TRAINING VISUAL ANALYSIS USING A CLINICAL DECISION-MAKING MODEL

By

Kailie Jae Kipfmiller

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Applied Behavior Analysis—Master of Arts

2018

ABSTRACT

TRAINING VISUAL ANALYSIS USING A CLINICAL DECISION-MAKING MODEL

By

Kailie Jae Kipfmiller

Behavior analysts visually analyze graphs to interpret data in order to make data-based decisions. Though front-line employees implement behavioral interventions on a daily basis, they are not often trained to interpret these data. A clinical decision-making model may aid front-line employees in learning how to interpret graphs. A multiple-baseline-across-participants design was used to evaluate the effectiveness of a clinical decision-making model on the percentage of correct clinical decisions interpreted from line graphs. All of the participants increased their percentage of correct responses after the introduction of the clinical decision-making model. Two of the 8 participants required additional feedback sessions to reach mastery criterion. The implications of these findings are discussed.

TABLE OF CONTENTS

LIST OF TABLES v

LIST OF FIGURES vi

KEY TO SYMBOLS vii

KEY TO ABBREVIATIONS viii

Introduction 1

Method 9
    Participants and Setting 9
    Materials 9
        Clinical outcomes and clinical decision-making model. 9
        Graphs. 10
        Module development. 12
        Treatment Integrity. 13
    Dependent measures 13
        Correct responses. 13
    Experimental Design and Procedure 14
    Procedure 14
        Baseline. 14
        Intervention. 15
        Brief feedback session. 16
        Intensive feedback session. 16
    Interobserver Agreement 17

Results 18
    Molly 18
    Zane 18
    Ally 18
    Riley 18
    Bre. 19
    Jessica 19
    Nate 19
    Kayla 19
    Error Analysis 20
        Baseline. 20
        Intervention. 20
    Session Duration 21

Discussion..22

APPENDIX...29

REFERENCES...39

LIST OF TABLES

Table 1 Participant demographic and background information. 30

Table 2 Parameters for the autoregressive formula for each clinical decision outcome. 31

LIST OF FIGURES

Figure 1. Examples of the four graph categories: top left, continue; top right, discontinue; bottom left, intervention is complete; bottom right, modify 32

Figure 2. Clinical decision-making model 33

Figure 3. Results for Molly, Zane, and Ally..34

Figure 4. Results for Riley and Bre 35

Figure 5. Results for Kayla, Nate, and Jessica 36

Figure 6. Results for the error analysis for individual participants.......................................37

Figure 7. Results for the average session duration per condition for individual participants 38

KEY TO SYMBOLS

®	Registered Trademark
™	Unregistered Trademark

KEY TO ABBREVIATIONS

BACB	Behavior Analyst Credential Board
RBT	Registered Behavior Technician
BCBA-D	Board Certified Behavior Analyst- Doctoral Level
BCBA	Board Certified Behavior Analyst
BST	Behavioral Skills Training

Introduction

In Applied Behavior Analysis (ABA), experimentation is used to identify the variables responsible for change in behavior (Cooper, Heron, & Heward, 2007). Experimental data are organized, interpreted, and communicated primarily through graphical display. Once graphed, behavior analysts use visual analysis to evaluate treatment effects (Sidman, 1960). Visual analysis entails inspecting graphed data for level, trend, variability, immediacy of effect, overlap, and consistency of data patterns (Gast & Ledford, 2014; Kratchowill, 2013). It is necessary for behavior analysts to conduct visual analysis on an ongoing basis to ensure that applied interventions are producing socially significant behavior change (Baer, Wolf, & Risley, 1968).

When implementing behavioral treatments for children with autism, behavior analysts often train and closely supervise front-line employees who are responsible for the direct implementation of behavior analytic interventions (see Sellers, Alai-Rosales, & MacDonald, 2016). One type of front-line employee is the Registered Behavior Technician ™(RBT®). An RBT is a paraprofessional who practices under the supervision of a Board Certified Behavior Analyst®, a Board Certified Behavior Analyst-Doctoral™, a Board Certified Assistant Behavior Analyst®, or a Florida Certified Behavior Analyst® (Behavior Analyst Certification Board®; BACB®). Entering data and updating graphs is a task that is likely performed by an RBT (BACB RBT Task List, 2016), and some have argued that RBT training requirements should also extend to interpreting data (Leaf et al., 2016). By interpreting data, front-line employees would be able to closely monitor data and be in a position to alert their supervisors, who may only be in contact with the treatment data on a weekly basis, to potential changes to be made to behavioral programming. Therefore, front-line employees may be able to implement more effective therapy for children with developmental disabilities by detecting the need for intervention change much sooner because they access data on an ongoing basis.

Though interpreting data may be an important skill for a front-line employee to have, only a few studies have empirically evaluated methods to teach visual analysis to front-line employees, such as teachers, staff, and students (e.g., Fisher, Kelley, & Lomas, 2003; Maffei-Almodovar, Feliciano, Fienup, & Sturmey, 2017; Stewart, Carr, Brandt, & McHenry, 2007; Wolfe & Slocum, 2015; Young & Daly, 2016). These studies involved training packages using graphic aids, brief trainings, written rules on structured criteria, and instructor modeling, and are described in detail below.

Fisher, Kelly, and Lomas (2003) conducted three experiments that evaluated methods to train staff members to analyze AB single-case design graphs (e.g., consisting of a baseline condition followed by a treatment condition) and identify a treatment effect. In Experiment 1, the authors compared three different methods of visual analysis to determine which methods most accurately detected treatment effects: split-middle (SM) method, dual criterion (DC) method, and the conservative dual criterion (CDC) method. In Study 2, Fisher et al. used verbal and written instructions with modeling to train staff members to use the DC method to accurately detect a treatment effect of AB graphs. Study 3 determined whether the training methods in Study 2 could be incorporated into a PowerPoint presentation to train a large group of staff members to interpret AB graphs.

Over the course of the three experiments, Fisher et al. (2003) found that while both the DC and CDC methods controlled for Type I errors, the CDC method better controlled for Type I errors than the DC method. Fisher et al. also found that a 10 to 15 min training procedure combined with written and verbal criteria and modeling successfully increased the accuracy of the visual analysis of staff members when interpreting these graphs for a treatment effect. Further, this training procedure was successfully incorporated into a format that was used to rapidly train a large group of staff members and improve their accuracy of determining treatment effects.

Stewart, Carr, Brandt, and McHenry (2007) extended Fisher et al. (2003)'s findings by evaluating the effectiveness of two different training procedures on training undergraduates who had no experience with visual analysis to interpret AB graphs. The first training procedure was a 12 min traditional lecture (i.e., a videotaped narrative instruction) on basic elements of visual data analysis based on the Cooper, Heron, and Heward (1987) textbook. The second training procedure was an 8 min videotaped lecture and brief demonstration on the CDC method used by Fisher et al. Results indicated that a traditional lecture on visual analysis did not lead to an improvement in interpreting AB graphs, while the training on the CDC method resulted in undergraduates accurately determining a behavior change using visual analysis. When the CDC criterion lines were removed from the graph, the accuracy of undergraduates determining a behavior change returned to baseline levels, suggesting that this method may only improve accuracy of visual inspection in the presence of visual aids. Both studies by Fisher et al. and Stewart et al. suggested that accuracy of visual analysis can be improved when visual aids (e.g., CDC lines superimposed on the graphs) and brief trainings with feedback are used.

Given that the traditional lecture, alone, in the Stewart et al. (2007) study was not effective in increasing the accuracy of visual analysis by undergraduate students, Wolfe and

Slocum (2015) compared the effectiveness of a recorded lecture, interactive computer-based instruction, and a no-treatment control group on the correct judgements of undergraduates who visually analyzed AB graphs. The interactive computer-based instruction consisted of four modules that provided explanations and demonstrations of skills used when conducting visual analysis (e.g., single-subject research, level change, slope change, and level and slope change) using narration and animation. Each module interspersed 80 mandatory practice items throughout the online modules and provided corrective feedback. During the interactive computer-based component, an experimenter was present to answer questions about technical issues but did not answer questions related to content.

The traditional lecture condition, which served as a comparison to the computer-based instruction, required the undergraduates to read a chapter on visual analysis in the Cooper, Heron, and Heward (2007) textbook and to view a 36 min recorded lecture on the same content from the computer-based instruction (Stewart et al., 2007). The experimenter provided the undergraduates with 20 optional paper-based practice items and was present in the room to answer undergraduate questions on content from the lecture. After each training, the undergraduates completed a series of questions on identifying changes in slope and level in AB graphs. The durations of both trainings were 105 min and both approaches were more effective than no-treatment. Contrary to Stewart et al. (2007)'s findings, results indicated that traditional lectures may be more effective than recorded instruction. The authors suggested these effects could have been due to the presence of the instructor to answer the undergraduates' questions.

Young and Daly (2016) examined the effects of an instructional package that included multiple components (e.g., prompts with and without prompt-delays and contingent reinforcement) to train visual inspection skills to undergraduates without experience in ABA. Participants interpreted AB graphs with and without superimposed CDC criterion lines and identified whether or not there was a treatment effect. Young and Daly's results suggested that the prompts (e.g., trend lines) and reinforcement procedures were effective in increasing the accuracy of identifying a treatment effect when conducting visual analysis, even when the prompts were removed.

Most recently, Maffei-Almodovar, Feliciano, Fienup, & Sturmey (2017) trained special education teachers to make instructional changes based on data-based rules during the visual analysis of discrete-trial percentage graphs. The authors found that behavioral skills training (BST) (i.e., instruction, modeling, rehearsal, and feedback) improved the special education teachers' accuracy of decision-making and reduced errors during graph analysis. The three special education teachers completed the behavioral skills training in 43.4 min, 52.8 min, and 66.2 min, respectively. These results extended previous findings by demonstrating that participants can make data-based decisions through BST.

Although previous research has demonstrated that various instructional packages were effective in increasing the accuracy of detecting treatment effects in graphs when using visual analysis, there are several limitations worth noting. One limitation is that research on training visual analysis only examined AB single-case design graphs. AB graphs, while certainly relevant and important, may not be the type of graphs that front-line employees likely come into contact with on a daily basis. Front-line employees may also likely view data within a single condition as opposed to ‘A’ (baseline) and ‘B’ (treatment) conditions separated by a phase change line. However, no studies exist that have evaluated teaching visual analysis of a single condition.

A similar limitation is that previous studies trained the participants to detect whether there was a treatment effect. It may be more clinically relevant to train front-line employees to use visual analysis to evaluate continuous time-series graphs for treatment progress and subsequently identify an appropriate course of clinical action (e.g., to discontinue an ineffective intervention). Another potential limitation may involve the cost of the training procedures reported in the above studies; the training procedures lasted anywhere from 8 min to 105 min and often required the physical presence of the trainer. Simplifying the necessary training resources and reducing reliance on trainer feedback is important to reduce cost associated with visual analysis training.

Given the above limitations of previous research, this study evaluated the effect of a low-cost resource delivered through an online platform to teach visual analysis of time-series graphs. Previously reported training methods, such as BST and traditional lectures, require a time-intensive approach to train employees to implement behavior analytic interventions with procedural integrity (Downs, Downs, & Rau, 2008). The cost of these resources, in terms of time and money, combined with the high rate of employee-turnover in the behavioral health and school settings result in time and money lost for the employer (LeBlanc, Gravina, & Carr, 2009). Online training programs, or interactive computer trainings, are a low-cost and time-efficient alternative training package once they are developed. Online training programs may also allow for a dissemination of information to a large number of people (Fisher et al., 2003), and have been demonstrated to teach staff to conduct behavioral interventions (e.g., Wishnowski, Yu, Pear, Chand, & Saltel, 2017). Given this information, and the demand for training a large number of RBTs, the first purpose of this study is to evaluate the effects of a computer-based training program in teaching front-line employees to conduct visual analysis.

One strategy that may be useful to streamline methods to teach front-line employees to interpret data and make decisions in applied settings is a clinical decision-making model. A clinical decision-making model asks a series of systematic questions that can be answered sequentially and may lead to optimal clinical decisions. Clinical decision-making models are becoming prevalent in behavior analytic literature, and there is a growing interest for using these models in a variety of clinical applications. For example, models have been developed for selecting between multiple function-based treatments for attention and escape-maintained behavior, as well as selecting between discontinuous measurement procedures for problem behavior (Fiske & Delmolino, 2012; Geiger, Carr, & LeBlanc, 2010; Grow, Carr, & LeBlanc, 2009). Additional authors, such as Brodhead (2015), proposed a clinical decision-making model for deciding when to address non-behavioral treatments while working with an interdisciplinary team, and LeBlanc, Raetz, Sellers, and Carr (2016) proposed a model for selecting measurement procedures given the characteristics of the problem behavior and restrictions of the environment. Though multiple clinical decision-making models have been developed to aid in behavior-analytic practice, to our knowledge, none have been empirically evaluated.

Therefore, the second purpose of this study is to develop a clinical decision-making model on visual analysis and empirically test the effects on decision-making. This evaluation may provide an example of a framework for empirically evaluating clinical-decision making models, which is much needed in behavior analytic literature and also provides a low-cost alternative to other forms of training (e.g., BST).

Method

Participants and Setting

Prospective participants completed a paper-based demographic survey prior to the study. The survey asked participants to describe basic demographic information, including age, gender, highest degree obtained, and academic major(s). Participants were also asked to report the number of courses taken involving behavior analysis, the name of the course, the number of months providing behavior analytic services to clients (if applicable), and to describe any prior training in visual analysis of graphs.

Eight adults (two males, six females, M_{age} = 25.8 years) participated in this study. Participants were required to complete this training as part of a 40 hr training at a university-based EIBI center serving children with autism spectrum disorder. According to a demographic survey the participants completed prior to the beginning of the study, Kayla reported classes in behavior analysis, and having experience in providing behavior analytic services to clients for 10 months. Ally, Bre, Nate, Molly, Riley, and Zane reported having some experience with graphs; however, none of the participants received specific training on the visual analysis of line graphs. Full participant demographics are reported in Table 1.

All participants completed the baseline and intervention conditions in their own home with their own personal computer. An experimenter was not present during experimental sessions.

Materials

Clinical outcomes and clinical decision-making model. The clinical outcomes and the clinical decision-making model was developed by the experimenter and two doctoral-level behavior analysts who have over 20 combined years of experience in behavioral treatments and visual analysis for young children with autism. The four clinical decisions to make upon visual inspection of graphical data were: a) continue intervention b) discontinue intervention c) modify intervention or d) intervention is complete. Graphs in the *intervention is complete* category consisted of an upward trend with or without variability and had three consecutive data points at 80% or above on the y-axis. Graphs in the *continue intervention* category consisted of an upward trend with or without variability and did not have three consecutive data points at 80% or above on the y-axis. Graphs in the *discontinue intervention* category consisted of a flat trend with or without variable data points ranging from 20% to 50%. Finally, graphs in the *modify intervention* category consisted of highly variable data (i.e., data points that ranged unpredictably from 20% to 100%). See Figure 1 for an example of the four types of graphs.

The clinical decision-making model (see Figure 2) sequentially asked and answered questions that lead the participant to an optimal clinical decision based on the analyzed data. For example, the participant may ask himself, “Are data trending upward?”. If yes, he would follow the corresponding arrow and continue on to the question “Are the last three data points at 80% or above?”. If the answer was no to the initial question, the participant would follow the corresponding arrow to the question “Are data trending downward?”. This process of asking and answering questions would continue until one of the four clinical outcomes was left.

Graphs. Graphs were generated prior to the study in Microsoft Excel™ using an autoregressive model similar to the one used by Wolfe and Slocum (2015), which was a modified version of the formula used by Fisher et al. (2003). Two-hundred and forty graphs were generated. Each graph consisted of 10 data points to represent a child's hypothetical percentage of correct responding during a 10 trial-block teaching session (e.g., 8 independent responses out of 10 trials resulted in a score of 80%). The x-axes were labeled "Sessions" and the y-axes were labeled "Percentage of Correct Responses". Grid lines were present for all graphs at intervals of 20. There were no phase change lines or figure captions presented for any graphs.

The four general categories of graphs were developed using the following procedures. First, the autoregressive equation was manipulated to identify parameters that would result in graphs that fit into one of four pre-determined categories (see Table 2). Starting with the first category, the predetermined parameters were input for that category into the autoregressive equation to generate at least 60 graphs. This process was then repeated for the remaining three categories.

After the graphs were initially developed for all four categories, graphs were converted to JPEG images, individually numbered from 1 to 240, and placed a folder. Then, a random number generator was used to select a graph from the folder for analysis. The first author and two experts (second and forth authors) examined the graph using the clinical decision-making model and categorized the graph into one of the four clinical decision outcomes. This process was repeated until all 240 graphs were sorted. There were two purposes to this procedure. The first purpose was to test the clinical decision-making model to ensure it worked as it was intended to. The second purpose was to reduce the probability that graphs were incorrectly sorted as a result of experimenter bias.

Both of the experts were BCBA-Ds with extensive training and practice using visual analysis to make clinical decisions. In the event a graph was generated that could not be categorized (e.g., we identified a graph with data points greater than 100%) we replaced that graph with a newly generated graph. A total of sixty graphs within each category were developed.

Module development. To evaluate the effects of the clinical decision-making model on the accuracy of clinical decisions, two modules in the learning platform Desire2Learn (D2L) ® were created. Module 1 was presented in a quiz format. Each question contained one of the 60 available graphs from one of the four pre-determined categories, and the instructions "Please select the correct answer", along with four possible answers that were presented in a randomized order: a) Intervention is Complete, b) Continue Intervention, c) Discontinue Intervention, d) Modify Intervention. Participants were able to view only one question at a time, and once they answered a question, they were presented with the next question. They were not able to return to any of the previous questions and were allowed unlimited time to complete the quiz.

The quiz was programmed to randomize the order of graphs from each category and was programmed so each of the 60 graphs, within each category, had an equal probability of being presented. Every session, each category of graphs was presented three times, for a total of 12 trials for each session completed of Module 1.

Module 2 was identical to Module 1, except the clinical decision-making model was simultaneously presented along with each question containing a graph. Participants were also given the opportunity to download the clinical decision-making model to accommodate participants with low-resolution or small computer monitors.

Prior to the beginning of the study, we tested both Module 1 and Module 2 to ensure the following: instructions appeared on the screen prior to starting the quiz, every page included the instruction “Please select the correct answer”, one graph was shown per question along with the four outcomes, the clinical decision-making model was shown above the graph in Module 2, and the participants’ results were not available for the participant to view after the quiz was completed.

Treatment Integrity. We measured treatment integrity to assess the degree that the module parameters appeared correctly. The observer recorded whether all components appeared during each question for each session completed (i.e., one graph, four answers, and the clinical decision-making model, if applicable). The experimenter summed the number of questions that appeared without error for all conditions and divided that number by the total number of questions. This number was multiplied by 100 to yield a treatment integrity percentage score.

We collected treatment integrity data for 100% of sessions across participants. Zane, Riley, and Nate had integrity of 100% across both conditions. Molly and Jessica’s integrity was 99.0% (range: 91.7-100%). Ally’s integrity was 99.2% (range: 91.7-100%). Bre’s integrity was 99.2% (range: 91.7-100%), and Kayla’s fidelity was 99.4% (range: 91.7-100%). Overall, the average fidelity was 99.1% (range: 98.8–100%).

Dependent Measures

Correct responses. The dependent measure was percentage of correct responses. A correct response was defined as any instance when a participant's answer to a quiz question matched to the correct clinical decision that corresponded to the graph depicted in that question. Percentage of correct responses was calculated by diving the number of correct responses by the total number of questions in a module (n = 12), multiplying it by 100, and converting the result into a percentage.

Experimental Design and Procedure

A multiple baseline across participants design was used to evaluate the effectiveness of the clinical decision-making model on selecting clinical decisions when conducting visual analysis.

Procedure

Baseline. The experimenter provided the participants with a document containing instructions explaining how to create a username and password to log onto D2L. After participants created a username and password, the experimenter sent an email with instructions to complete a pre-determined number of sessions of Module 1.

The first screen that followed the log in screen displayed the D2L homepage. The participants were instructed to select the link labeled "Visual Analysis". The next screen displayed the main screen with both Module 1 and Module 2. The participants were instructed to select the link labeled Module 1 beneath the header labeled "Content Browser". Module 2 was locked, and participants were not able to view Module 2 until completing baseline. The screen following the main screen provided the participants with an introduction and instructions prior to beginning the quiz in the baseline condition. The instructions were presented through D2L and specifically read:

> In this module, you will be shown a series of graphs and asked to select between 1 of 4 possible clinical decisions, based on information in that graph. Please answer each question to the best of your ability. **You may not use outside materials, and you may not consult with your peer**s (emphasis added). Please complete this quiz on your own.

Participants were then presented with a link labeled "Start quiz" at the bottom of the screen. Upon starting the quiz, the first question displayed a graph with the instructions "Please select the correct answer". Four clinical decision-making outcome options were provided beneath the graph. Once the participants selected an answer, they selected a link labeled "Next" to move on to the next question. Participants could not review previous questions and no corrective feedback was given after each answer. This process continue until 12 questions were completed. The participants were able to click on the button labeled "Submit Score" once they answered all 12 questions containing graphs. The participants never received any feedback on the accuracy of their answers during Module 1.

Intervention. Upon completion of baseline, participants were instructed to complete a pre-determined number of quizzes in Module 2. Module 2 was similar to Module 1 with the following modifications. Each question contained the clinical decision-making model. Prior to the beginning of each quiz, the participant was presented with the following instructions:

> In this module, you will be shown a series of graphs and asked to select between 1 of 4 possible clinical decisions, based on information in that graph. With each graph, we have provided you with a visual aid. Please use the visual aid to answer each question to the best of your ability. Because some computer monitors may have low resolution, we

recommend you also download the visual aid (link below). Please disable your pop-up blocker before you download the visual aid.

A hyperlink was available in the instructions for the participants to download the clinical decision-making model. Otherwise, this condition resembled that of baseline. The participants never received any feedback on the accuracy of their answers during Module 2.

Brief feedback session. Participants received brief feedback if they did not improve percentage of correct responding during the intervention phase. Nate and Kayla were the only participants exposed to this session. In the feedback session, the experimenter presented the participant with a paper-based copy of the clinical decision-making model and four randomly-selected graphs from each clinical decision category (i.e., intervention is complete, continue intervention, discontinue intervention, and modify intervention). The experimenter then asked the participant to vocally explain his or her decision at each point of the clinical decision-making model to come to a clinical decision regarding each graph. The experimenter addressed questions (e.g., defined variability and described examples of variable data) and errors in decision-making by redirecting to specific boxes in the clinical decision-making model until the participant reached the correct decision. The feedback session occurred only once, and lasted no longer than 5 min.

Intensive feedback session. A second feedback session occurred if the participants did not increase accuracy of correct responding after the first feedback session. Kayla was the only participant exposed to this condition. In this condition, the participant completed one session of Module 2 with the experimenter (however, this module was not scored or depicted in Kayla's results). The participant used the same talk-out-loud procedure as in the brief feedback session;

however, the participant completed a total of 12 graphs instead of four. The experimenter clarified questions (e.g., data may be variable and also have an upward or flat trend) and provided feedback when necessary.

Interobserver Agreement

An independent observer used participant data exported from D2L to ensure the answers to each quiz question were transcribed correctly. An agreement was defined as an identical score recorded by D2L and the independent observer. A disagreement was defined as any instance where the observer did not score the same answer scored by D2L. Interobserver agreement was calculated by dividing the number of agreements by the total number of agreements plus disagreements and multiplying by 100 to yield a percentage. The experimenter used a random number generator to select 30% of each participant’s sessions across all conditions. Once a session was selected, the observer examined each individual graph to ensure that D2L accurately scored the answers as correct or incorrect. Total Interobserver agreement across all dependent variables was 100%. That is, there were no errors in transcription.

Results

Figures 3, 4, and 5 depict the results for participants' correct responses when making clinical decisions based on graphs. Five of the seven participants' performance increased after the introduction of the clinical decision-making model. Kayla and Nate required additional feedback to increase their correct responding. Results for each participant are described below.

Molly

During baseline, Molly averaged 20% correct responding across five sessions, (range: 16-41%). After the introduction of the intervention, the first seven sessions averaged 88% (range: 83-92%). The subsequent four sessions averaged 98% (range: 92-100%).

Zane

During baseline, Zane averaged 50% correct responding across eight sessions (range: 33-67%). After the introduction of the intervention, we observed an immediate increase in correct responding. In six sessions, two sessions were at 92% and four sessions were at 100%.

Ally

During baseline, Ally demonstrated variable responding, averaging 33% across 12 sessions (range: 33-83%). The first six sessions averaged 50% (range: 33-67%). The subsequent six sessions averaged 71% (range: 50-83%). After the introduction of the intervention, the first session was 83%. The subsequent seven sessions averaged 95% (range: 92-100%).

Riley

During baseline, Riley averaged 72% correct responding across five sessions. The first session in baseline was 92%. The subsequent four sessions were 67%, 75%, 67%, and 58%, respectively. Riley demonstrated an immediate increase in correct responding after the introduction of the intervention. We observed stable responding at 100% across three sessions.

Bre

During baseline, Bre averaged 72% correct responding across eight sessions (range: 58-83%). Bre demonstrated an immediate increase in correct responding after the introduction of the intervention. We observed stable responding at 100% across three sessions.

Jessica

During baseline, Jessica averaged 67% correct responding across ten sessions (range: 50-100%). Sessions one, two, three, and four were 83%, 67%, 50%, and 58%, respectively. Session five was 100%, with the subsequent five sessions averaging 63% (range: 50-83%). After the introduction of the intervention, the first two sessions were 75%. The third session was 67%, and the subsequent three sessions were 100%.

Nate

During baseline, Nate averaged 67% correct responding across eight sessions (range: 50-75%). After the introduction of the intervention, the first session was 50%. The subsequent nine sessions averaged 80% (range: 67-100%). Following the 18th session, Nate was exposed to the brief feedback condition. The first session in the brief feedback condition was 100%, with the subsequent four sessions averaging 96% (ranging 92-100%).

Kayla

During baseline, Kayla averaged 61% correct responding across five sessions (range: 50-83%). After the introduction of the intervention, the first two sessions were 83%. Sessions three and four were 100%. The subsequent six sessions averaged 86% (range: 75-92%). Kayla was exposed to the Brief feedback condition following session 15. Following the brief feedback session, Kayla averaged 93% correct responding across six sessions (range: 83-100%). Kayla was then exposed to the intensive feedback condition after session 21. The first session following the intensive feedback condition was 100% and the second session was 92%. The subsequent three sessions were 100%.

Error Analysis

An independent observer recorded the questions that each participant marked incorrectly for each session. The observer recorded the category of the questions missed (i.e., intervention is complete, continue intervention, discontinue intervention, modify intervention). The percentage of errors was calculated by dividing the number of errors for a particular category by the sum of total errors for individual participants. These results are summarized in Figure 6.

Baseline. There were three patterns of responding observed in baseline. Jessica, Kayla, Ally, Molly, and Zane made errors across all categories in the first pattern of responding, with the least percentage of errors occurring in the discontinue intervention category, except for Zane, who had the highest amount of errors in this category. Bre and Riley had a higher percentage of errors in the intervention is complete and modify intervention categories. Nate had the highest percentage of errors during the intervention is complete category compared to the other three categories.

Intervention. Across all participants, the highest percentage of errors made during intervention occurred for the discontinue intervention category, and the least percentage of errors were made in the intervention is complete category. Jessica and Kayla made the most errors during the discontinue intervention category. Kayla and Molly made errors in the discontinue intervention, intervention is complete, and modify intervention categories. Zane made similar percentages of errors in the continue intervention and modify intervention categories, and Nate made errors in all four categories, with the highest percentage of errors in the modify intervention category. Bre and Riley made no errors in the intervention conditions.

Session Duration

The average session duration for each participant was calculated by taking the sum of the duration per session and dividing by the total number of sessions for each condition (see Figure 7). The longest and shortest durations for each condition were removed for each participant to account for outlier data that may have occurred (e.g., leaving the computer and coming back to the session at a later time). For all eight participants, the average time to complete a session was 5 min 36 s (range: 1 min 54 s-15 min 18 s) in baseline. The average time to complete a session in intervention was 5 min 44 s (range: 2 min 14 s-8 min 33 s). For Ally, Bre, Jessica, Nate, Riley, and Zane, the addition of the flowchart increased the mean length of time to complete each session by 2 min 45 s (range: 38 s-3 min 9 s). For Kayla and Molly, the addition of the flowchart decreased the mean length of time to complete each session by 7 min 15 s and 4 min 26 s, respectively.

Discussion

This study evaluated the effect of a low-cost resource delivered through an online platform to teach visual analysis of time-series graphs. A secondary purpose of the study was to empirically test the effects of the clinical decision-making model on decision-making. The results support the effectiveness of using a clinical decision-making model to increase the percentage of correct clinical decisions when using visual analysis to interpret line graphs. All eight of the participants increased accuracy of correct responding after the introduction of the clinical decision-making model. Six of the 8 participants increased their accuracy without any feedback or instruction from an instructor. Nate and Kayla required additional feedback from an instructor. Implications of these findings, along with directions for future research, are discussed below.

Between the baseline and intervention conditions, we observed differences in immediacy of treatment affect across participants. Riley, Zane, and Bre demonstrated immediate treatment effects once exposed to the clinical decision-making model, and Molly, Ally, and Jessica demonstrated a delayed treatment effect. The exact reasons for this obtained difference in immediacy of effect is unclear. It is possible that previous experiences in visual analysis (e.g., involvement in science-based laboratories) in undergraduate courses may be one explanation for the immediacy of effect for Zane, Bre, and Riley. However, Molly and Ally also reported previous experience in visual analysis. Another explanation for differences in immediacy of effect within the intervention condition may be that feedback may have been inadvertently provided to participants who did not immediately achieve correct responding by requiring them to complete additional sessions. That is, by asking participants to complete more sessions, they may have “understood” they were doing something wrong. However, asking Nate and Kayla to complete extra sessions did not appear to affect their responding. Finally, it is possible that some participants required additional time with the model to understand which decision to select based on the information given. For example, Jessica initially did not show improvement when given the clinical decision-making model, then demonstrated 100% accuracy in responding for the last three data points. Jessica reported in an informal conversation, following the completion of the study, that she initially read the box stating, “Are the last 10 data points at 50% or below?” on the clinical decision-making model incorrectly for the first three sessions. She then realized this mistake and subsequently responded correctly the remaining three sessions. In summary, the exact reasons for these differences in immediacy of effect across all participants are unclear and present a future area of inquiry in future research.

The remaining two participants, Nate and Kayla, required an additional brief feedback session, while Kayla required an intensive feedback session following the brief feedback session. These results are consistent with Stewart et al., (2007), Wolfe and Slocum (2015) and Young and Daly (2016), where feedback was provided to some participants to increase accuracy of responding, as well as previous research on training front-line employees (e.g., Higbee et al., 2016; Pollard, Higbee, Akers, & Brodhead, 2014). During the feedback sessions, both Nate and Kayla reported some confusion on the definition of “variable” and ‘flat”. Once the experimenter provided definitions and examples, Nate increased his percentage of accurate decision-making and subsequently completed the study. During the intensive feedback session with Kayla, the experimenter provided neutral affirmation for the correct responses and provided corrective feedback for the graphs she did not answer correctly. Specifically, the experimenter provided the feedback that data may still trend upward or be flat although data are also variable. Kayla also indicated that additional variables (e.g., completing the modules late at night, managing family responsibilities) may have affected her accuracy when completing the sessions.

Given the information obtained during feedback sessions with Nate and Kayla, another explanation for differences in responding, across participants, may be that some participants found it difficult to detect trends (e.g., upward trend, downward trend, or no trend) while observing variable data. This finding is in line with previous research that found novice raters accurately detected the trend when there was little to no variability in the data, but accuracy decreased as the variability of the data increased (Nelson, Van Norman, & Christ, 2016). Therefore, it might be necessary to initially teach participants to detect trend lines of graphs with and without variability. Teaching participants this skill will potentially eliminate the need for additional feedback sessions following the introduction of the clinical decision-making model.

The error analysis conducted for each condition provided useful information about the percentages of errors the participants were making in each clinical decision outcome category. During the baseline condition, participants made the most errors when shown a graph that was defined as the intervention is complete. After the introduction of the clinical decision-making model, the participants made the least percentage of errors in this category, with participants increasing accuracy to an average of 88.5%. These results suggest that the clinical decision-making model was most effective in providing the participants with clear guidelines of when to consider an intervention complete. However, after the introduction of the clinical decision-making model, the average percentage of errors for discontinue the intervention increased by 15.5%. This number is likely somewhat affected due to Jessica initially misreading the clinical decision-making model (as described above). Kayla also made the most errors in this category. A possible explanation may be that the discontinue the intervention graphs contained variable data that may have been confused with another category, modify the intervention. It is possible that Kayla was only attending to the variability of the data and not attending to where the data points lie (i.e., she may have selected modify intervention in the presence of variable data, despite the data points being below 50%).

The remaining participants made similar percentages of errors during the continue intervention and modify intervention categories. This finding may be due to participant errors in detecting trend when data were variable, not attending to the grid lines on the graphs before selecting an answer, or difficulty in determining whether the data point fell above or below the 80% grid line. For example, when observing three consecutive data points to determine if the intervention was complete, one data point may have appeared to be located at the 80% grid line, when in fact it was below 80%.

The error analysis data provides information on strengths and modifications needed to the clinical decision-making model. Based on the error analysis, the percentage of errors decreased for the intervention is complete category after the participants were given the clinical model. As a result, a strength of the clinical model could be that the participants could easily detect when an intervention was complete based off the graphs. A limitation of the model could be our use of the terms “flat” and “variable”. In the error analysis, the highest percentage of errors occurred for the discontinue category. A likely possibility for this finding may be that the participants selected modify intervention any time they saw that the data were variable, when the correct answer would have been to discontinue intervention.

The intervention in the current study provided a brief, low-cost alternative to other training methods, such as BST or other trainings where the instructor would be present. In previous research, large group trainings took approximately 10 to 15 min (Fisher et al., 2003) and 8 min to 12 min (Stewart et al., 2007) to complete. The current study's average duration to complete a session using the clinical decision-making model across all participants was 7 min 57 s. In other research studies that utilized alternative training methods, time to train participants lasted anywhere from 43.3 min (Maffei-Almodovar, Feliciano, Fienup, & Sturmey, 2017) to 105 min (Wolf & Slocum, 2015). Given this information, the results of this study suggest that the clinical decision-making model is more time efficient than training in previous research, and does not require the presence of an instructor, unless additional participant feedback is needed.

There are a few limitations of this study. First, we used an autoregressive formula (Fisher et al., 2003; Wolfe & Slocum, 2015; Young & Daly, 2016) to generate graphs based on hypothetical rather than actual student data. Graphs were randomly generated to remove experimenter bias. Thus, it is unclear if the participants could generalize their performance to actual student data, and future research may consider adding in a condition where the participants analyze actual data (e.g., Maffei-Almodovar, Feliciano, Fienup, & Sturmey, 2017). Another limitation may be that our participants analyzed completed data sets, whereas front-line employees will more frequently visually analyze on-going data sets and possibly make clinical decisions sooner than after ten sessions. Future research may evaluate conducting visual analysis on an on-going basis to more closely resemble applied practice.

The content of the clinical decision-making model may need to be operationally defined to reduce potential participant confusion. For example, some participants (e.g., Kayla) reported they were confused about the terms “variability” and “flat trend” and did not know that data paths could be variable and still be a flat trend. Slight modifications of the clinical decision-making model may result in fewer errors.

Also, another limitation is that the experimenter was not present for the training. Because of this, it is it is unclear how long the participants referred to the clinical decision-making model and whether or not they used outside sources, such asking friends, for answers. However, all participants demonstrated changes in behavior when the clinical decision-making model was presented, therefore, if participants did access outside sources, those sources likely had negligible effects A final limitation is that the clinical decision-making model was not removed in the current study. Therefore, it is unclear if participant accuracy would have maintained in absence of the model. It is worth noting, however, that the model was meant to be permanent and could be considered a permanent reference for front-line employees. Relatedly, it is unclear how much the participants were relying on the model each session, especially towards the end of the study. It is possible that once participants learned the “rules” conveyed in the model, they decreased their reliance on that model. Regardless, future research may examine whether or not the clinical decision-making model can be removed while maintaining participant’s accuracy of responding.

With regard to the hypothetical graphs we generated, it is unclear as to whether or not the gridlines assisted in each participant's data analysis. Future research may evaluate whether displaying gridlines on the graphs contributed to the accuracy of decision-making upon visual analysis. Second, the graph size of the modules changed in resolution based on the participant's computer. During the current study, a hyperlink was provided to access the clinical decision-making model in a Microsoft Word document in addition to viewing the model on the D2L screen. Future research may ensure that graphs appear at an optimal resolution on different browsers prior to having participants complete the modules.

Despite these limitations, the results of the study have important implications for research and practice. This study demonstrated the effectiveness of an efficient, low-cost tool in assisting in accurate analysis of visual data of front-line employees. Also, the current study is the first, to our knowledge, to empirically evaluate the effects of a clinical decision-making model on clinical decisions related to behavior-analytic practice. A clinical decision-making model may become a tool used in clinical settings for beginning practitioners to accurately and rapidly make treatment related decisions. This may reduce the amount of ineffective treatments that are in place, and also prevent the overtraining of specific targets (i.e., thereby ensuring therapy is focused on skill deficits), capitalizing on effective programming for children who may benefit from behavioral therapy, such as children diagnosed with autism spectrum disorder. The current study also contributes to previous research that successfully utilizes computer-based instruction to teach front-line employees to conduct visual analysis (Fisher et al., 2003; Wolfe & Slocum, 2015; Young & Daly, 2016).

APPENDIX

Table 1

Participant demographic and background information.

	Gender	Degree	Major	Experience	
				BA	VA
Molly	Female	Bachelor's	Communications	No	Yes
Zane	Male	Bachelor's	Neuroscience	No	Yes
Ally	Female	Bachelor's	Psychology, Human Biology	No	Yes
Riley	Female	Bachelor's	Physiology	No	Yes
Bre	Female	Bachelor's	HDFS	No	Yes
Jessica	Female	Bachelor's	Special Education	No	No
Nate	Male	Bachelor's	Psychology	No	Yes
Kayla	Female	Bachelor's	Psychology	10 months	No

Note. BA = Behavior analysis; VA = Visual analysis; HDFS = Human Development and Family Studies. Kayla reported taking Introduction to Applied Behavior Analysis as a course. However, her description of the class did not describe any specific training in clinical decision-making or the visual analysis of graphs. Ally, Bre, Nate, Molly, and Riley reported having experience in graphs. However, their explanations did not describe specific interpretation of line graphs, except for Bre, who plotted data points on line graphs, but not interpret the data. Molly and Zane reported having experience in interpreting audio wave graphs and bar graphs, respectively.

Table 2

Parameters for the autoregressive formula for each clinical decision outcome

Parameters	Intervention Complete	Continue Intervention	Discontinue Intervention	Modify Intervention
y-intercept	20 – 30	20 – 36	20 – 30	40 - 60
Slope	7	6	0	0
Auto-c	0	0	0	-0.8
Variability	(-3 – 3), (-10 – 10)	(-3 – 3), (-10 – 10)	(-2 – 2), (-4 – 4)	(-10 – 10)

Note. Auto-c = autocorrelation. The parenthesis in the variability row indicate the degree of the variability of the graphs (e.g., the range (-3 – 3) reflected data with minimal variability, and the range (-10 – 10) reflected a larger degree of variability in the data).

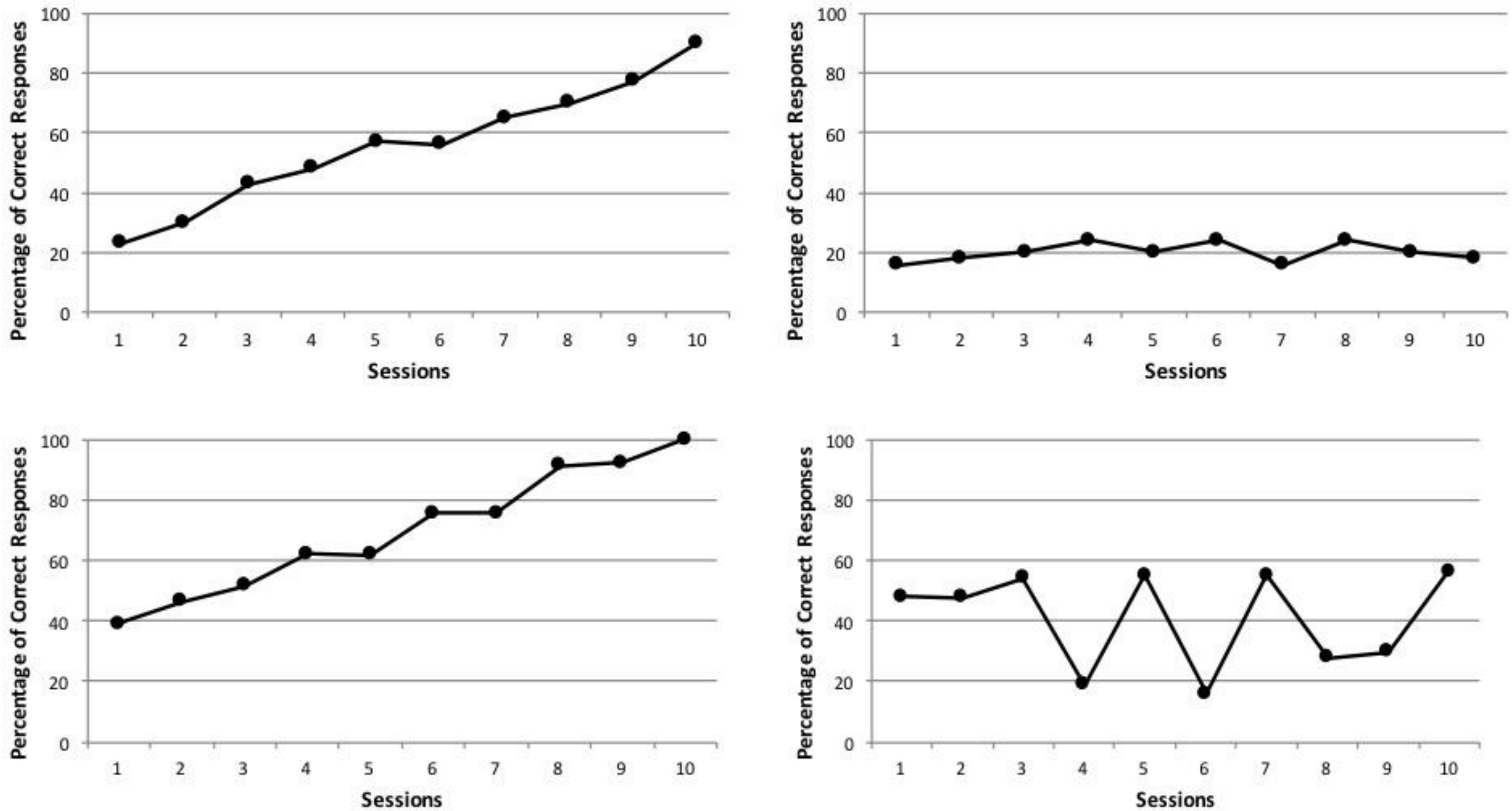

Figure 1. Examples of the four graph categories: top left, continue; top right, discontinue; bottom left, intervention is complete; bottom right, modify.

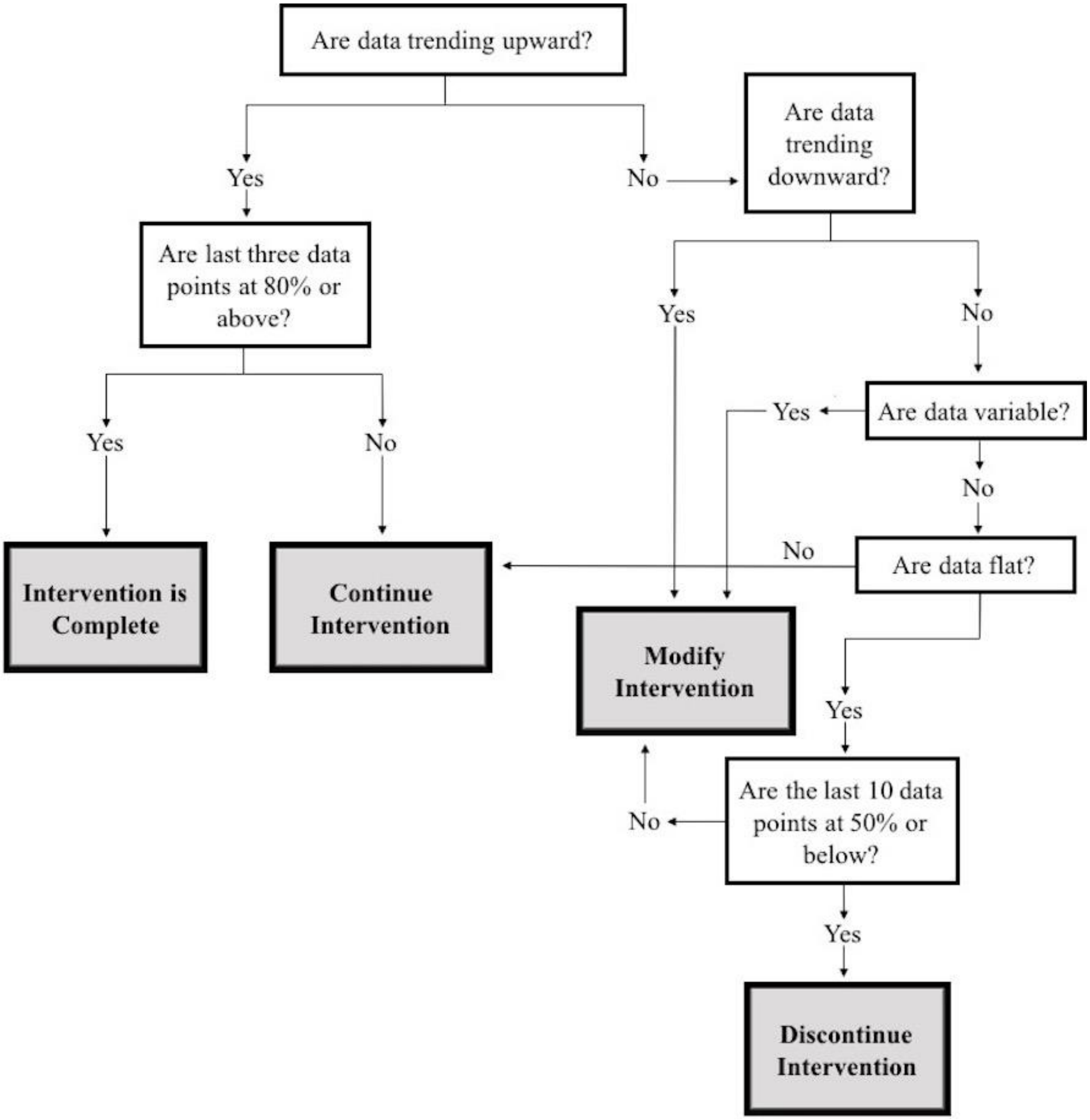

Figure 2. Clinical decision-making model.

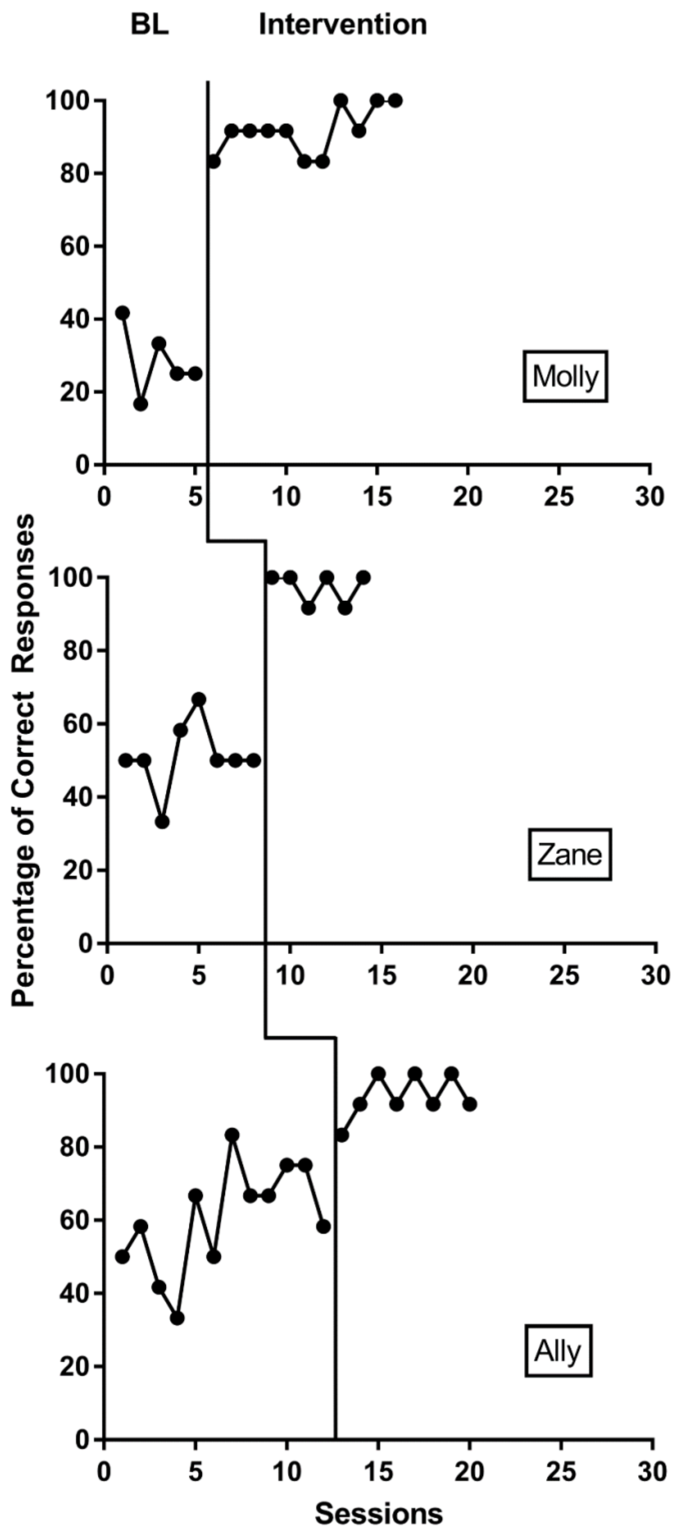

Figure 3. Results for Molly, Zane, and Ally.

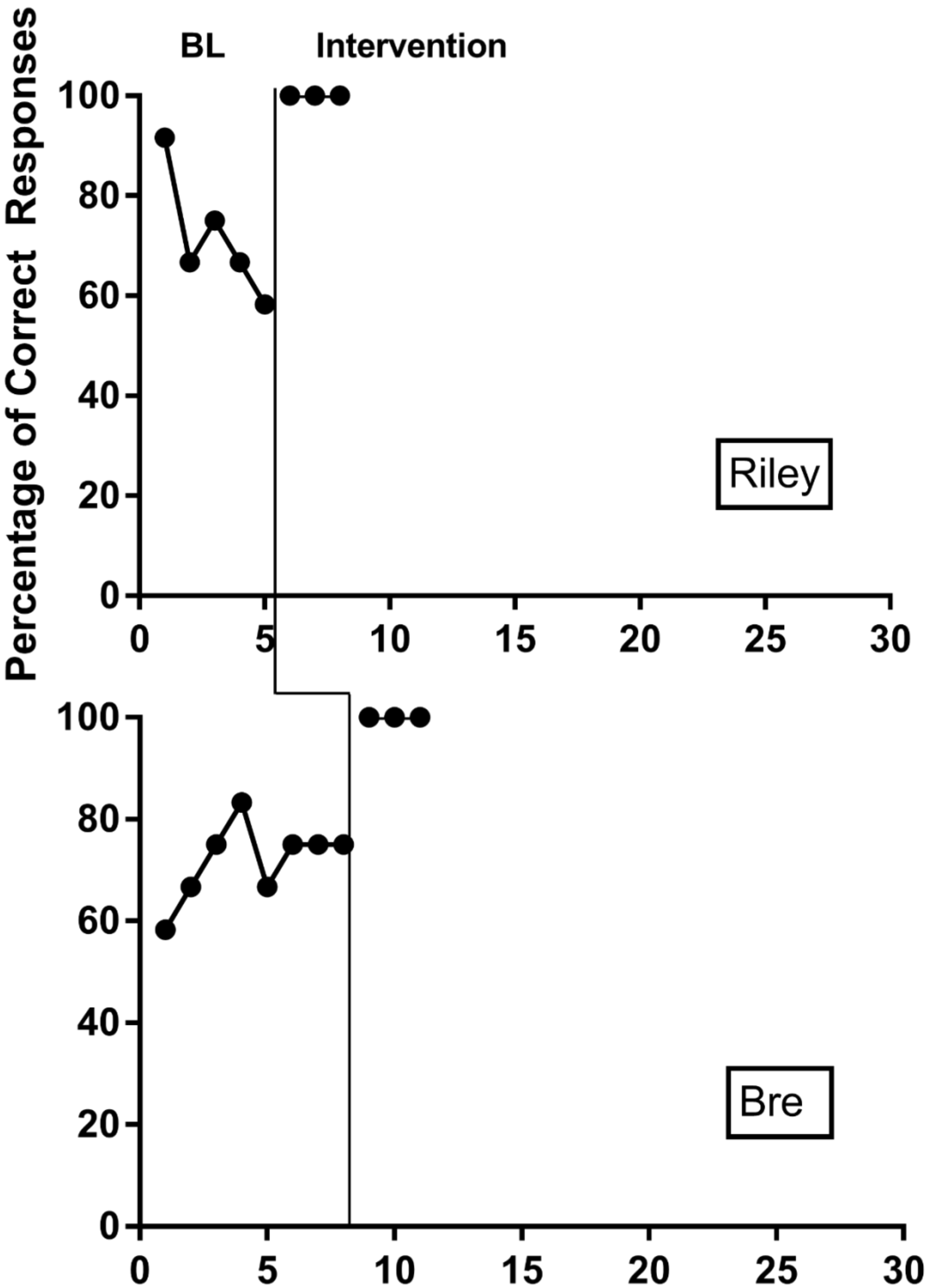

Figure 4. Results for Riley and Bre.

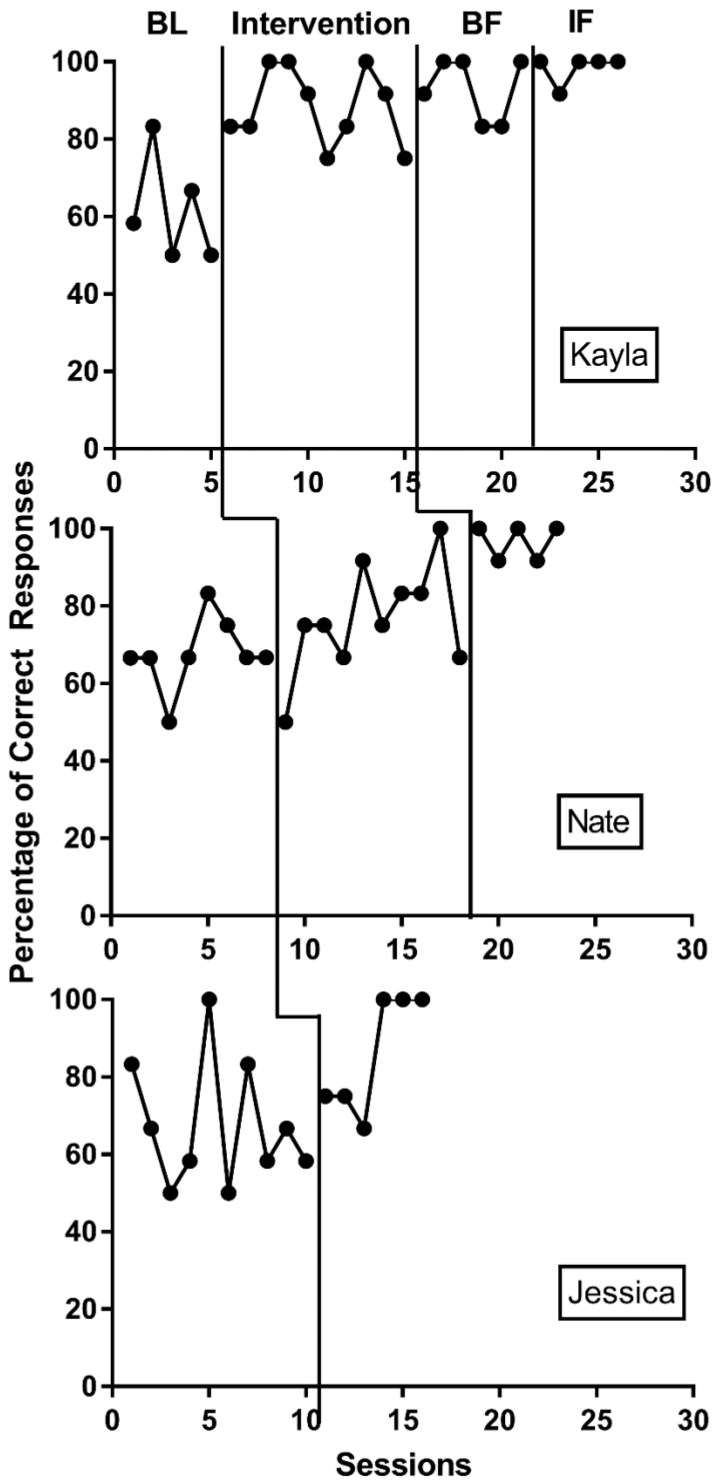

Figure 5. Results for Kayla, Nate, and Jessica.

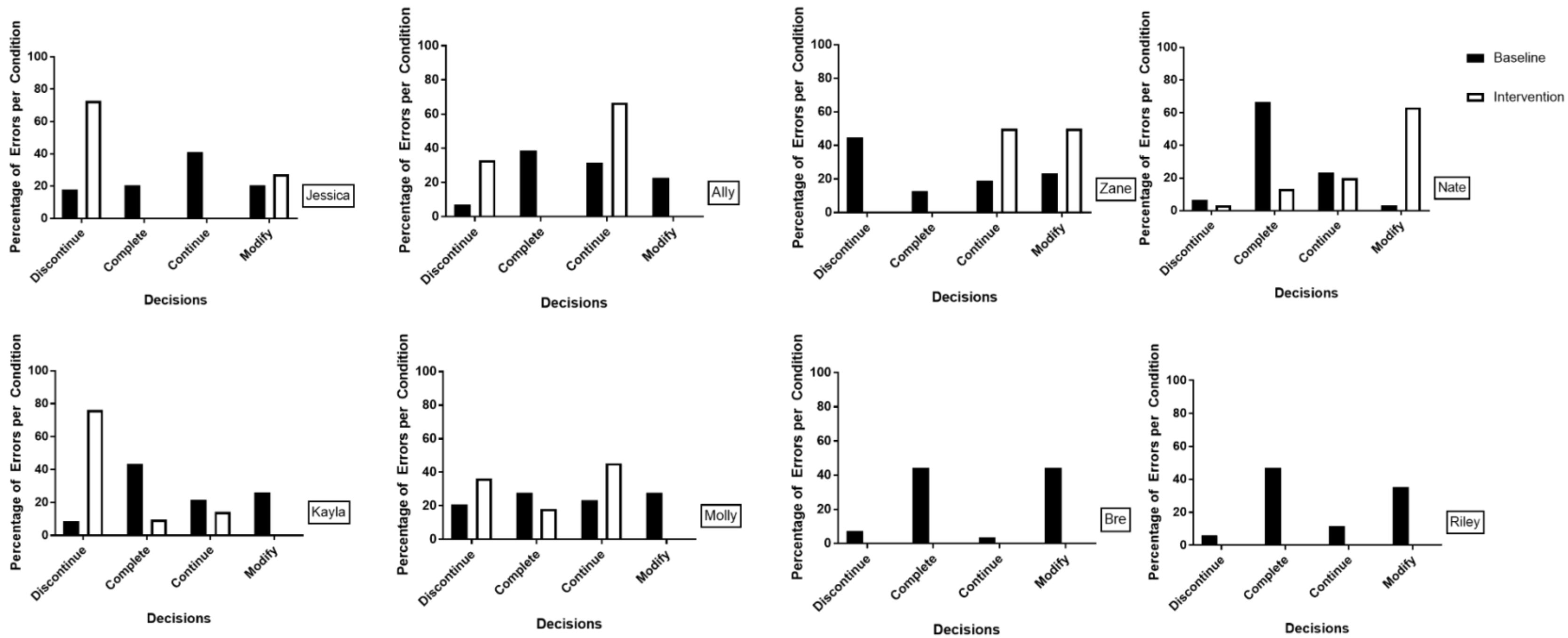

Figure 6. Results for the error analysis for individual participants.

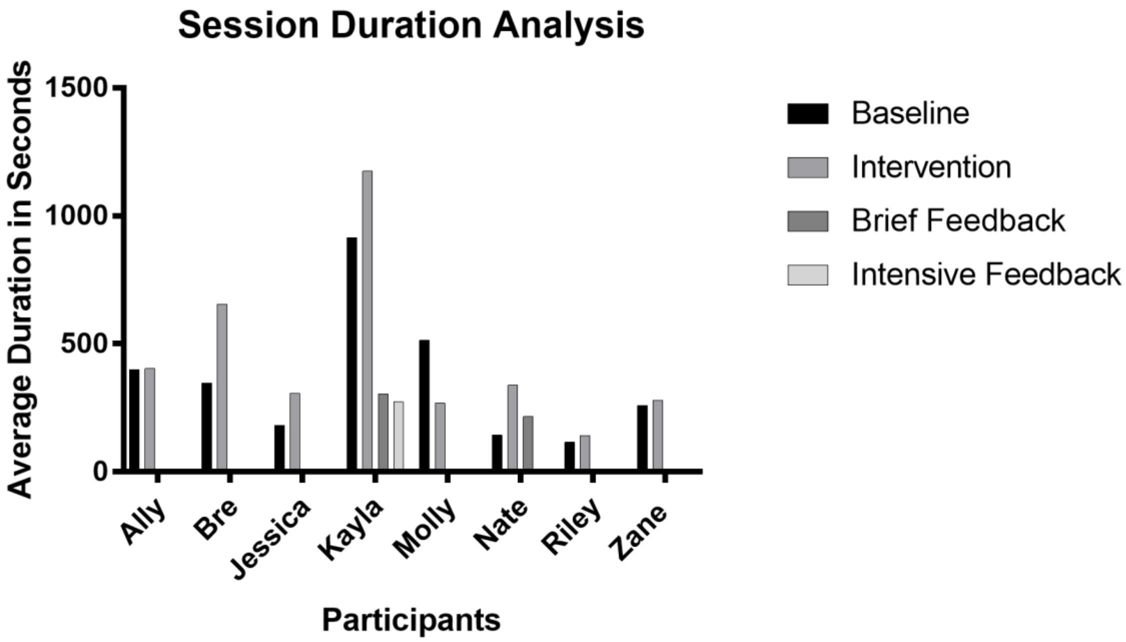

Figure 7. Results for the average session duration per condition for individual participants.

REFERENCES

REFERENCES

Baer, D. M., Wolf, M. M, & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1,* 91-97. doi: 10.1901/jaba.1968.1-91

Behavior Analyst Certification Board (2016). *Registered Behavior Technician™ (RBT®) task list.* Retrieved from http://bacb.com/rbt-task-list/ on 4/17/2017.

Brodhead, M. T. (2015). Maintaining professional relationships in an interdisciplinary setting: Strategies for navigating nonbehavioral treatment recommendations for individuals with autism. *Behavior Analysis in Practice, 8, 1,* 70-78. doi: 10.1007/s40617-015-0042-7

Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied Behavior Analysis.* Columbus, OH: Merrill Publishing Company.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson/Merrill-Prentice Hall.

Downs, A. Downs, R. C., & Rau, K. (2008). Effects of training and feedback on discrete trial teaching skills and student performance. *Research in Developmental Disabilities, 29,* 235-246. doi: 10.1016/j.ridd.2007.05.001

Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis, 36,* 387-406. doi:10.1901/jaba.2003.36-387

Fiske, K., & Delmolino, L. (2012). Use of discontinuous methods of data collection in behavioral intervention: guidelines for practitioners. *Behavior Analysis in Practice, 5*, 77-81. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3592492/

Gast, D. L., & Ledford, J. R. (2014). Visual analysis of graphic data. In D.L. Gast & J.R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed.) (pp. 176-210). New York, NY: Routledge.

Geiger, K. A., Carr, J. E., & LeBlanc, L. A. (2010). Function-based treatments for escape-maintained problem behavior: a treatment selection model for practicing behavior analysts. *Behavior Analysis in Practice, 3*,1, 22-32. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3004681/pdf/i1998-1929-3-1-22.pdf

Grow, L. L., Carr, J. E., & LeBlanc, L. A. (2009). Treatments for attention-maintain problem behavior: empirical support and clinical recommendations. *Journal of Evidence-Based Practices for Schools, 10,* 70-92. Retrieved from https://www.researchgate.net/publication/255949566_Treatments_for_attention-maintained_problem_behavior_Empirical_support_and_clinical_recommendations

Higbee, T. S., Aporta, A. P., Resenda, M. N, Goyos, C., & Pollard, J. S. (2016). Interactive computer training to teach discrete-trial instruction to undergraduates and special educators in Brazil: A replication and extension. *Journal of Applied Behavior Analysis, 49,* 780-793. doi:10.1002/jaba.329

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*, 1, 26-38. doi: 10.1177/0741932512452794

Leaf, J. B., Leaf, R., McEachin, J., Taubman, M., Smith, T., Harris, S. L. … & Waks, A. (2016). Concerns about the Registered Behavior Technician™ in relation to effective autism intervention. *Behavior Analysis in Practice,10*, 2, 157-163. doi:10.1007/s40617-016-0145-9

LeBlanc, L. A., Gravina, N., & Carr, J. E. (2009). Training issues unique to autism spectrum disorder. J. L. Maston (Ed.). *Applied behavior analysis for children with autism spectrum disorders* (pp. 225-235). Auburn, AL: Springer Science +Business Media, LCC. doi: 10.1007/978-1-4419-0088-3_13

LeBlanc, L. A., Raetz, P. B., Sellers, T. P., & Carr, J. E. (2016). A proposed model for selecting measurement procedures for the assessment and treatment of problem behavior. *Behavior Analysis in Practice, 9,* 77-83. doi: 10.1007/s40617-015-0063-2

Maffei-Almodovar, L., Feliciano, G., Fienup, D. M., & Sturmey, P. (2017). The use of behavioral skills training to teach graph analysis to community-based teachers. *Behavior Analysis Practice*, 1-8*.* doi:10.1007/s40617-017-0199-3

Nelson, P. M., Van Norman, E. R., & Christ, T. J. (2016). Visual analysis among novices: Training and trend lines as graphic aids. *Contemporary School Psychology, 21,* 92-102. doi: 10.1007/s40688-016-0107-9

Pollard, J. S., Higbee, T. S., Akers, J. S., & Brodhead, M. T. (2014). An evaluation of interactive computer training to teach instructors to implement discrete trials with children with autism. *Journal of Applied Behavior Analysis, 47,* 765-776. doi:10.1002/jaba.152

Sellers, T. P., Alai-Rosales, S., MacDonald, R. P. F. (2016). Taking full responsibility: The ethics of supervision in behavior analytic practice. *Behavior Analysis in Practice, 9,* 299-308. doi: 10.1007/s4617-016-0144-x

Sidman, M. (1960). *Tactics of scientific-research: Evaluating experimental data in psychology.* New York, NY: Basic Books.

Stewart, K. K., Carr, J. E., Brandt, C. W., & McHenry, M. M. (2007). An evaluation of the conservation dual-criterion method for teaching university students to visually inspect AB-design graphs. *Journal of Applied Behavior Analysis, 40,* 713-718. doi: 10.1901/jaba.2007.713-718

Young, N. D., & Daly, E. J. (2016). An evaluation of prompting and reinforcement for training visual analysis skills. *Journal of Behavioral Education, 25,* 95-119. doi: 10.1007/s10864-015-9234-z

Wishnowski, L. A., Yu, C. T., Pear, J., Chand, C., & Saltel, L. (2017). Effects of computer-aided instruction on the implementation of the MSWO stimulus preference assessment. *Behavioral Interventions,* 1-13. doi:10.1002/bin.1508

Wolfe, K., & Slocum, T. A. (2015). A comparison of two approaches to training visual analysis of AB graphs. *Journal of Applied Behavior Analysis, 48,* 472-477. doi: 10.1002/jaba.212