Comparison of Data Collection Methods in Discrete-Trial Training Programs

A Thesis Presented
By
Laura Woods

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 2013
NORTHEASTERN UNIVERSITY

Bouvé College of Health Sciences Graduate School

Thesis Title: Comparison of Data Collection Methods in Discrete-Trial Training Programs

Author: Laura Woods

Department: Counseling and Applied Educational Psychology

Approved for Thesis Requirements of Master of Science Degree
Comparison of Data Collection Methods in Discrete-Trial Training Programs

By

Laura Woods

B.S., Cabrini College

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 2013
Comparison of Data Collection Methods in Discrete-Trial Training Programs

Table of Contents

Abstract……………………………………………………………………………2

Introduction…………………………………………………………………………3

Discrete-Trial Training……………………………………………………………3

Continuous versus Discontinuous Measurement………….……4

Related Articles………………………………………………………………………4

Problem Statement and Experimental Question………………..……6

Method………………………………………………………………………………6

Participants…………………………………………………………………………6

Setting and Materials……………………………………………………………7

Independent Variables ……………………………………………………………7

Dependent Variables……………………………………………………………7

Interobserver Agreement………………………………………………………8

Experimental Design……………………………………………………………9

Procedures…………………………………………………………………………9

Results……………………………………………………………………………11

Discussion………………………………………………………………………13

References……………………………………………………………………….15

Figures…………………………………………………………………………….16

Appendices……………………………………………………………………20
Abstract

The purpose of the current study was to replicate and extend previous studies that compared continuous and discontinuous measurement of discrete trial training responses. Two young girls diagnosed with an autism spectrum disorder participated. An alternating treatments design was used to compare the number of sessions necessary to meet each measurement system’s designated mastery criterion. For both participants, mastery criteria were met faster during the continuous measurement method.

Keywords: continuous measurement, discontinuous measurement, discrete trial training
Comparison of Data Collection Methods in Discrete-Trial Training Programs

Discrete-trial training (DTT) is a recommended Applied Behavior Analysis practice. It is a “method for individualizing and simplifying instruction to enhance children’s learning” (Smith, 2001). With children with autism, teachers and therapists have used DTT to successfully teach new forms of behavior, discriminations, advanced skills, and managing disruptive behaviors. Despite this utility, DTT has several disadvantages. Extra efforts are typically needed to enable the student’s generalization across teachers, settings, and stimuli, implementation can require a significant amount of hours per week, and teachers may need intensive training to deliver it properly.

Smith (2001) listed several important components of intensive DTT. Typically, DTT is conducted in a distraction-free setting. During each trial, an instructional antecedent is presented, which is followed by an opportunity for the student to respond. The instructor then delivers an error correction procedure or the programmed reinforcer, as is appropriate. At the beginning of each trial, the instructor delivers the response; for example “what is it?” paired with any necessary materials. Next, an instructional discriminative stimulus is delivered simultaneously with a most intrusive prompt, which enables the participant to emit a correct response. For example, the teacher may model the desired response, to which the student responds. Afterwards, the teacher delivers a consequence. The consequence could be programmed reinforcement for a correct response, or an error correction procedure for an incorrect response. During the subsequent intertrial interval, the teacher records data, waits until the next trial, and possibly delivers previously learned skills to sustain behavioral momentum. Sessions typically involve collecting ten trials of data for each target behavior, analyzing the number of correct responses, and calculating the correct percentage. As instruction progresses,
the teacher fades out the assistance: i.e., the prompt, until the student responses correctly to only
the discriminative stimulus that reflects the terminal behavior.

DTT data can be collected either continuously or discontinuously. In continuous
measurement, the learner’s responses and prompt levels are recorded for every trial, which
allows for a comprehensive ongoing account of their performance across learning opportunities.
While continuous measurement is a useful direct measure; it can be impractical to use, since the
data collector often must record more than one target response. An alternative that is often used
to estimate behavior is discontinuous measurement or momentary time sampling (Meany-
Daboul, Roscoe, Bourret, Ahearn, 2007). Discontinuous measurement records a portion of the
learning opportunities such as the first-trial only or a few trials (Lerman, Dittlinger, Fentress, &
Lanagan, 2011). Although it is efficient, this method has been found to produce longer intervals
between measures (Hanley, Cammilleri, Tiger, & Ingvarsson, 2007). To date, three studies have
compared the effects of continuous and discontinuous measurement within DTT (Cummings &
Carr, 2009; Najdowski et al., 2009; Lerman, Dittlinger, Fentress, & Lanagan (2011).

Cummings & Carr (2009) compared correct responding during acquisition and follow-up
for continuous (trial-by-trial) and discontinuous (first-trial-only) measurement probes.
Participants were six children with autism spectrum disorders ranging from 4 to 8 years old.
Each child was taught one hundred skills which were randomly assigned to either the continuous
or discontinuous measurement condition. Training took place during twenty-trial sessions, all
ten target trials interspersed with ten previously mastered skills. Data were recorded for either
all ten target trials or the first trial depending on the experimental condition. Mastery criteria
were set at either 100% correct responding for the two consecutive days. This criterion applied
to the number of trials in which data was collected, either all ten trials or for the first trial. Three
weeks after the children achieved skill mastery, follow-up probes were conducted. Results demonstrated that the children reached the mastery criterion, specified for discontinuous measurement quicker than they met the criterion specified for the continuous measurement condition. In the discontinuous measurement condition, participants met the mastery criterion in one hundred and eight fewer trials, which would have resulted in fifty-seven fewer minutes of teaching. However, continuous measurement led to better skill maintenance during follow-up probes. Cummings & Carr’s (2009) recommendations for best practices included utilizing continuous measurement unless session duration is a concern.

Najdowski et al. (2009) replicated and extended the Cummings & Carr’s study by changing mastery criteria. These authors taught eleven children four skills each of which were randomly assigned to the two measurement systems. A more stringent criterion, 100% correct responding for three consecutive sessions, was applied during the discontinuous measurement condition. During the continuous method, 80% accuracy for three consecutive sessions was required. Unlike the Cumming & Carr (2009) participants, the children met mastery in the same number of sessions for both conditions.

Lerman, Dittlinger, Fentress, & Lanagan (2011) replicated and extended the two previous studies. Therapists taught eleven participants two or three skills each. Sessions involved eight or nine instructional trials which were interspersed with maintenance trials. Mastery using an 88% independent correct responding criterion on continuous measurement was compared to two different discontinuous criteria, 100% percent independent correct responding on either the first or first three trials. Three different mastery criteria were compared. Lerman, Dittlinger, Fentress, & Lanagan concluded that their data suggested that first-trial measurement provides a rough estimate of overall performance. However, it led to premature conclusion of skill mastery,
especially when criterion is set at two consecutive sessions. The three-trial subsets corresponded to the continuous measurement with similar changes in performance, especially with specific prompt recordings. Lerman, Dittlinger, Fentress, & Lanagan suggested continuous measurement for a detailed performance account and a monitoring system of teacher’s instruction. However, Lerman, Dittlinger, Fentress, & Lanagan recommended using discontinuous measurement for shorter session durations and stronger data reliability. Future research might investigate the use of a small subset of trials instead of first-trial only, and a more stringent mastery criterion than three sessions is warranted.

The purpose of the current study was to compare the effects of continuous and discontinuous measurement within DTT programs while replicating and extending Cummings & Carr’s (2009) study by using more stringent mastery criteria.

Method

Participants

The participants were Madison and Alexis, both of whom were diagnosed with autism. Madison was 10 years old, nonverbal, and used a static display paper-based communication book. Alexis was 8 years old and a static display paper-based communication book supported her verbal limitations. Both participants were students at a public elementary school, and received intensive one-on-one instruction based on the principles of Applied Behavior Analysis. They attended school for six hours a day, five days a week, and had individualized educational programs (IEPs) that outlined their academic objectives. Madison’s IEP contained a prepositions objective and Alexis’ IEP contained a spelling high frequency words objective. Both girls had previous success learning through DTT methods. Six teachers participated in this study. All of
the teachers had at least one year of experience implementing discrete-trial programs, and four out of the six were pursuing master’s degrees in special education.

**Setting**

All sessions took place in a public school self-contained classroom. Within the classroom, both girls had their own designated work area. Each area contained a table, two chairs, and dividers with limited decorations.

**Materials**

Materials included a data sheet, visual stimuli, clear container, toys, and a procedural integrity data sheet. The data sheets for both participants included directions at the top of the page on each step of implementing the discrete trial, definitions of correct and incorrect behavior, definitions of error correction and reinforcement, and examples of maintenance trials. Materials for Madison’s discrete trials included her communication book, token board and tokens, a clear container that was turned upside down, the same six toys, and large-sized, medium-sized, and small-sized visual stimuli for both prepositions (Appendix 1). Materials for Alexis’ trials included a similar data sheet, a token board and tokens, and dark and blank visual stimuli (Appendix 2).

**Independent Variables**

The independent variables within the current study were continuous and discontinuous measurement of DTT programs.

**Dependent Variables**
The dependent variables were the percentage correct and the number of sessions to the designated mastery criteria for continuous and discontinuous measurement. During discontinuous measurement sessions, a quasi-random method determined the subset of trials for data collection. An appropriate subset was defined as three randomly selected trials from the 3rd to the 8th trials (e.g. 3rd, 5th, and 8th).

**Interobserver Agreement**

Interobserver agreement was assessed on two variables: participant’s response and number of sessions to criterion. Interobserver agreement for correct responding was collected during the training sessions. Interobserver agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements, and multiplying this number by 100. An agreement was defined as both data collectors scoring the same data for the student’s response. Interobserver agreement was collected across 100% of baseline sessions for Madison, and 100% of baseline sessions for Alexis. Interobserver agreement for both Madison’s and Alexis’ baseline sessions were 100%.

During intervention sessions, Interobserver agreement data were collected in 46% of Madison’s sessions and in 59% of Alexis’ sessions. Agreement averaged 93% (range 85% to 100%) during Madison’s sessions and averaged 100% in Alexis’ sessions respectively. Interobserver agreement was also collected during 100% of follow-up probes for both participants and averaged 100% in all cases.

The number of sessions to reach mastery criterion was also measured. A second teacher analyzed the data collected, and recorded the number of sessions until the participants reached the designated criterion (Appendix 3). The same interobserver agreement formula was used as
the one previously stated. An agreement was defined as both teachers recording the same number of sessions until reaching mastery criterion. Interobserver agreement was collected across 100% of intervention sessions and follow-up probes for both participants. Agreements averaged 100% for both conditions and participants.

**Experimental Design**

A concurrent alternating treatments experimental design within participants was to compare the effects of the measurement methods. The time of day the sessions and the measurement method implemented were randomly selected.

**Procedures**

Each session consisted of ten teaching trials interspersed with three to five previously mastered skills. Data were only collected for the teaching trials.

**Baseline**

In the baseline trials, the teacher identified spelling words and prepositions that the participant responded to with accuracy below 33%. On each trial, the teacher set up the materials, presented the initial instruction, waited five seconds for a response, and delivered either reinforcement or no correction for correct or incorrect responses, respectively. After these trials, the teacher presented several trials in which the participants made previously mastered responses to instructions; she reinforced all correct responses.

On the mastered trials, Madison’s instruction was “where is the ______?” and Alexis’ was “spell ______.” For Madison, a correct response was defined as Madison touching the toy one time and then the corresponding icon in her communication book. For Alexis, a correct
response was accurately spelling the requested word. The teacher reinforced responding made on the mastery trial by saying, “Great job!” and delivering a token. After earning ten tokens, each participant received five minutes of an activity that she had indicated was preferred.

Based on the baseline performances, the prepositions “behind” and “next to” and the spelling words “field” and “should” were selected as the stimuli to be presented on Madison and Alexis’ teaching trails.

**Teaching Procedures**

During the teaching trials, the teacher set up the materials, presented the verbal instruction with the designated prompt, waited five seconds for a response, delivered reinforcement or the correction procedure, and recorded the data. Following the teaching trails she presented a varied amount of mastered skill instructions. The teacher faded the prompts after the participant’s responses met the designated mastery criterion.

Stimulus fading techniques were used to teach the responses to the participants. Madison’s prompts were faded in size over three steps, which will be referred to large, medium, and small stimuli. These stimuli were taped to Madison’s communication book, but not simultaneously (Appendix 1). Responses in the presence of the small stimulus were considered independent (Appendix 1).

Alexis’ prompts were printed words which were faded in appearances from black letters to blank in the final stage (Appendix 2). During the teaching phases, the teachers delivered a token and praise such as “that’s next to!” The correction procedure involved removing all materials, setting them up exactly the same way, giving an immediate gesture to the toy then the icon, and saying “that’s next to.” The mastery criterion for continuous measurement was three
consecutive sessions above 90% accuracy, and discontinuous measurement’s was four consecutive sessions at 100% accuracy.

**Follow-Up Probes**

When a skill met mastery criterion, the teachers stopped presenting the stimuli for three weeks. After three weeks, they repeated the baseline conditions. These follow-up probes were repeated until three consecutive sessions above 90% accuracy for continuous measurement and four sessions at 100% accuracy for discontinuous measurement.

**Results**

Total percentage correct for Madison's discrete trial programs is depicted in Figure 1. The circles represent continuous measurement for the target preposition "behind." Madison did not correctly label the preposition during baseline; however, correct responding occurred immediately during the first teaching phase. During sessions 7 and 8 responding decreased to 70% and 90%, which appeared to be caused by lack of attending to the materials. Correct responding was defined as touching the animal then touching the communication icon; however, during these sessions she only touched the icon. When the error correction procedures were implemented, Madison's correct responses quickly increased. Her accuracy remained at 100% for the rest of the medium-sized stimulus, small-sized stimulus, and the three-week follow-up phases. During the continuous measurement condition, Madison met mastery criterion in 11 sessions.

Madison’s percentage correct during the discontinuous measurement condition is represented in Figure 2. The teachers taught the participant the preposition "next to". During baseline, she did not respond correctly. When the large visual stimulus was introduced, Madison
immediately began responding with 100% accuracy. Responding during session 6 decreased to 66% correct due to a confounding variable. For one of the three trials measured, Madison touched the toy, engaged in motor stereotypy, and then touched the communication icon. Following error correction, her correct responding increased. After maintaining 100% accuracy for 4 consecutive sessions, a medium visual stimulus was introduced. Madison only correctly responded on 66% of the trials; anecdotal observations suggest that she did not attend to the toy and the icon. However, her correct responding increased following error correction. Madison's responses remained at 100% for the rest of the medium and visual stimulus phases, and during the three-week follow-up phase. When using discontinuous measurement, her responding met mastery criterion in 15 sessions.

Total percentage correct for Alexis' spelling programs is depicted in Figure 3. The circles represent continuous measurement for the target preposition "should." Alexis did not correctly label the preposition during baseline; however, correct responding occurred immediately during the first intervention phase. During session 7, responding decreased to 80% due to misspelling the target. Correct responding was defined as spelling the word "should"; however, during this session she spelled it "shod" for two trials. However, through error correction procedures Madison's responses quickly increased. Her responses remained at 100% for the rest of the blank visual phase, independent phase, and during the follow-up three weeks after mastery. When using continuous measurement, Madison's responding met mastery criterion within 10 sessions.

Discontinuous measurement of Alexis' spelling programs is represented in Figure 4. The target was "field". During baseline, she did not spell the word correctly. When the dark visual was introduced, she immediately began responding with 100% accuracy. After maintaining
100% accuracy for 4 consecutive sessions, a blank visual was introduced. Madison's responses remained at 100% for the rest of the blank visual phase, independent phase, and during the follow-up three weeks after mastery. When using discontinuous measurement, her responding met mastery criterion in 12 sessions.

**Discussion**

In this study, errorless discrete trial training procedures were used to teach functional academic skills to two young girls with autism. The findings related to number of sessions to mastery criterion were inconsistent with those of Cummings & Carr (2009) and Najdowski et al. (2011). Cummings & Carr (2009) found discontinuous measurement to be quicker, Najdowski et al. (2011) demonstrated no difference between measurements, and the present study determined continuous measurement to be faster.

The results of this study have important implications. For both participants, mastery criterion was met quicker when measured continuously for three consecutive sessions. However, the difference between session masteries was similar for both participants. Madison's responding reached mastery criterion faster by only four sessions, and Alexis's was quicker by only two sessions with continuous measurement. Since the difference is not significant, it could be suggested that both measurement methods can be used when teaching discrete trial training programs. This finding supports Lerman, Dittlinger, Fentress, & Lanagan’s (2011) suggestion of using continuous measurement for complete accounts and discontinuous measurement for data reliability.

There are a number of limitations to this study. First, it can be anticipated that errorless teaching would not lead to a significant difference between numbers of sessions to criteria. However, due to varying number of consecutive days for mastery criterion, results had the
potential to be significant. Programs that required errorless teaching were the only programs available with the chosen participants, because the other IEP objectives would have carryover effects. For example, another option included teaching counting to 1 in the first phase and then counting to 1 or 2 in the second phase. Secondly, this study only analyzed the participants' abilities to reach mastery criterion and maintain the subsequent skill for three weeks. There was limited generalization training with other stimuli, and no generalization across locations.

Madison's objects during her prepositions programs varied; however, the same six toy animals were utilized. Furthermore, data was not collected with Madison's two target prepositions in her communication book at the same time. Since the same toy animals were used for both targets, a true assessment of mastery would involve discriminating between the two concurrently. Alexis' programs did not involve generalization across stimuli or locations. Various methods could have been used such as writing her spelling words, finding correctly spelled words in an array of misspelled options, and practicing the skill in her subseperate classroom and her inclusion classroom.

The results of this study suggest several areas to be studied in future research. Firstly, researchers might use teaching objectives that are not errorless. It would be interesting to demonstrate similar results with a different teaching style. Secondly, research could compare the measurement systems in a variety of ways. Only teaching 100% and 33% of the discrete trials instead of teaching all of the trials despite measurement system could be an option. Also, investigating duration of teaching trials and gathering data collector's feedback would be useful.

The data collected in the current study exemplify the importance of comparing measurement methods within discrete trial training. Thus, it is important to continue this research especially with the frequent use of this teaching procedure.
References


Figure 1. Data for Madison’s programs. Circles represent continuous measurement for percentage correct.
Figure 2. Data for Madison’s programs. Squares represent discontinuous measurement for percentage correct.
Figure 3. Data for Alexis’ programs. Circles represent continuous measurement for percentage correct.
Figure 4. Data for Alexis’ programs. Squares represent discontinuous measurement for percentage correct.
Appendix 1

Madison’s visual stimuli

Large-Sized Stimuli

Medium-Sized Stimuli

Small-Sized Stimuli (Independent)

“Behind” taped into – Madison’s communication book
Appendix 2

Alexis’ visual stimuli

field

should
Appendix 3

Interobserver Agreement Data Sheets:

### Data Collection

<table>
<thead>
<tr>
<th>Trials</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>+</td>
<td>-</td>
<td>I</td>
<td>+</td>
<td>-</td>
<td>I</td>
<td>+</td>
<td>-</td>
<td>I</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Continuous?**  
**Discontinuous?**  
_____% correct

Did your data collection and percentage calculation match the primary teacher’s data?  
Yes  
No

### Number of sessions to mastery criterion

| Continuous mastery criterion: move to next prompt level when it is 90% for 3 consecutive sessions | Percent correct for Session 1: | Percent correct for Session 2: | Percent correct for Session 3: | Percent correct for Session 4: | Percent correct for Session 5: | # of sessions to criterion: |  
---|---|---|---|---|---|---|  
| Discontinuous mastery criterion: move to next prompt level when it is 100% for 4 consecutive sessions | | | | | | |  
| | | | | | |  
| | | | | | |  
| | | | | | |  
| | | | | | |  

Do you agree with primary teacher’s number of sessions?  
Yes  
No