WHAT WERE YOU EXPECTING? OBSERVER EXPECTATION EFFECTS ON DATA RECORDING AS A RESULT OF TRAINING HISTORY

A Thesis Presented to the Faculty of California State University, Stanislaus

In Partial Fulfillment of the Requirements for the Degree of Master of Science in Psychology

By
John Burt
August 2016
CERTIFICATION OF APPROVAL

WHAT WERE YOU EXPECTING? OBSERVER EXPECTATION

EFFECTS ON DATA RECORDING AS A RESULT

OF TRAINING HISTORY

by

John Burt

Signed Certification of Approval page is on file with the University Library
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>List of Tables</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>v</td>
</tr>
<tr>
<td>Abstract</td>
<td>vi</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Reactivity</td>
<td>1</td>
</tr>
<tr>
<td>Observer Drift</td>
<td>2</td>
</tr>
<tr>
<td>Complexity</td>
<td>2</td>
</tr>
<tr>
<td>Expectations</td>
<td>3</td>
</tr>
<tr>
<td>Training</td>
<td>6</td>
</tr>
<tr>
<td>Purpose of Study and Hypotheses</td>
<td>7</td>
</tr>
<tr>
<td>II. Method</td>
<td>9</td>
</tr>
<tr>
<td>Participants</td>
<td>9</td>
</tr>
<tr>
<td>Materials</td>
<td>9</td>
</tr>
<tr>
<td>Procedure</td>
<td>10</td>
</tr>
<tr>
<td>III. Results</td>
<td>14</td>
</tr>
<tr>
<td>IV. Discussion</td>
<td>16</td>
</tr>
<tr>
<td>Research Limitations and Confound</td>
<td>17</td>
</tr>
<tr>
<td>Directions for Future Research</td>
<td>17</td>
</tr>
<tr>
<td>Conclusion</td>
<td>19</td>
</tr>
<tr>
<td>References</td>
<td>21</td>
</tr>
<tr>
<td>Appendices</td>
<td></td>
</tr>
<tr>
<td>A. Informed Consent</td>
<td>26</td>
</tr>
<tr>
<td>B. Training Instructions</td>
<td>28</td>
</tr>
<tr>
<td>C. Test Videos Behavioral Definition and Instructions</td>
<td>33</td>
</tr>
<tr>
<td>D. Debriefing Form</td>
<td>34</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Advanced and Novice Group Observation Errors for Video 1</td>
<td>14</td>
</tr>
<tr>
<td>2. Advanced and Novice Group Observation Errors for Video 2</td>
<td>15</td>
</tr>
</tbody>
</table>
ABSTRACT

There are many biases that affect the collection of data on human behavior. One such bias is expectations. Research has shown that this bias can be reduced sometimes with training. The current study looked at how two different amounts of training time and an expectation statement affected data collection accuracy. In congruence with some previous research the current study found that those with more training were more accurate than those who had less training before the expectation statement was given, $U = 1612, p = .043, r = .18$. After the expectation statement was given the group that received more training was significantly more accurate than those who received less training, $U = 1640, p = .058, r = .17$. However, this may or may not be the result solely of the expectations effect on data collection. Reasons for these findings were discussed.
CHAPTER I

INTRODUCTION

The observation of human behavior has long been a cornerstone of applied behavior analysis (ABA) (Kazdin, 1977). Because a “true” record of behavioral phenomena rarely exists, researchers and practitioners commonly use the consistency, or “reliability” of observers’ data records to estimate behavioral occurrence (Cozby, 2001). One type of reliability, often referred to as interobserver agreement (IOA), involves the systematic comparison of data records to determine the extent to which two observers agreed that behaviors occurred at the same point in time, and can be calculated using several methods (Kelly, 1977). IOA is often expressed as a percentage of agreement, with high agreement coefficients (e.g., 90%) indicating that the record is accurate (Kazdin, 1977).

Human observation, and subsequently IOA, has been shown to be influenced by several variables, including reactivity (Reid, 1970; Romanczyk, Kent, Diament, & O'Leary, 1973), observer drift (Kent, O'Leary, Diament, & Dietz, 1974; O'Leary & Kent, 1973), behavioral complexity (e.g., Mash & McElwee, 1974; Taplin, 1973, Reid, 1970), and expectancies (Kent & O'Leary, 1974; Weiner, 2010; Redfield, 1976).

Reactivity

Reactivity pertains to a change in participant behavior because they are being observed (Heppner, Wampfold, & Kivlighan, 2008). Kent, Kanowitz, O'Leary, and Cheiken (1977) observed the data of 12 participants who were trained in the use of a
nine-category observation code for disruptive classroom behavior. Half of the participants were told that they were being videotaped and that their data collection would be analyzed. On average, the knowledge that the participant was being observed increased reliability 29 points, from 50 percent correct to 79 percent correct.

**Observer Drift**

Observer drift occurs when recording becomes less accurate due to an inconsistent application of the criteria used to record behavior (e.g., change in behavioral definition) (Kazdin, 1978). Kent, O'Leary, Diament, & Dietz (1974) compared the data collected by 9 pairs of observers. Nine observers were shown two videos. In the baseline phase the video was shown with no change in behavioral definition (e.g., a standard operational definition of aggression). However, in the treatment phase the behavioral definition of aggression was changed to be more ambiguous (e.g., less concrete operational definition). An example of this would be in the baseline phase of treatment hitting was defined as any contact with hand to face body, etc. However, in the treatment phase the definition became any “unwanted” touching. Accuracy scores were compared between the baseline and treatment phase. The data were significantly more accurate in the baseline phase than in the treatment phase.

**Complexity**

Complexity refers to the number of target behaviors and/or target persons being assessed by the behavioral coding system, (Wood, 1986) and studies have shown that IOA decreases when complexity increases (Mash & McElwee, 1974;
Reid, 1970). The influence of complexity has been investigated by Mash and McElwee (1974). They trained 48 observers to score dyadic taped verbal interactions using either four or eight response categories. Because the interactions were preprogrammed (meaning participants were already trained) and known in advance by researchers, observer accuracy could be assessed. Observer’s level of accuracy using the four-category system was eighteen percent higher on average than observers using the eight-category system. Thus, the number of response categories in an observational system influences observer accuracy.

**Expectations**

Another influential variable, and one focus of the current study, involves observers’ expectations about the behavioral content that will be observed. Observers are typically exposed to many details about a client prior to recording data, including information on therapeutic techniques, client diagnoses and behavioral histories, and treatment status. Any one of these factors may prime expectations about what will occur during the observation (e.g., being told that a certain behavior is more prevalent.)

In behaviorally oriented research, the motivation variables (such as water to a thirsty person) are called establishing operations (Skinner, 1957; Michael, 1982). Establishing operations (EOs) have two effects: they produce a change in the reinforcing effectiveness of some stimulus, object, or event, and they alter the frequency of the behavior that has been reinforced by that stimulus, object, or event (Michael, 1982; 2000). For example, the EO of food deprivation raises the
effectiveness of food consequences as reinforcers and the EO of food satiation lowers the effectiveness of food consequences as reinforcers. Food deprivation EOs will evoke behavior that has produced food in the past and food satiation EOs will inhibit food seeking behavior.

The behavioral view of expectation related observing involves antecedent behaviors (usually some self-statements about the upcoming event) and how these self-statements relate to what is observed. In essence, these expectations make the expected behavior more salient and perceived to be more reinforcing as a stimulus (an EO effect) and therefore it is recorded more often during observation periods (Michael, 1982; 2000).

Kent and O'Leary (1974) looked at observer accuracy when an observer was told to expect an increase in the behavior they were observing. Half of the observers were told that disruptive behavior would increase in a given video and the other half were not told to expect an increase. All observers were given an operational definition of disruptive behavior and taught how to record this behavior. Also, observers viewed the same classroom videotapes, showing a decrease in disruptive behavior during treatment. In general, the observers who were not told about an increase in disruption recorded a greater reduction in behavior than those who were told to expect an increase. By contrast, Kent et al. (1974) and Redfield (1976) showed that expectations regarding treatment outcomes did not influence observer recording accuracy. Therefore, the data on the influence of expectancies is unclear.

Expectations combined with performance feedback for the observer may also
influence observation. For example, O'Leary, Kent, and Kanowitz (1975) told participants to expect that a treatment involving a token economy would reduce disruption from one phase of treatment to the other (baseline to treatment phase). Observers viewed videotapes of children in a classroom in the two phases of treatment. Both phases of treatment had the same number of behavioral occurrences. The experimenter was in the room with the observers providing positive comments (approval) of recordings if reductions in target behaviors were observed during treatment, and negative comments (disappointment) if no change or an increase in the target behaviors was observed. Results showed that expectations plus feedback had a significant impact on accuracy. There was a significant difference of 50 percent better accuracy among participants when approval and disappointment were not used. A conclusion one might draw from this is that observers can be influenced by misleading feedback.

Cooper, Finley, and Good (1982) studied teachers’ perception of the student’s ability and students’ test scores. Thirteen female third- through sixth-grade teachers in a Midwestern city were observed. The students reading test results were measured using a standardized reading evaluation. Third through fifth graders were administered the Reading Comprehension subtest of the Iowa Test of Basic Skills. Sixth graders were given the Mastery Reading subtest of the Missouri Statewide Testing Program. Results indicated that teachers' perceptions of a student's ability correlated strongly with student achievement (in January $r = .72$; in May $r = .66$). Perceived ability was also related to achievement change. High perceived ability in
January predicted more achievement gain between January and May \((r = .16)\). The data showed a significant discrepancy between perceived and actual ability. In other words, the more overestimated (or less underestimated) the student's ability was in January the more achievement gains from January to May \((r = .24)\). Results indicate that a participant’s perception (expectation) can significantly be different than actual scores.

Research has produced mixed results regarding the influence of expectations on recording accuracy (Kass & O'Leary, 1974; Kent et al., 1974). However, few studies have examined observer characteristics that might interact with expectations to influence recording accuracy. Previous experience might be expected to influence recording accuracy, since both observation and data collection involve complex behavioral repertoires. If making accurate behavioral observations is not part of an observer’s learning history, they may be less attentive to the actual behaviors and more influenced by what others say about the behaviors. Expectations and experience can and do interact. Meaning a person’s experience and expectation can affect how they collect data.

**Training**

A fourth area of influence on behavioral recording involves the characteristics of the training provided to a behavioral observer. These characteristics might include quality or length of the training process. One goal of training is to ensure that observers adhere to the definitions of behavior and record behavior at a consistent level of accuracy (Reid, 1970).
Dempsey, Iwata, Fritz, and Rolider (2012) studied the effect video training has on the accuracy of observations of targeted behaviors. A feature of the video training that made it stand out was the progression from a simple behavioral phenomenon to a complex one. Using a three to six video series they found that those with more training did significantly better than those who received less training.

It is assumed that once training mastery is achieved, observers will continue to apply the same definitions of behavior and record accurately in the future. In other words, observers should be more resistant to biases because of their past training.

(e.g., Kent et al., 1974, 1977; O'Leary & Kent, 1973; Reid, 1970; Reid & DeMaster, 1972; Taplin & Reid, 1973)

**Purpose of Study and Hypotheses**

It has been nearly forty years since the bulk of the research on variables that effect observer reliability was conducted. With the increase in diagnoses of autism spectrum disorders (McKenzie, et al., 2015), and the increased funding for treatment for problem behaviors (self-injurious behavior, aggression etc.) research emphasizing variables that influence observer accuracy is needed because accurate assessments of behavior will help behavioral therapists determine an effective treatment strategy.

The current study examined length of training time on expectation effects related to data collection. It was hypothesized that participants receiving less training would, on average, have a lower accuracy score than the ones receiving more training. It was also hypothesized that observer accuracy scores of the participants receiving less training (Novice group) would decrease as a result of the manipulated
expectation when compared to the scores of the participants receiving more training (Advanced group).
CHAPTER II

METHOD

Participants and Setting

The study involved 128 undergraduate students recruited from a psychology course at a local junior college. With the instructor’s permission, participants were recruited from an undergraduate psychology class. Participants participated during the class period and all participants received experimental credit from their course instructors, as well as the possibility of getting a gift card for their participation. Demographics were not collected from the participants due to time constraints.

Materials

Participants used a paper and pencil recording system. They were trained using DVD based training videos on a computer.

Observer training was conducted using the OBSERVE program (Dempsey et al., 2012), which involved a series of 5-min videos depicting commonly observed behavioral content that systematically increase in the number of target behaviors, behavioral occurrences, and individuals to observe. For example, Video 1 involved 15 occurrences of one target behavior, whereas, Video 3 involved 45 occurrences of three target behaviors. Prior to training, participants were instructed on how to use the observer training and data recording system.
Design

A 2 X 2 mixed factorial group design was used. The two independent variables were amount of training (novice and advanced) and expectation (no expectation and expectation). The repeated measure was expectation. The amount of training differed from Novice to Advanced. Novice received 15 minutes of training while the advanced group received 30 minutes of training.

Novice Group

Observers in this group recorded data on videos 2, 3, and 4 of the video training sequence, and received a history of recording 9 different target behaviors involving 90 behavioral occurrences. Video 4 contained simulated SIB that was to be counted by the participants.

Advanced Group

Observers in this group recorded data on videos 1 through 6 of the video training sequence, and received a history of recording 21 different target behaviors involving 310 behavioral occurrences.

Expectation

Expectation was introduced with the following statement, “You will see a lot of self-injurious behavior in this video” (given to both groups before they viewed the second test video). The dependent variable was the accuracy scores of the participants from two test videos following the last training video. The accuracy scores were derived by comparing the accuracy scores of the participants to the researcher’s score.
Procedure

Overview

All participants attended a single experimental session. All sessions took place in the classroom, containing the typical student desks, chairs, paper and pencil for recording data.

General Training Procedure

Before participating in the study the participants were given an informed consent form and asked to read and sign it (See Appendix A). Participants were then instructed to participate in observation training. Participants were taught using verbal instruction and the OBSERVE program to conduct observations. Two groups were created by randomly assigning participants to either the novice group training condition or the advanced group training condition. There were 59 participants in the Advanced group and 69 in the Novice group.

Observer training was conducted using a video-based observer training system. This system was used in the past to train observers in less than an hour (Dempsey, Iwata, Fritz, and Rolider, 2012). Differential training histories were created in participants by training about one-half of the participants to a “novice” level, and the remaining participants to an “advanced” level.

Participants progressed through the video training sequence using the following procedure to complete each video. First, participants were read a document by the researcher that described the video in terms of its session content along with relevant target behaviors, operational definitions, and measures (Appendix B). The
researcher explained each relevant target behavior, operational definition and answered any questions from the participants before starting the videos. Participants next recorded data on the training video in one five-minute video session while continuing to have access to the written document. This procedure was repeated for each participant for each training video they watched.

Following training, participants in each group recorded data on self-injurious behavior on data sheets (see Appendix C) observed in test video 1. Test video 1 came from the website Youtube.com and showed a young boy under the age of 5 engaging in self-injurious behavior for around 1 minute.

After participants completed the data recording task for test video 1, the researcher made the following expectation statement to each group; “You will see a lot of self-injurious behavior in the following video”. Participants in each group then recorded observations of self-injurious behavior from test video 2. Test video 2 came from the website Youtube.com and was shortened using Edpuzzle.com to show only the first minute of the 15-minute video. This video showed an adolescent boy (under the age of 15) engaging in self-injurious behavior. Upon completion of this task, the completed data sheets were collected (participants passed them down the row). After all data sheets were collected, a debriefing form was passed out (Appendix D). After receiving the debriefing form, participants were excused from the classroom. The Novice group participated first and went through the procedure (viewing all training videos and test videos) while participants in the advanced group waited outside the classroom. When the Novice group finished they were excused and the Advanced
group came in and did the same as the novice group with the addition of the extra training videos. Upon completion of their participation, the Advanced group was dismissed.

**Measure of Outcome and Expectation Manipulation**

The correct number of occurrences was determined by the researcher who viewed the videos several times. He stopped and rewound the videos to check their counts on every occurrence (e.g.: He recorded one occurrence of SIB, rewound the video and double checked then proceeded to the next occurrence and repeated the rewinding process from there on for each occurrence.). The researcher determined that test video 1 had 41 occurrences of SIB and test video 2 had 45 occurrences of SIB.
CHAPTER III

RESULTS

Analyses focused on testing the two hypotheses. The accuracy was calculated using a computer spreadsheet program (Windows Excel) where each participant’s raw score was subtracted from the researcher’s correct score yielding a difference. This difference was how observer error was measured. These error numbers were then analyzed statistically.

A computer based statistical analysis program, SPSS 23 was used throughout the experiment to analyze the data. The error score distributions for both videos were checked for normality using the Shapiro Wilkes test. Since the distribution was not normal, a non-parametric test (Mann Whitney U) was used for both videos. For video 1 a Mann-Whitney U Test indicated that observer error was lower on average among those in the advanced group (Mdn = 21.0) than those in the novice group (Mdn = 24.0), U = 1612.00, p = .043, r = .18. This difference was statistically significant. The r value indicates a small effect size. See Table 1.

Table 1

*Advanced and Novice Group Observation Errors for Video 1*

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mdn</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice</td>
<td>69</td>
<td>24</td>
<td>1</td>
<td>36</td>
</tr>
<tr>
<td>Advanced</td>
<td>59</td>
<td>21</td>
<td>1</td>
<td>37</td>
</tr>
</tbody>
</table>
The Mann-Whitney U Test was used to analyze the difference between the two groups. This indicated that observer error was lower among those in the advanced group \((Mdn = 5)\) than those in the novice group \((Mdn = 7)\), \(U = 1640, p = .058, r = .17\). This result was statistically significant. The \(r\) value indicates a small effect size. See Table 2.

Table 2

*Advanced and Novice Group Observation Errors for Video 2*

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mdn</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice</td>
<td>69</td>
<td>7</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>Advanced</td>
<td>59</td>
<td>5</td>
<td>0</td>
<td>22</td>
</tr>
</tbody>
</table>

To test for a difference among error scores of all participants from video 1 to video 2, a Wilcoxon Signed-ranks test was used. A Wilcoxon Signed-ranks test indicated that the overall error scores from video 2 were less \((Mdn = 5)\) than those in video 1 \((Mdn = 22)\), \(z = 9.23, p < .001, r = .82\). This result was statistically significant. The \(r\) value indicates a large effect size.
CHAPTER IV
DISCUSSION

It was hypothesized that more observer training would produce more accurate observations. This was supported by the results from test video 1. Novice participants were significantly less accurate than Advanced participants. Advanced participants received twice the amount of training than the Novice participants. Hence more observer training produced more accurate observations. This was also found in previous research (Kent et al., 1974, 1977; O'Leary & Kent, 1973; Reid, 1970; Reid & DeMaster, 1972; Taplin & Reid, 1973; Cooper et al., 1982; Dempsey et al., 2012).

It was also hypothesized that the expectation manipulation would have a negative effect on the novice group. This was supported by the results of test video 2. Novice participants were significantly less accurate than Advanced participants. These results are similar to previous studies looking at the effect of expectation on behavior (Kent & O'Leary, 1974), as well as the effect training has on behavior (Kent et al., 1974, 1977; O'Leary & Kent, 1973; Reid, 1970; Reid & DeMaster, 1972; Taplin & Reid, 1973; Cooper et al., 1982; Dempsey et al., 2012).

However, when the error scores of all participants from video 1 and the error scores of all participants from video 2 were statistically examined, it was found that the overall error scores from video 2 were significantly lower than those in video 1. The effect size was large in these results meaning the effect was a reliable one. This indicates that there was a bigger difference between the two videos than between the two groups. Thus, the between
groups effects of the expectation manipulation are confounded with the differences between the two videos and make conclusive interpretation of expectation effects impossible. Previous research (Kent et al., 1974; Redfield & Paul, 1976) report mixed results related to expectation manipulations indicating that more research is needed in this area.

**Research Limitations and Confounds**

**Expectation Manipulation**

A potential limitation was related to the expectation manipulation. It was hypothesized that observer accuracy scores from participants receiving less training (Novice group) would decrease as a result of the manipulated expectation when compared to the participants score receiving more training (Advanced Group). This was true when looking at each group in the same video. For example, when video 2 was viewed after the expectation statement was given, the Advanced group’s scores were more accurate than the Novice group scores. However, the across video comparison revealed that all participants did better on video 2 than video 1. This possibly means that participants simply improved by practice on test videos and that the expectation manipulation did not affect their accuracy. Future research should explore utilizing stronger and more varied expectation manipulations in the test video portion of the experiment to investigate potential effects.

**Directions for Future Research**

Future researchers should look at minimizing the effects of confounds and limitations discussed in the previous section (ex: expectation manipulation).
Another area for future research is using a group approach to training (as done in this study). This approach could be used where greater numbers of observers are needed in applied settings. Future research might discover more efficient and effective group training methods.

Another area for future research is to investigate the nature of the manipulated expectations. Instead of just saying “There will be a lot of SIB in this video”, one could expand upon the verbiage being used and the behavior being studied. One example of this would be to use terms such as “good” or “bad” or “difficult” or “easy” to describe the person being observed. It has been observed by this researcher that these terms are used quite frequently within the behavioral community to describe clients to staff from other staff members or even supervisors to staff. Previous research on expectations has shown this bias to negatively affect accuracy of data collection (Kent & O'Leary, 1974). One could explore the problem we have with expectations related to bias and how this could be solved through more rigorous training (e.g., increasing the time one is trained as was done with the Advanced group in this study).

The types of behaviors being examined should be broadened in future research. SIB is not the only problem behavior. Physical aggression, hand mouthing, pica, just to name a few, could be examined to assess observer training methods and expectations as well as improve how we treat people with these types of problem behaviors. Each of these types of behaviors present their own unique challenges for creating accurate observation and recording training methods.
Conclusion

Despite almost 60 years of work in Applied Behavior Analysis (ABA) there are still many things we don’t know about how biases influence data collection. Based on this researcher’s findings, video training is effective in improving accuracy of data collection. Whether or not it has an effect on expectations is unclear. This means that due to confounds in the current research, one cannot say conclusively that the video based training had an effect on participants that are influenced by an expectation.

The training methods used in this study are useful to data collection tasks everywhere. The approach where participants were first shown simulated behaviors in the training phase and then tested with videos of actual SIB is very effective. Most participants were able to identify and count most of the SIB during test video 1 and 2 in the current study. This means that someone with a maximum of 30 minutes of training with simulations should be able to be accurate in a realistic setting. The amount and intensity of the test video SIB was very similar to what is prevalent in populations of individuals diagnosed with Autism Spectrum disorders and other developmental disabilities.

There is a need to train large numbers of observers at the same time. This study successfully trained a large number of participants with significant results. One may be able to use this training approach to train a larger group than what is typically done (a group of 20 compared to a group of 1 to 5).
REFERENCES
REFERENCES

and experimental field studies at the level of data and empirical concepts.

*Journal of Applied Behavior Analysis, 1*(2), 175-191.
doi:10.1901/jaba.1968.1-175

Cooper, H. M., Findley, M., & Good, T. (1982). Relations between student
achievement and various indexes of teacher expectations. *Journal of
Educational Psychology, 74*(4), 577 doi: 10.1037/0022-0663.74.4.577

Research (7th Ed.).*California: Mayfield Publishing Company.

training revisited: A comparison of in vivo and video instruction. *Journal Of


Hops, H., Walker, H. M., Fleischman, D., Nagoshi, J. T., Omura, R. T., Skindrud, K.,
636-644. doi:10.1037/00220663.70.4.636

Johnson, S. M., & Bolstad, O. D. (1975). Reactivity to home observation: A
comparison of audio recorded behavior with observers present or absent.


APPENDICES
APPENDIX A

INFORMED CONSENT

1. This research study will examine factors that are related to the training of data collectors. If you agree to participate, you will be asked to watch a training video, and then take data based on a behavioral event.

2. You are free to discontinue your participation at any time without penalty. Even if you withdraw from the study, you will receive any entitlements that have been promised to you in exchange for your participation, such as experimental extra credit and possibly an incentive of a ($5 dollar Starbucks gift card).

3. Participation in this research study does not guarantee any benefits to you. However, possible benefits include the fact that you may learn something about how research studies are conducted and you may learn something about this area of research (i.e., factors that are related to the training of data collectors).

4. You will be given additional information about the study after your participation is complete.

5. If you agree to participate in the study, it will take about 60 minutes to complete the study.

6. All data from this study will be kept from inappropriate disclosure and will be accessible only to the researchers and their faculty advisor.

7. The present research is designed to reduce the possibility of any negative experiences as a result of participation. Risks to participants are kept to a minimum. However, if your participation in this study causes you any concerns, anxiety, or distress, please contact the Student Counseling Center at (209) 667-3381 to make an appointment to discuss your concerns.

8. This research study is being conducted John Burt. The faculty supervisor is Dr. Bruce Hesse, Professor, Department of Psychology and Child Development, California State University, Stanislaus. If you have questions or concerns about your participation in this study, you may contact the researchers through Dr. Hesse at (209) 667-3255.

9. You may obtain information about the outcome of the study at the end of the academic year by contacting Dr. Hesse.

10. If you have any questions about your rights as a research participant, you may contact the Campus Compliance Officer of California State University Stanislaus at IRBadmin@csustan.edu.
11. You will be provided with a blank, unsigned copy of this consent form at the beginning of the study.

12. By signing below, you attest that you are 18 years old or older.

13. By signing below, you are indicating that you have freely consented to participate in this research study.

PARTICIPANT’S SIGNATURE: ______________________ DATE: _________
## Video Session 1: Error Correction

Data are recorded on the behavior of the student, who sits across from the teacher. Teacher presents sight words to the student, and student responds correctly or incorrectly. Student incorrect responses result in three correction responses. _Distracters: Card presentation by therapist, incorrect response by student, correction procedure by therapist._

<table>
<thead>
<tr>
<th>Target</th>
<th>Operational Definition</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Response</td>
<td>Correct pronunciation of sight word within 3 seconds of presentation. Score after emission of response (Except during error correct).</td>
<td>Frequency</td>
</tr>
</tbody>
</table>
**Video Session 2: Match to Sample:** Data are recorded on the behavior of the student, who sits across from the teacher. The teacher presents a sample and two comparison cards to the student, and the student responds correctly (choose target) or incorrectly (choose distracter). The student’s correct responses result in praise, and the student’s incorrect responses result in no programmed consequences.

*Distracters: Card presentation by therapists, praise for correct response by therapist.*

<table>
<thead>
<tr>
<th>Targets</th>
<th>Operational Definition</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Correct Response</strong></td>
<td>Pointing to or touching card that matches sample card. Score after emission of response.</td>
<td>Frequency</td>
</tr>
<tr>
<td><strong>Incorrect Response</strong></td>
<td>Pointing to or touching card that does not match sample card. Score after emission</td>
<td>Frequency</td>
</tr>
</tbody>
</table>
**Video Session 3: Classroom Observation:** Data are recorded on the behavior of the target student (female with pony tail) who sits second from front right, and teacher (female with green shirt) who sits at the desk in front of classroom during in-class math work period. The target student’s request for assistance (hand rise) result in verbal attention (assistance) by teacher. Target student also engages in on and off task behavior. 

Distracters: Teacher working at desk, off-task behavior of target student, behavior of other students in class.

<table>
<thead>
<tr>
<th>Targets</th>
<th>Operational Definition</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Communication (hand raise student)</strong></td>
<td>Elevation of either hand above head. Score after emission of response</td>
<td>Frequency</td>
</tr>
<tr>
<td><strong>Assistance (teacher)</strong></td>
<td>Verbal attention to target student by teacher. Score at beginning of verbal emission</td>
<td>Frequency</td>
</tr>
</tbody>
</table>
**Video Session 4: (Demand Session of Functional Analysis):** Data are recorded on the behavior of target child (male) and therapist (female). Therapist initiates instructions using a three-step prompting sequence, consisting of a verbal prompt (followed by a 5-s delay), a modeled prompt (followed by a 5-s delay), and a physical prompt (physical guidance to complete task). Child aggression (hitting) and self-injurious behavior (SIB; head-hit) result in 30-s escape from therapist instruction. Compliance results in praise. Disruption results in no programmed consequences.

*Distracters: Gestural and physical prompts in three-step prompting sequence by therapist, child compliance, child disruption.*

<table>
<thead>
<tr>
<th>Targets</th>
<th>Operational Definition</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIB (head-hit) (child)</td>
<td>Forceful contact between one of participant’s hands and participant’s head from distance of 3” or greater. Score after emission of response.</td>
<td>Frequency</td>
</tr>
<tr>
<td>Aggression (child)</td>
<td>Forceful contact between one of child’s hands and therapist’s body. Score after emission of response.</td>
<td>Frequency</td>
</tr>
<tr>
<td>Verbal Instruction (therapist)</td>
<td>Initial verbal instruction in three-step prompting sequence. Score after emission of response.</td>
<td>Frequency</td>
</tr>
<tr>
<td>Escape (therapist)</td>
<td>Termination of demand by therapist 1 by removing task materials and turning away. Score after emission of response.</td>
<td>Frequency</td>
</tr>
</tbody>
</table>
**Video Session 5: (Functional Communication Training):** Data are recorded on the behavior of target child (male), therapist 1 (TH1, female seated next to him) and therapist 2 (TH 2, female standing next to end of table). TH1 initiates instructions using a three-step prompting sequence, consisting of a verbal prompt (followed by a 5-s delay), a modeled prompt (followed by a 5-s delay), and a physical prompt (physical guidance to complete task) TH 2 prompts child to engage in communicative response (“STOP” card exchange) that results in a break from TH 1’s instruction. Child disruption results in extinction, whereas a card exchange results in 20-s escape from TH 1’s instruction *Distraction: Gestural and physical prompts in three-step prompting sequence by TH 1, movement by TH 2, child compliance.*

<table>
<thead>
<tr>
<th>Targets</th>
<th>Operational Definition</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disruption</strong> <em>(child)</em></td>
<td>Hitting objects from a distance of 3” or greater, pulling items off the wall, ripping items, throwing items a distance of 2” or greater. Score after emission of response.</td>
<td>Frequency</td>
</tr>
<tr>
<td><strong>Independent communication</strong> <em>(child)</em></td>
<td>Handing “STOP” card to TH 1. Score after emission of response. But not if communication prompted.</td>
<td>Frequency</td>
</tr>
<tr>
<td><strong>Escape (TH 1)</strong></td>
<td>Termination of demand by TH 1 by removing task materials and turning away. Score after emission of response.</td>
<td>Frequency</td>
</tr>
<tr>
<td><strong>Verbal Instruction</strong> <em>(TH 1)</em></td>
<td>Initial verbal instruction in three-step prompting sequence by TH 1. Score after emission of response</td>
<td>Frequency</td>
</tr>
<tr>
<td><strong>Prompt (TH2)</strong></td>
<td>Verbal prompt to “ask for a break” by TH 2. Score after emission of response.</td>
<td>Frequency</td>
</tr>
</tbody>
</table>


APPENDIX C

TEST VIDEOS BEHAVIORAL DEFINITION AND INSTRUCTIONS

<table>
<thead>
<tr>
<th>Self injurious behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operational Definition</strong> - any hitting with hand, kicking with leg, head-banging or biting from one person to self (example John hit himself with his palm one time on the face). One occurrence is the start and stop of each defined behavior- Ex: one slap in the face equals one occurrence of self-injurious behavior</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record every occurrence of the behavior in frequency form meaning once an occurrence starts and stops (1 hit =1 occurrence) a slash shall be made in the box below</td>
</tr>
</tbody>
</table>

| Frequency of self-injurious behavior |
Thank you for participating in this study! We are interested in understanding the relationship between training level and the manipulated expectation given to a participant. We predict those who have been trained longer will record more accurately in the training phase as well as the manipulated expectation phase than those who received less training.

All the information we collected in this study will be kept safe from inappropriate disclosure, and there will be no way of identifying your responses in the data archive. We ask that you do not discuss the nature of the study with others who may later participate in it, as this could affect the validity of our research conclusions.

If you have any questions about the study or would like to learn about the results of the study, you may contact me (John Burt) through my research supervisor, Dr. Bruce Hesse, at (209) 667-3255. You may also learn more about the results of the study by attending the thesis defense of this study which will be announced via interuniversity email. If you have questions about your rights as a research participant, you may contact the Campus Compliance Officer of CSU Stanislaus at IRBadmin@csustan.edu. If participation in the study caused you any concern, anxiety, or distress, you may contact the Student Counseling Center at (209) 667-3381 (If a student at CSU Stanislaus) If you are a student at Modesto Junior College please contact the Student Counseling Center at (209) 575 6550

If you would like to learn more about this research topic, we suggest the following references:
