosed with Pervasive Developmental Disorder who had some experience using a time delay prompt during discrete trial sessions. All participants

had attended the private school where this research was conducted for two or more years.

### *Setting and Materials*

All sessions were conducted in an assessment room containing a table and two chairs. A trained experimenter conducted all sessions. The participants' responses were recorded using a data sheet scored by the experimenter (see Appendix A for sample data sheet). A second observer collected data on procedural integrity (see Appendixes B, C, and D).

Outlines of the individual states in the United States of America were used as stimuli. The orientation of a state was altered so that they appeared on their side or upside-down. The 50 U.S. states were sorted into categories based on their physical characteristics: States that were tall, slender, and contained little to no straight sides were put in one category. States that appeared almost circular with no straight sides were put into another. States that did not fit into one of those two categories were not used. Five states from each of the qualifying categories were used for the experiment. One state from each category was then randomly chosen as the target stimulus, and the other four states served as distractors.

Each state outline was displayed in a white square measuring 2.25 inches, and three states were arranged on a black background measuring 8 inches by 11 inches. Each array of three stimuli included the target stimulus and two assigned distractors. They were randomly positioned, centered horizontally, and separated by 0.75 inches. During all sessions, a presentation binder containing 10 laminated stimulus pages was used.

*Inter-observer Agreement and Procedural Integrity*

A second observer independently collected data on inter-observer agreement (IOA) and procedural integrity during 34.3% of baseline sessions, 31.9% of training sessions, and 100% of maintenance probe sessions. IOA was assessed on a trial-by-trial basis by dividing the number of trials in agreement by the total number of trials and then multiplying by 100. For all participants, mean agreement was 99% during baseline sessions (range: 96.7% - 100%), 100% during training sessions, and 100% during the maintenance probe sessions.

During procedural integrity measures, an independent observer completed an integrity checklist to document whether all steps of the procedure were applied as planned, and if the target stimulus was presented in the correct position (e.g. left, right, or center). Correct positioning for all sessions were 100%. Procedural integrity scores were calculated by dividing the number of steps completed correctly by the total number of possible steps multiplied by 100. The mean integrity score across all phases was 99.8% (range: 98.2% - 100%).

*Procedure*

The experimenter conducted preference assessments with each participant in their respective classrooms. Preferred edibles based on teacher report were used. A paired stimulus preference assessment was conducted (Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I., 1992), and the two most preferred edibles were used as reinforcer in this study.

*Baseline:* During baseline sessions, participants were brought to the assessment room and asked to sit in a chair at the table. The experimenter closed the

door and was seated at the table across from the participant. The experimenter placed the presentation binder 8 to 10 inches in front of the participant. The binder was then opened to the first page and the discriminative stimulus ( $S^D$ ), "Touch \_\_\_\_\_," was delivered. If the participant attempted to respond prior to the  $S^D$  being delivered, their hands were blocked, and the experimenter then asked him to exhibit a ready response by either folding his hands on his lap or placing his hands palms down on the table. Once the ready response was exhibited, the  $S^D$  was delivered. If the student touched the correct stimulus, a (+) was recorded. If he touched an incorrect stimulus, a (-) was recorded. If he did not respond within 5 seconds of the  $S^D$ , an NR (no response) was recorded and the trial was considered incorrect. There were no programmed consequences in place during baseline sessions.

*Training:* Training sessions were conducted identically to baseline sessions with the addition of reinforcement for correct responses and conducting a correction procedure for incorrect responses. If the student responded incorrectly or did not respond within 5 seconds of the  $S^D$  being delivered, the experimenter turned away and recorded a (-) or (NR) on the data sheet. The  $S^D$  was then redelivered with an immediate prompt to the correct response, followed by reinforcement. During training sessions, correct responses were reinforced on a fixed ratio-1 (FR1) schedule of reinforcement. Reinforcement consisted of verbal praise (i.e., "Good job") paired with a preferred edible. Incorrect responses were followed by a correction procedure where the experimenter prompted the participant to the correct response. Verbal praise paired with an edible reinforcer was then delivered following the correction procedure. This intervention was administered to Participants 1 and 3 throughout the

study and to Participant 2 until session 19. During session 19, the experimenter noted that Participant 2 would point to a stimulus without attending to the presented array, then put out his hand to receive reinforcement. It was hypothesized that Participant 2 learned the contingency that a response, whether independently correct or prompted, would result in reinforcement. As a result, a second intervention (Intervention II) was put into place for Participant 2. Only correct and independent responses resulted in an edible reinforcer paired with verbal praise. An incorrect response was followed by a correction procedure and verbal praise only. Teachers of Participant 2 stated that verbal praise was reinforcing to the student, however, edibles were a more potent reinforcer. The purpose of this change was to ensure that Participant 2 was being reinforced with an edible for attending to the presented array of stimuli and selecting the correct stimulus.

*Maintenance Probe:* Each participant's maintenance probe was conducted approximately 1-month after the criterion was met. Maintenance probes were identical to baseline sessions such that correct responses received no reinforcement, and no correction procedure was administered for incorrect procedures.

#### *Mastery Criteria*

Two arbitrarily chosen mastery criteria were used based on typical IEP goals of other students attending the school. The low criterion was trained until an average of 75% correct and independent responding ( $\pm 5\%$ ) were achieved based on three consecutive sessions. The high criterion was trained until an average of 95% correct and independent responding ( $\pm 5\%$ ) were achieved based on three consecutive sessions.

## Results

Figure 1 summarizes the results for the three participants. For each participant, baseline levels were at low or chance levels of correct responding. During the training phase, Participant 1 was the only participant to achieve both the high and low mastery criteria demonstrating rapid rates of acquisition across both criteria. Participant 1 achieved mastery criteria within five sessions during the low criterion, and he achieved mastery criteria during the high criterion within 10 sessions. Average percent correct responding based on the final three sessions during the low criterion was 73.3% and 93.3% in the high criterion. Scores for the maintenance probes, conducted 30 days following mastery criterion, were 20% correct for the high criterion and 50% correct for the low criterion.

For Participant 2, levels of correct responding during Intervention I remained at baseline levels during the beginning of the initial training sessions. Toward the end of Intervention I, a decreasing trend in correct responses became apparent. Intervention II was then implemented, and the participant achieved 100% correct responding by the fourth training session for both criteria. Participant 2 averaged 100% correct responding based on the final three sessions of both criteria. The maintenance probes conducted 30 days after the participant achieved mastery criteria were 100% and 90% correct for the high and low criteria respectively.

Participant 3 demonstrated rapid rates of acquisition during Intervention I. Following five sessions of training, this participant achieved 100% correct in both criteria. Correct responding during the final three sessions for both conditions averaged 100%. Maintenance probes, conducted 35 days following his achievement

of mastery criteria due to scheduling conflicts, were 70% correct for the high criterion and 30% correct for the low criterion.

### Discussion

The purpose of the current study was to compare the effects of two different mastery criteria on the maintenance of skills. It was hypothesized that the high criterion would result in higher rates of correct responding during a maintenance probe, which was conducted at least one month after the mastery criteria was achieved. This was postulated due to the increased opportunities for reinforcement with the use of additional training sessions associated with the higher criterion. It was also hypothesized that the low criterion would result in lower rates of correct responses during the maintenance probe due to fewer training sessions and reinforcement opportunities.

Participant 1 was the only participant who achieved both the low and high criteria during training. The maintenance probe data show that mastery levels of correct responding were not maintained during either condition. An interesting finding was that Participant 1 scored higher during the low criterion maintenance probe. During training of the low criterion, reinforcement was only available for five sessions when Participant 1 met mastery criteria and training was stopped. During the high criterion, reinforcement was available for ten sessions when mastery criteria were reached. Although reinforcement for correct responding was available for double the amount of sessions in the high criterion, Participant 1 performed better in the low criterion maintenance probe. It is possible that because reinforcement during the maintenance probe was not delivered, the participant may not have been

motivated to respond correctly. Also, Participant 1 may not have been able to retain the information that was taught 30 days prior due to a lack of practice or lack of properly fading reinforcement.

During the training phase, Participants 2 and 3 achieved beyond expectations for the low criterion. They also achieved 100% for the high criterion. There are a number of variables that may have produced these results. First, there was only one target response for each condition. Once the participants learned that the correct response resulted in reinforcement, they may have been motivated to exhibit a correct response. Additionally, Participants 2 and 3 demonstrated that a motor response was going to be expected of them since they pointed to the stimuli during the baseline phase when no motor response was required. Perhaps they had a learning history of being presented with an array of stimuli with reinforcement for touching one of the stimuli. Participant 1, on the other hand, exhibited low to no motor response during baseline and had to be taught the touching response.

Although Participants 2 and 3 did not have a low criterion response, they were kept in the study to use their scores as a comparison for the maintenance effects of a high criterion. Although this was not the original purpose of the study, this information is still beneficial to the literature since little research has been published on the effects of mastery criteria on maintenance skills in children with developmental disabilities.

During Intervention I, Participant 2 exhibited low levels of correct responding. It was noted that he was not attending to the stimuli, and touching any area on the page following the discriminative stimulus with the expectation to receive his edible

reinforcement. The experimenter concluded that Participant 2 learned the contingency that even incorrect responses resulted in prompting and reinforcement. It was hypothesized that this student responded incorrectly because it required less effort to give the wrong response and be prompted to the correct response, than to attend to the task at hand and respond correctly to receive reinforcement. Intervention II was put into place for Participant 2 to reinforce only the correct and independent responses. Once in place, Intervention II resulted in the participant achieving 100% for both conditions in six sessions.

The maintenance probe was conducted 30 days following the achievement of mastery criteria for Participant 2. He maintained the mastery criteria for both stimuli during the 30-day maintenance probe. Correct responding was 100% during the stimuli originally designated as high criteria. For the stimuli that were assigned to the low criterion, the participant achieved 90% correct. He responded incorrectly during 1 trial during the low criteria, and then immediately corrected himself although the experimenter did not comment on the incorrect response. Results from Participant 2 suggest that teaching procedures used during training may have had more of an effect on maintenance performance than setting a higher mastery criterion. The effects of teaching procedures will be investigated further following the discussion involving Participant 3.

After low to chance levels of responding in baseline, training began for Participant 3. The percentage of correct responses increased rapidly during the training sessions. Participant 3 achieved 100% correct responding in both conditions after eight sessions of training. The maintenance probes, conducted 35 days after

mastery criteria was achieved, were 70% correct and 30% correct responding for the high criterion and low criterion, respectively. Although Participant 3 scored 100% for three consecutive sessions in both conditions during training, he scored higher during the maintenance probe of the high criterion. This is an unexpected finding, as we would expect to see similar rates of correct responding during both maintenance probes.

The unexpected and inconsistent results across participants indicate that mastery criteria may not control the maintenance of skills to the degree commonly accepted and that other factors, such as the type of teaching procedure, may be better predictors of long term maintenance of skills. Participant 2 was the only participant to maintain levels of responding at the high and low criteria during maintenance probes. This may be due to the requirements of Intervention II where the participant had to attend and touch the correct stimulus when directed in order to receive reinforcement. Participant 3 was reinforced for any response, independent or prompted. Less effort was required from him to retain what he had learned. In conclusion, greater student effort is needed to learn when teaching procedures require independent and correct responses are used, but worthwhile gains are made in skill maintenance.

A limitation of the current study was the variability of teaching procedures that were used. A prompt delay method of teaching was chosen for this study for its simplicity in training a number of experimenters quickly. Although Participants 1 and 3 had some experience with prompt delays, Participant 2 had limited exposure to this teaching procedure. Errorless teaching was a method more commonly used with

Participant 2 in his regular classroom. This may have been a reason for the slow rates of acquisition during the Intervention I phase of training.

Another limitation was the change in the exclusion criterion that was implemented toward the end of the study. Evidence of acquisition at the conclusion of 10 training sessions was the original criterion that participants needed to demonstrate in order to remain in the study. Participant 2 began to show a decreasing trend towards the end of Intervention I after 10 training sessions. It was during the tenth training session that the experimenter noted that Participant 2 would request reinforcement despite being inattentive and making incorrect responses. Participant 2 was kept in the study, citing fault in the reinforcement contingency that was in place.

A third limitation of the study involved the use of an FR1 schedule of reinforcement. This schedule of reinforcement was used in this study to ensure all participants were being reinforced equally for exhibiting the same response. Typically, rich schedules of reinforcement are faded so students are not made dependent on reinforcers to exhibit correct responses. This was not explored due to this study's experiential purpose.

Results of this study show that mastery criteria may not be a predictor of how a student with developmental disabilities will perform during maintenance trials. Teaching procedures and reinforcement schedules are variables that may influence how a student performs a task after reinforcement for correct responses has ceased. Future research should examine the role of teaching procedures on different tasks, and how students perform those tasks after a period of no reinforcement. It would

also be beneficial to the field of applied behavior analysis if research on the maintenance effects of differing schedules of reinforcement were assessed.

## References

- Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*, 491-498.
- Johnston, J. M. & O'Neil, G. (1973). The analysis of performance criteria defining course grades as a determinant of college student academic performance. *Journal of Applied Behavior Analysis, 6*, 261-268.
- Luiselli, J. K., Russo, D. C., Christian, W. P., & Wilczynski, S. P. (2008). Skill acquisition, direct instruction, and educational curricula. In J. K. Luiselli (Ed.), *Effective practices for children with autism* (p. 196). New York, NY: Oxford University Press.
- Mastery. 2009. In *Merriam-Webster Online Dictionary*. Retrieved March 21, 2009, from <http://www.merriam-webster.com/dictionary/mastery>
- Reiser, R. A., Driscoll, M. P., Farland, D. S., Vergara, A., Tessmer, M. C. (1986). The effects of various mastery criteria in student performance and attitude in a mastery-oriented course. *Educational Communication and Technology: A Journal of Theory, Research, and Development, 34*, 31-38.
- Semb, G. (1974). The effects if mastery criteria and assignment length in college-student test performance. *Journal of Applied Behavior Analysis, 7*, 61-69.

Appendix A

Sample data sheet.

**SCORING**  
**BSL:**  
 += correct & ind.  
 -= incorrect & ind.  
**Train:**  
 += correct & ind.  
 +=p= correct & prompted  
 -= incorrect & ind.  
**MAINT:**  
 += correct & ind.  
 -= incorrect & ind.

D		Date:	prompt delay:	Time:	IOA: Y / N	Score
1.		A	B		C	
2.		C	A		E	
3.		E	A		B	
4.		A	B		E	
5.		E	C		A	
6.		E	D		A	
7.		D	A		B	
8.		B	A		D	
9.		A	D		E	
10.		C	B		A	
Session #: _____		B T M		Experimenter: _____		% correct: _____

E		Date:	prompt delay:	Time:	IOA: Y / N	Score
1.		C	A		E	
2.		A	D		B	
3.		D	E		A	
4.		A	D		E	
5.		D	A		C	
6.		C	A		B	
7.		B	C		A	
8.		A	E		B	
9.		C	A		D	
10.		E	B		A	
Session #: _____		B T M		Experimenter: _____		% correct: _____

F		Date:	prompt delay:	Time:	IOA: Y / N	Score
1.		A	B		C	
2.		E	A		B	
3.		B	A		D	
4.		D	A		B	
5.		A	D		E	
6.		E	D		A	
7.		C	B		A	
8.		E	C		A	
9.		C	A		E	
10.		A	B		E	
Session #: _____		B T M		Experimenter: _____		% correct: _____

A		Date:	prompt delay:	Time:	IOA: Y / N	Score
1.		C	A		B	
2.		E	B		A	
3.		D	A		C	
4.		A	D		E	
5.		D	E		A	
6.		C	A		E	
7.		A	D		B	
8.		A	E		B	
9.		B	C		A	
10.		C	A		D	
Session #: _____		B T M		Experimenter: _____		% correct: _____

B		Date:	prompt delay:	Time:	IOA: Y / N	Score
1.		E	D		A	
2.		D	A		B	
3.		A	B		C	
4.		B	A		D	
5.		E	C		A	
6.		C	A		E	
7.		A	B		E	
8.		C	B		A	
9.		E	A		B	
10.		A	D		E	
Session #: _____		B T M		Experimenter: _____		% correct: _____

C		Date:	prompt delay:	Time:	IOA: Y / N	Score
1.		D	E		A	
2.		C	A		D	
3.		A	E		B	
4.		A	D		E	
5.		D	A		C	
6.		C	A		B	
7.		E	B		A	
8.		C	A		E	
9.		B	C		A	
10.		A	D		B	
Session #: _____		B T M		Experimenter: _____		% correct: _____

## Appendix B

Procedural integrity data sheet used during baseline sessions.

Student: \_\_\_\_\_

Completed by: \_\_\_\_\_

Scoring: Record a + if the experimenter displays the prescribed behavior as stated.  
Record a – if the experimenter does not.

BSL:

Date: \_\_\_\_\_

1.  No session has been conducted with the student for at least a half hour prior to the current session
2.  The correct stimulus set is in the presentation binder as prescribed on the datasheet.
3.  Has the student sit in the chair in the assessment room.
4.  Places stimuli flip book 8-10 inches from the student on the desk.
5.  Presents the Sd (i.e. "Touch \_\_\_\_\_")
6.  Records a + if the student touches the correct stimulus. Records a – if the student touches the wrong stimulus. If the student does not respond within 5 seconds, records a "NR" for no response.
7.  Correct responses are not reinforced and there is no correction procedure during baseline.
8.  Flips to the next page and repeats steps 3 – 5 until session is completed.
9.  Once the student completes all 10 trials, the flipbook is removed

## Appendix C

Procedural integrity data sheet used during training sessions.

Student: \_\_\_\_\_

Completed by: \_\_\_\_\_

Scoring: Record a + if the experimenter displays the prescribed behavior as stated.  
Record a – if the experimenter does not.

Training:

Date: \_\_\_\_\_

1.  No session has been conducted with the student for at least a half hour prior to the current session
2.  The correct stimulus set is in the presentation binder as prescribed on the datasheet.
3.  Has the student sit in the chair in the assessment room.
4.  Places the flip book 8-10 inches from the student on the desk.
5.  Presents the Sd (i.e. Touch \_\_\_\_\_”)
6.  The experimenter prompts the student using the prescribed time delay prompt
  - a. 0s delay: immediately prompts the student to the correct response after stating the Sd.
  - b. 2s delay: wait 2 seconds before prompting the student to the correct response.
  - c. 4s delay: wait 4 seconds before prompting the student to the correct response.
7.  Experimenter records data on the datasheet.
8.  Flips to the next page and repeats steps 4 – 7 (as needed).
9.  Once the student completes all 10 trials, the flipbook is removed

## Appendix D

Procedural integrity data sheet used during maintenance probes.

Student: \_\_\_\_\_

Completed by: \_\_\_\_\_

Scoring: Record a + if the experimenter displays the prescribed behavior as stated.  
Record a – if the experimenter does not.

Maint.

Date: \_\_\_\_\_

1.  At least one month has passed since the participants last session.
2.  The correct stimulus set is in the presentation binder as prescribed on the datasheet.
3.  Has the student sit in the chair in the assessment room.
4.  Places stimuli flip book 8-10 inches from the student on the desk.
5.  Presents the Sd (i.e. "Touch \_\_\_\_\_")
6.  Records a + if the student touches the correct stimulus. Records a – if the student touches the wrong stimulus. If the student does not respond within 5 seconds, records a "NR" for no response.
7.  Correct responses are not reinforced and there is no correction procedure during the maintenance probe.
8.  Flips to the next page and repeats steps 3 – 5 until session is completed.
9.  Once the student completes all 10 trials, the flipbook is removed

Figure Caption

*Figure 1.* Percentage of correct and independent responses across baseline, training, and maintenance probe phases.

