6-2019

Training Functional Analysis Skills with Instruction with Video Modeling and Video-Self Monitoring

Haley Ciara Hughes

TRAINING FUNCTIONAL ANALYSIS SKILLS WITH INSTRUCTION WITH VIDEO MODELING AND VIDEO-SELF MONITORING

by

Haley Ciara Hughes

A thesis submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Master of Arts
Psychology
Western Michigan University
June 2019

Thesis Committee:

R. Wayne Fuqua, Ph.D., BCBA-D, Chair
Jonathan Baker, Ph.D., BCBA-D
Jessica Frieder, Ph.D., BCBA-D

ACKNOWLEDGMENTS

I would like to thank my academic advisor, Dr. Wayne Fuqua, for his guidance and support in my graduate career. I also wish to show my appreciation for Dr. Jonathan Baker and Dr. Jessica Frieder, the other members of my thesis committee, for their ongoing support in the development and implementation of my research. I have a great degree of gratitude for all of my academic peers, supervisors, and friends for their feedback as I organized this project, trouble-shot issues during data collection, and prepared my oral defense.

I also wish to thank my family, in particular, my father, step-mother, grandparents, and mother for their continued encouragement in my academic endeavors. I would not be half the professional that I am today without their guidance in raising me to have the expectations, perseverance, and goals that I have to this day.

Haley Ciara Hughes

TRAINING FUNCTIONAL ANALYSIS SKILLS WITH INSTRUCTION WITH VIDEO MODELING AND VIDEO SELF-MONITORING

Haley Ciara Hughes, M.A.

Western Michigan University, 2019

When developing plans to reduce a challenging behavior, Board Certified Behavior Analysts (BCBAs) have an ethical obligation to first conduct a functional assessment (PECC, 2014, 3.01a), the goal of which is to identify the controlling variables for such behaviors and to use that information in training more appropriate replacement behaviors (PECC, 2014, 4.08b). It is important to train aspiring behavior analysts to implement an experimental functional analysis (FA) as it yields more accurate results than other types of functional assessment (Iwata & Dozier, 2008). Despite being considered a gold standard for training a variety of skills, behavioral skills training (BST) is often very time intensive on the part of the trainer (Iwata et al., 2000). Video self-monitoring (VSM) may be an effective alternative to train students, as it has been shown to be effective in the training of practitioners on other ABA practitioner skills (Belfiore, Fritts, & Herman, 2008; Field, Frieder, Mcgee, Peterson, & Duinkerken, 2015). Using a multiple baseline across subjects design, this study evaluates the efficacy of several alternative training strategies for training FA skills for undergraduate students, including interventions that featured instruction plus video modeling (IVM), and VSMN, with and without performance feedback. Results revealed that IVM produced a notable improvement in performance for all four participants, but that this intervention alone was not sufficient. All participants showed further performance improvement with the addition of VSMN and VSMN plus feedback.

TABLE OF CONTENTS

Table of Contents – Continued

Table of Contents – Continued

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Functional Behavior Assessment and Functional Analysis

The Professional and Ethical Compliance Code (PECC) of the Behavior Analysis Certification Board (BACB) emphasizes behavior analytic assessment and the implementation of treatments that are informed by functional behavior assessment (FBA) (BACB, 2016b). When behavior analysts develop a behavior-reduction program, they have an ethical obligation to first conduct a functional assessment (PECC, 2014, 3.01a) and use those results to design a function-informed intervention plan. FBA typically refers to an umbrella term that comprises indirect assessment, direct assessment, and experimental functional analysis (Cooper, Heron, & Heward, 2007). Iwata and Dozier (2008) state that despite the popularity of indirect and direct assessment, experimental functional analysis (hereafter referred to as "FA") is the most accurate way to identify a causal relationship between environmental variables and problem behaviors. As a result, FAs are considered to be a crucial component in training programs for applied behavior analysis (ABA) practitioners.

The FA consists of experimental manipulations of "test conditions" selected as plausible causal variables for a problem behavior (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994). The researchers can then demonstrate a "cause" of the problem behavior by observing changes in challenging behavior as those contextual variables (e.g., reinforcement and other relevant variables) are systematically manipulated and replicated across multiple sessions (Iwata & Dozier, 2008). There is wide recognition that ABA practitioners (BCBAs, special educators, and other professionals who implement ABA interventions) need to develop competence in the design, implementation and interpretation of FAs skills. Thus, there is a pressing need to develop and evaluate training strategies that are both effective and efficient at developing and maintaining competence in FA skills.

Importance of Training Practitioners

Training competent ABA practitioners is important both to the clients that they serve and to the field of behavior analysis. In their seminal article, Baer, Wolf, and Risley (1968) identify the defining features of applied behavior analysis, features that include the provision of services that are effective and technological (e.g., described in sufficient detail to allow replication). Moreover, in addition to a client's right to an effective treatment, Van Houten et al. (1988) also outline the right to treatment by a competent behavior analyst.

Over recent years, many individuals have sought and become credentialed as BCBAs (Deochand and Fuqua, 2016). However, as of 2017, nearly 35% of aspiring practitioners are still having difficulty in passing the credentialing exam (BACB, 2016a), which may indicate an issue with the quality of academic and experiential preparation to become a BCBA. Passing rates (the number of people from each institution who take and pass the BCBA exam) vary across institutions also, ranging from 25% to 100%, suggesting that university-based training programs for behavior analysis may vary significantly in their efficacy in preparing students to pass this essential component of the certification process. Even those students who pass the BCBA certification exam may still need supplemental training to apply crucial ABA practitioner skills. Of course extensive supervised experience is also a crucial component of an effective training program and also an essential component of the credentialing process. However, there is an increasing large number of evidence-based ABA practitioner skills to be trained and practicum training experiences may be limited by the range of clients and nature of behavioral goals at any given practicum placement. As a result, it seems only prudent to supplement practicum training of ABA practitioner skills with more structured training (e.g., the use of simulations) to insure that aspiring BCBAs acquire essential skills, such as how to design and conduct a FA, at a level of mastery and fluency that is needed for high quality practice.

Additionally, the credentialing process does not have a data-based system to evaluate the competence of BCBAs as they engage in applied or clinical practice. In recent years, the expectations for BCBA supervisors to provide structured training of practitioner skills has been clearly articulated (e.g., Sellers, Valentino, & LeBlanc, 2016). This is a valuable step in assuring competent practitioners but this process is still new and has not yet been subject to empirical evaluation as to the extent of ABA skills that are covered in experiential training and the mastery level that is obtained by BCBAs in training through this process. Thus, it is important to identify a training method that is efficient and effective because training in an academic setting only occurs for a limited amount of time. Even those individuals who pass the BACB exam and complete practicum courses may need supplemental training to apply more advanced practitioner skills, such as functional analysis, in real world settings or to expand their scope of competence to new areas of practice.

Thus, it is important to identify a training method that is efficient and effective in training essential practitioner skills because academic training occurs for a limited amount of time. While most universities have practicum experience as a requirement, it is still possible that students will not fully develop competence in clinical practice skills, either because of a lack of efficacy of practicum instruction or the range of skills taught is limited to the range of clients to which the student is exposed. As behavior analysts, it is part of the ethical code of conduct to train behavior analytic students competently, which would require an evidence-based method of instruction and training (Bailey & Burch, 2016, PECC, 2014, 5.04). Fortunately a number of training strategies have been developed and evaluated, these will be selectively reviewed in the next section.

Methods of Training

Behavioral Skills Training

Behavioral Skills Training (BST) is a multicomponent package consisting of instruction, modeling, rehearsal, and feedback (Ward-Horner & Sturmey, 2012). It is empirically supported and highly recommended as an effective training method for a wide variety of skills. Parsons, Rollyson, and Reid (2012) have described the essential components of BST. Instruction may consist of written and vocal explanation of the components of a skill using operational definitions. Next, modeling may consist of the trainers demonstrating the skill for the participants in a "role-playing" scenario either in person or via video. This should require that the instructors be well versed in the script of the modeling scenarios. Then, the participants would be required to practice the skill by role-playing a number of scenarios. Finally, the participants receive specific feedback regarding their performance (e.g., what they completed correctly and incorrectly). The components of rehearsal and feedback are then repeated until a mastery criterion is achieved (Parsons, Rollyson, & Reid, 2012; Reid, 2012).

Iwata et al. (2000) evaluated the effects of a BST program on the implementation of functional analysis test conditions for undergraduate students. Researchers played the role of clients while following scripts during the five-minute, simulated FA conditions. Prior to the BST intervention, the key components of the each condition and a videotaped simulation of the conditions were reviewed with participants, and then participants completed a 20-item quiz to ensure master of the academic content. In this study, there was no separate evaluation of the academic instruction and video simulation. Following this "classroom-like" training, participant

FA performance was probed again while participants had access to the outline of each of the FA conditions, with the addition of experimenter feedback on their performance in-vivo. Moreover, if participants scored below a set mastery criteria, the video of the session was replayed, and the experimenter highlighted correct and incorrect features of the participant's performance in the video. All participants improved their performance in the FA skills following BST.

While BST is a common strategy for staff training (Graudins, Rehfeldt, DeMattei, Baker, & Scaglia, 2011; Hogan, Knez, & Kahng, 2014; Barnes et al., 2014; Iwata et al., 2000; Fetherston & Sturmey, 2014), there are some limitations to its use. Trainer and trainee time allocation can be quite extensive especially if additional booster sessions or remedial training is necessary (Barnes et al., 2014). For example, the instruction, quiz, and video model took approximately 1.5 hours total for a group of participants, with 10-30 minutes of feedback on performance in Iwata et al. (2000)'s study. This may be impractical in some university settings that feature large enrollment classes. where instructor time cannot be allocated to all students to give adequate, individualized feedback. Further compounding this issue, there are 87 online programs with a verified course sequence as of 2019 (BACB, n.d.). If BST is the only effective way to teach, then this poses a barrier for online programs in educating students as the online format may not be conducive for direct observation and individualized feedback on practitioner skills. As a result, alternative methods to BST should be pursued that have more ecological validity in the standard classroom environment.

Interteaching

Interteaching is a multi-component package that requires students to complete preparation materials prior to class, break up into small student/peer groups to discuss the material. The instructor then facilitates facilitating discussion between students and their peer partners. Students are then given an opportunity to practice skills. Finally, the instructor provides a brief lecture to go over any difficulties the students encountered. Students scored higher on quiz scores with interteaching without points and with points than lecture alone in a study of the effects of interteaching on examination scores (Filipiak, Rehfeldt, Heal, & Baker, 2010). However, student mean examination scores never rose above 85% in any condition, indicating that, on average, students were still not fluent in the material. Moreover, while interteaching did not require as much time on the instructor, student assessments contained only written examinations, which assess verbal repertoires. Without a systematic evaluation of student performance skills that map onto the written questions, it is not possible to state whether the training had any impact on student performance repertoires.

Instruction

In a typical university environment, classroom instruction alone is commonly used as a method of training. In a typical university environment, classroom instruction is commonly emphasized as a method of training. However, in a component analysis of behavioral skills training written instruction was not found to be an effective skills training strategy when implemented alone (Ward-Horner & Sturmey, 2012). So, as a method of training in a university environment, it may be cost-effective and common, but the efficacy of written instruction alone as a methodology of training is questionable and merits supplementation by additional training strategies.

Video Modeling and Video Self Modeling

A variety of video modeling interventions have been developed and evaluated for training practical skills for students. Among the research literature, video modeling interventions typically require that a participant observe a video of an individual (either themselves or a peer/expert) performing accurately, without any type of response requirement while observing the video. Two primary types of video modeling include video modeling (VM), using another individual as a model (e.g., peer or expert), and "video self modeling" (VSM), which involves using self as a model.

With VM, participants are asked to observe a video of a competent individual engaging in the targeted behavior. Following observation of the video, they are asked to perform the required skill again. VM has been shown to be an effective intervention at teaching skills to staff in some studies (Catania, Almeida, Liu-Constant, & Reed, 2009). However, some researchers report that feedback is warranted to improve performance to mastery some of the time (Ward-Horner & Sturmey, 2012).

Similarly, VSM entails a student observing videos of himself or herself engaging in the targeted behavior accurately (Cihak & Schrader, 2008). Cihak and Schrader (2008) found that some participants derived greater benefit from being their own video model when compared with observing another individual as the model. They hypothesized that the relative benefits of VSM, in comparison to VM, may have been a result of the shared similarities between participant and model being identical in VSM. However, the majority of research on VSM has been conducted with training skills to individuals with developmental disabilities (Gelbar, Anderson, McCarthy, & Buggey, 2012), and to the author's knowledge, no studies have systematically evaluated VSM to train FA skills. Although VM and VSM are considered evidence-based practices among teaching individuals with autism (Bellini & Akullian, 2007), their status as an evidence-based strategy for staff training has not been systematically evaluated. Additionally, as some research has suggested that supplemental feedback is necessary when training staff on complex skills, researchers continue to recommend the use of BST to train practitioners particular ABA skills (Ward-Horner & Sturmey, 2012; Parsons et al., 2012).

Video Self-Monitoring

Self-monitoring requires that an individual monitor his or her own behavior after being trained on the targeted behavior (Allinder, Bolling, Oats, & Gagnon, 2000). In one variation of self-monitoring, researchers can adjust the methodology of self-monitoring slightly by videotaping participant performance, and having students rate their behavior. This is a process titled video self-monitoring (VSMN). VSMN (sometimes referred to as video feedback or video self-evaluation) is a form of self-monitoring where the student performs a skill, watches a video of himself or herself performing the skill, collects data on his or her behavior, and evaluates his or her performance (Belfiore, Fritts, & Herman, 2008). This is different than video self-modeling as in video self-modeling, the student is typically shown a doctored video of himself or herself performing a skill typically to fluency (i.e., with no errors). By contrast, VSMN would include two additional components: correct and incorrect performance and monitoring his or her own performance on a task analysis.

Although there is not a standardized manner to conduct VSMN, VSMN does have some defining features. Typically, the researcher trains the participant how to use the checklist of necessary components for the target skill prior to having the participant score his or her behavior independently (Downs, Miltenberger, Biedronski, & Witherspoon, 2015). Currently, there is not a standard method for how much training to provide participants, or what the training should look like. Furthermore, there is not standardization for what level of accuracy that participants must obtain in scoring their own video demonstrations and how to train that level of scoring accuracy. For example, the training may include having the participant practice self-scoring with the sheet alongside the researchers until data collected by the participant matches researcher data to some pre-determined degree of accuracy (e.g., 90%). However, some other participants may need less instruction or no additional practice to become accurate self-recorders. Following any preliminary training regarding the scoring sheet, whether that be practice self-recording or instruction only, the researcher then shows the participant a video of himself or herself performing a skill. The participant is asked to self-record with the checklist for a specified amount of time while watching the video. Then, the researchers can either provide feedback (Belfiore et al., 2008) on their scoring accuracy or no feedback on their accuracy (Downs et al., 2015). In the study conducted by Bishop, Snyder, and Crow (2015), feedback was seen as superior to no feedback. Ultimately, prior research shows that there is some variety to how VSMN approaches have been implemented across studies, which may cause some issues with replication.

Although the specific methodology in the studies may vary, VSMN has been shown to be applicable to a range of skills, such as increasing the number of social initiations of children with autism (Deitcham, Reeve, Reeve, & Progar, 2010), improving the accuracy of yoga poses of participants, (Downs et al., 2015), and teaching skills to staff (Belfiore et al., 2008). By observing behavior after it has occurred, participants are able to more closely monitor their behavior. Researchers have also implemented VSMN to teach staff to perform discrete trial teaching with children with autism (Belfiore et al., 2008), implement embedded instructional learning trials (Bishop, Snyder, & Crow, 2015), and increase praise frequency of teachers at Head Start (Wright, Ellis, & Baxter, 2012).

Belfiore, Fritts, and Herman (2008) implemented VSMN to train providers to implement discrete trial instruction following training of the components of the self-monitoring checklist.

All three participants reached mastery within eight sessions or less, with performance dropping significantly in maintenance for two of the participants. Furthermore, Bishop, Snyder, and Crow (2015) evaluated four conditions: 1.) VSMN without prior training on essential components of the target task, 2.) VSMN without ongoing feedback on form following a lengthy training on the form and the components of learning trials, 3.) VSMN with ongoing feedback on scoring accuracy, and 4.) maintenance without feedback. All three participants showed more improvement in performance with ongoing feedback on coding (third condition), but performance was slightly variable and did not maintain for two of the participants. Finally, Wright, Ellis, and Baxter (2012) analyzed the effects of VSMN of praise statements on Head Start teacher praise frequency both immediately after the video recording and the next day. They discovered that the frequency of praise for both immediate and delayed video self evaluation was higher than the control group.

VSMN has been shown to be an effective intervention, with most staff reaching mastery criterion within a short period of time (Belfiore et al., 2008). Belfiore et al. (2008) went on to note that the VSMN procedure required less time on the part of the researchers than other strategies, such as BST. However, in some cases, in-depth training on the self-monitoring form may take a longer period of time, as demonstrated by the 2-hour training on coding in the study conducted by Bishop et al. (2015). As such, it may be interesting to determine if this amount of pre-training to self-record would be necessary to video self-monitor. Bishop et al. (2015) also incorporated feedback sessions on scoring accuracy following each session during one of their phases, which increased the session time. During some of the experiments, issues arose with maintenance of skills (Belfiore et al., 2008; Bishop et al., 2015). As the intervention is not labor intensive, Belfiore et al. (2008) suggested frequent VSMN two to three times a week, which may keep performance within mastery criterion.

One published study has evaluated VSMN to teach graduate students to conduct an experimental FA, but as one component of a multi-component package. Field, Frieder, Mcgee, Peterson, and Duinkerken (2015) first trained participants using a small, 30-minute instruction on the written protocol of each FA condition based off of Iwata et al. (1982/1994) consisting of the experimenter reading off each of the components of the protocol while participants followed along on their own written copies. If performance did not improve to mastery, participants observed a peer video model for all four FA conditions, and scored the performance of the peer in just the targeted FA condition, which they called video observation and rating. For the remainder of this discussion, video observation and rating will be titled video peer monitoring (VPMN) for the purposes of clarity. Aside from the original instruction, Field et al. (2015) provided no supplemental training on how to record using the sheet and no feedback regarding accuracy in scoring, although they did collect data on participant scoring accuracy. VPMN differed from VSMN in that the condition required participants to view a video of a peer and monitor the peer's performance depicting only examples of correct performance, whereas VSMN requires a participant to observe his or her own performance depicting examples and nonexamples. Both VSMN and VPMN require the participant to view and score performance. If performance still fell short of mastery level, Field et al. (2015) implemented VSMN and 71.42% of those individuals reached mastery criteria as a result. This is interpreted as evidence to support that VSMN may be effective in training this skill, which Field et al. (2015) hypothesized may have been a result of the use of non-examples in the VSMN portion of the training.

Table 1 below highlights some of the standard differences between video self-monitoring, video peer-monitoring, video self-modeling, and video modeling that are generally found in common research methodology.

Table 1.

Overview of Standard Video Training Methods

Typical Components	Training Method			
	Video modeling (VM)	***Video self-modeling (VSM)***	***Video self-monitoring (VSMN)***	***Video peer-monitoring (VPMN)***
Who is in the video?	Peer or expert	Self	Self	Peer or expert
Types of performance	Correct	Correct	Correct and incorrect	Correct
Scoring of behavior(s) in video?	No	No	Yes	Yes

It should be noted that Table 1 provides a standard overview of each of the primary video methods. Moreover, scoring of correct and incorrect performance during VSMN is not a defining feature of VSMN, but it has seemed to characterize this research area. The way that VPMN is described in Table 1 refers specifically to the procedures described in Field et al. (2015). As such, there are variations of all of the above methodologies from these standard formats.

However, the study conducted by Field et al. (2015) had several limitations that should be addressed in future literature. To start, they conducted one baseline probe, so it is not possible to rule out the possibility that continued exposure to repeated baselines assessments could have produced increases in performance, even without an intervention. Moreover, the 30-minute instruction may not have been effective for most participants as it covered only the task analyses. As noted previously, instruction alone is typically not sufficient to achieve mastery performance (Ward-Horner & Sturmey, 2012). It is possible that with a more similar structure to the instruction provided in Iwata et al. (2000), greater performance gains may have been achieved.

Additionally, each script in Field et al. (2015)'s study included only five target behaviors per five minute condition. As such, it is possible that the conditions may not have been as difficult as those probed by others. For example, Iwata et al. (2000) integrated fifteen target behaviors, some inappropriate behaviors, and appropriate behaviors. There is potential that the ease of the scripts may have led to ceiling effects in participants following their integration of video monitoring as a result (Field et al., 2015).

Furthermore, Field et al. (2015) did not report which specific steps were missed in each FA condition making it is impossible to pinpoint skills that were especially resistant to training and thus a potential target for focused training. Moreover, in the VPMN condition, participants rated the behavior of a peer video model who engaged in perfect performance. As such, the VPMN condition did not include examples and non-examples of performance. It was not until the VSMN condition was integrated into Field et al. (2015)'s study that non-examples were included in the video, which may have been a reason for the lack of success for some participants. This begs the question of whether a more effective alternative to VPMN would be to expose participants to VSMN right away if their performance is not adequate.

Overall, VSMN requires that a participant observe a video of himself or herself engaging in a targeted skill, then score his or her performance using a checklist. Benefits of VSMN include that it is effective in teaching some practitioner skills and it is generally less demanding on trainer time than BST, one version of cost-efficiency. Some of the drawbacks to using VSMN include that it is not as well-researched as BST as a method of practitioner instruction, and the methodology in some of the studies is not always consistent (i.e., there is no training manual). Finally, as VSMN possesses some overlap with the practice and feedback components of BST, it merits evaluation as an independent training procedure that would allow students to provide individualized self-feedback to supplement instructor feedback.

In summation, some studies have found BST to be effective in teaching practitioners to implement functional analysis methodology (Chok, Shlesinger, Studer, & Bird, 2012; Iwata et al., 2000). However, due to the intensive nature of some BST training programs, it may not always be feasible to implement BST training in a classroom environment or in other settings where resources and time are scarce. As noted previously, VSMN provides an alternative to BST training and has also been shown to be effective at teaching staff skills (Belfiore et al., 2008; Bishop et al., 2015; Wright et al., 2012). However, only one study to date has attempted to assess the applicability of VSMN to teach college students to implement functional analysis methodology (Field et al., 2015). The limitations in Field et al. (2015) include that there were not multiple true baseline measures, the instruction was limited to 30-minutes, the independent variable was not applied across all of the functional analysis conditions it could have been applied to, and the lack of script difficulty. Thus there is a need to replicate the research on VSMN and address some of the methodological limitations of prior research.

Moreover, as VSMN showed promising results likely as a result of the inclusion of incorrect performance, VSMN should be studied without the VPMN condition occurring beforehand to evaluate participant performance. Therefore, the present study is a replication and extension of the Field et al. (2015) study with the purpose of analyzing the results of an instruction and video model (IVM) treatment package and VSMN on the implementation of the standard functional analysis skills among undergraduate novice students at Western Michigan University. Moreover, this study is a systematic replication and extension of the Field et al. (2015) study.

METHODS

Setting and Participants

The researchers conducted sessions for the present study at Western Michigan University in a small, private room. The entire project went through HSIRB approval (see Appendix A for approval letter). Seven undergraduate students in the psychology department volunteered to participate in the study; however, three of the students withdrew from the study prior to enough data being collected for analysis. As such, the researchers present the data from the four remaining participants in the current study. Three of these participants were female. All of the participants were freshman, 18-20 years of age, and Caucasian.

Inclusionary criteria included availability of two to three sessions per week and being a novice to functional analysis. It was necessary to identify participants with little to no knowledge of functional analysis conditions in order to evaluate the effectiveness of the independent variable so as to limit outside extraneous variables. This research was conducted with undergraduate students, as they are not likely to have encountered functional analysis methodology in their practica opportunities. Exclusionary criteria included having conducted or observed the implementation of a functional analysis or having a preference not to video self-monitor. Next, the researchers met with the participants and obtained information regarding their academic coursework, goals, and self-reported knowledge regarding functional behavior assessment. Participant demographic information are summarized in Table 2.

Table 2.

Participant Demographic Information

Questions	**Participant Responses**			
	P1	***P2***	***P3***	***P4***
Courses				
Number of psychology courses taken or currently enrolled in.	3	1	3	1
Have your courses involved a demonstration of how to conduct a functional analysis?	No	No	No	No
Have your courses included how to select function-based treatment?	No	No	No	No
Have your courses included how to collect direct assessment data (Antecedent Behavior Consequence (ABC) data, scatterplot data)?	No	No	Yes	No
Have your courses or work experience included how to collect indirect assessment data (functional assessment interview (FAI), questionnaires like QABF)?	No	No	Yes	No
How many of your classes have covered functional behavior assessments?	3	0	0	1
How many of your classes have covered functional analyses?	2	0	2	1
Work Experience				
Do you have any work experience where you have obtained training on functional analysis or functional behavior assessment?	No	No	No	No
Do you have any experience with problem behavior?	No	No	No	No
How many functional behavior assessments have you participated in?	2	0	0	0
How many functional analyses have you participated in?	1	0	0	0
If you had to rate your experience level in functional analysis (see above definition), how would you rate it at this moment.	Minimal	None	Minimal	Minimal
Future Goals				
Interested in obtaining BCBA certification.	Yes	No	Yes	Yes
Are you pursuing a career where you may have to conduct a functional analysis?	Yes	No	Yes	Yes
Are you pursuing a career where you might have to conduct a functional behavior assessment?	Yes	No	Yes	Yes

As demonstrated by Table 2 above, participants were all similar demographically. Participants reported having minimal levels of experience with functional analysis at best.

However, some of the answers for participants appeared contradictory, as demonstrated by P1's report of having covered functional behavior assessment and functional analysis in a course, but not having covered direct assessment data collection, indirect assessment, or functional analysis. Although the questionnaire had definitions of the difference between functional behavior assessment and functional analysis, it is possible that some participants may have misinterpreted some of the questions. Anecdotally, some participants reported having briefly covered what they thought was functional behavior assessment in a class or two, but that they were "unsure" even after reading the definitions.

Materials and Apparatus

Participants were paid $25 in Meijer gift cards for their participation in the study at the conclusion of the study if they participated in the study in its entirety. During some conditions of the study, the researchers required participants to review paper materials (e.g., task analyses, articles), or to self-monitor using paper-and-pencil data collection. The researchers used paper-and-pencil recording for performance measures, treatment integrity, and inter-observer agreement. Materials available during sessions for participants to select included journal magazines, a bear, the board game Sorry, coloring pages with coloring utensils, playing cards, Uno cards, and math sheets.

The lead researcher's password-protected personal computer was used to record all sessions. The computer was placed behind the researcher and the participant in the same place during each session, with both individuals in full view. Following all sessions, the researchers stored data on a password-protected flash drive kept in a locked filing cabinet inside of the Behavioral Medicine Laboratory (2704 Wood Hall). Additionally, the researchers placed a white noise blocker outside the room in order to preserve confidentiality of participants.

The researchers drafted seven scripts using the same format. Each script contained the attention, demand, tangible, and control functional analysis conditions, and totaled three minutes for each. The order of functional analysis conditions was randomized for each script. Each condition included the same number of target behaviors, inappropriate behaviors, and appropriate behaviors in random order. Table 3 below summarizes how many of each type of behavior (i.e., appropriate, inappropriate, target) were probed within each condition, which was held consistent for all scripts of that respective condition.

Table 3.

Number of Behavior Types in Each Script

Functional Analysis Condition	**Behavior Type**		
	Number of inappropriate behaviors	***Number of appropriate behaviors***	***Number of target behaviors***
Attention	3	7	6
Demand	3	4	6
Tangible	3	4	6
Control	3	7	6

The same three inappropriate behaviors were probed across all conditions in the same script. However, another three inappropriate behaviors were used in a different script. One of the inappropriate behaviors was always a verbal inappropriate behavior (e.g., "I hate you") and two of the inappropriate behaviors were physical behaviors (e.g., stomping feet). Appropriate behaviors were often the same behavior to ensure consistency across scripts (e.g., stating help and break in the demand condition). Three additional appropriate behaviors ("hey you," "check this out") in addition to the standard appropriate behaviors (e.g., "attention" in attention condition and "play" in control) were added to the attention and control conditions respectively to make these conditions more similar to the everyday life. The same target behavior was probed across all conditions in the same script, and never involved simulated physical aggression.

Finally, the demand condition also involved one probe for each script of non-compliance lasting approximately five to ten seconds.

Following the draft, the lead researcher created audio clips using iMovie while reading aloud from the script. These audio clips were saved as audio files, and transferred to Dropbox. Each script was saved separately under a folder titled with the script number, and inside the folder for the script number contained 4 audio files (each representing a different condition) labeled by condition and probe number (e.g., if a researcher opened the folder titled "script 1," the fourth probe would be labeled as "1.4"). An example of one of the conditions within a script can be found in Figure 1 below.

Six of the scripts were to be used during probes with participants, and the final script was used to create the video model. The lead researcher and a research assistant created four 3-minute video clips to serve as a video model for each condition containing the same materials that participants had at their disposal during sessions. These video clips were combined together using iMovie to create one video. For all performance probes, the lead researcher playing the role of the client used headphones connected to an Android phone with the Dropbox application to ensure consistency across sessions. The researchers assigned each session number a script (e.g., session 1 is script 3), and all participants received scripts in the same order.

Target behavior: *Self injury in the form of biting hand* Condition: *Control* Appropriate request: *"Play."* Other: *Play with toys in room when nothing is scripted*		
Time	**Behavior**	**Probe of…**
00:02	Bite hand	Target
00:20	Say "play"	Appropriate
00:26	Say "look at it!"	Appropriate
00:33	Say "play"	Appropriate
00:50	Bite hand	Target
1:07	Pulls own hair	Other inappropriate
1:30	Bite hand	Target
1:35	Bite hand	Target
1:37	Say "check it out!"	Appropriate
1:50	Slaps table	Other inappropriate
2:01	Say "look what I did"	Appropriate
2:09	Say "play with me please"	Appropriate
2:30	Bite hand	Target
2:40	Bite hand	Target
2:45	Yells "you're mean"	Other inappropriate
2:56	Say "play"	Appropriate

Figure 1. Sample Section of a Script. The columns from left to right in this picture denote the 1.) time a behavior was to occur, 2.) the specific behavior that the individual playing the role of the client was to emit, and 3.) what the probe entailed.

Research Assistant Training

Three undergraduate research assistants and one graduate student were recruited for this study to help collect treatment integrity and inter-observer agreement data. Training sessions for treatment integrity for in-session integrity data consisted of verbal instruction of the treatment integrity document corresponding to a given condition. Following verbal instruction, the research assistant was given an opportunity to practice during a session with a participant and feedback was provided. Following a 100% percent accuracy score, research assistants could collect treatment integrity data in-sessions on their own.

Specific conditions were assigned to each research assistant. Training sessions for inter-observer agreement consisted of one 1:1 meeting to go over the rubric for scoring for a specific condition. Following the verbal instruction, the researcher showed a video of a participant's performance that was not to be used for inter-observer agreement. The research assistant was asked to score criteria (e.g., target behaviors in the video) and explain why he or she scored in that manner. When all of the probes in a given video were correctly labeled as correct or incorrect and scored accordingly by the research assistant, the lead researcher allowed the research assistant to begin scoring inter-observer agreement independently for that respective condition.

Dependent Variables

Skill Mastery Measures

The primary dependent variable was the percentage correct scores on the task analyses in each condition. The task analysis components are outlined in Appendix B in the inter-observer agreement (IOA) chart, and these same FA components were used to score participant behaviors. These task analysis components were adopted and edited, with written permission (see Appendix C), based upon research conducted by Alissa Conway (Conway, 2018). The researchers tallied a frequency count in each column of the task analyses for the condition, and calculated percentage correct scores with the following equation for each condition: total correct divided by the total correct plus incorrect, and that outcome was multiplied by 100. For each FA condition, mastery criteria included 90% across the last two consecutive sessions within an experimental condition. The researchers recorded these data at a separate time, following sessions, by observing the videos of performance.

This study also has a number of secondary dependent measures. The components of the task analyses that were missed were reported as percentages as well to allow for analysis of component skills frequently missed. In order to obtain percentage data on component skills, the researchers took the total correct for a criteria within a condition divided by the total correct plus incorrect for that criteria within that condition, multiplied by 100. Then, the researchers averaged all participant scores across all conditions with the exception of baseline to obtain a total average for each criteria across participants. Additionally, the percent accuracy between participant's scores and researcher's scores was reported for video self-monitoring condition sessions and video self-monitoring plus feedback sessions. The researchers calculated percentage accuracy between scores by first identifying the percentage correct score for the participant's self-monitoring scores, following the same equation as percentage correct on task analyses. Then, the researchers lined the participant percentage correct with the researcher percentage correct, and divided the smaller number by the larger number to identify the percent agreement. Data collection on the preceding dependent variables occurred following the conclusion of a probe while observing a video of the participant's performance.

Finally, the training was timed and averaged across all participants to identify the total training time length for each session. These data were collected in vivo during sessions and performance probes.

Social Validity Measures

The researchers administered a brief questionnaire during the final session for each participant, which included subjective ratings on video self-monitoring, feedback, and applicability to future career goals (see Appendix D).

Inter-Observer Agreement

As noted previously, the primary researcher observed video recorded sessions in order to collect data on the primary dependent variable of percentage correct. Separately, a second researcher observed the videos without having access to the primary researcher's scores for the participant's condition and calculated his or her own frequency count for each criteria for his or her respective condition. Prior to the second researcher's data collection, the primary researcher randomly selected six to seven videos per participant for observers to score by using a random number generator, and provided a chart to the researchers assigned to secondary data collection for videos to observe.

Overall, inter-observer agreement was 93.89% across 34.72% of sessions (range: 72.22%-100%), which is the average score across all participants across all conditions scored. The average agreement in the attention, demand, tangible, and control conditions included 97.59%, 95.50%, 93.07%, and 89.42% respectively. As some of the criteria were a result of a frequency count and others were a result of occurrence or non-occurrence data for each condition, agreement was calculated differently for some criteria. If a section was ever labeled as not applicable, this was also counted toward the IOA score if both observers agreed that it was not applicable. The methodology of how agreement was calculated for each criterion can be found in Appendix C. In the far right column of this Appendix, reports for the average inter-observer agreement across all criteria for that specific condition are reported, and at the bottom of each condition, the averaged inter-observer agreement for that respective condition is reported.

Procedural Integrity

During sessions, a second researcher collected within-session procedural integrity data on the primary researcher's accuracy of implementation of the task analyses of each condition (e.g., video self-monitoring, baseline). The researchers calculated these percentages by dividing the total correct over the total correct plus incorrect within a session. The procedural integrity sheets were tailored to outline the steps of each condition, and the primary researcher had a copy of the procedural integrity to reference during sessions. Overall, procedural integrity within-sessions was 99.93% across 79.22% of sessions.

Following data collection, the primary researcher collected script procedural integrity data by reviewing task analyses and counting the number of inappropriate, appropriate, and target behaviors that occurred during a condition. These numbers were compared to the proposed number of behaviors in each category according to the script. The researcher then divided the smaller number by the bigger number and multiplied the outcome by 100 to obtain the script procedural integrity for that condition. The percentages identified in each of the four conditions were averaged together to obtain the total script procedural integrity for that session. Although the script outlined times for the target behaviors to occur, data were not collected on whether the behavior occurred at the precise time outlined in the script. Overall, procedural integrity on the scripts was 99.63% across 30.55% of sessions. The primary researcher also compared the number of behaviors she scored with the number of behaviors scored by a secondary researcher on his or her task analysis. The researcher divided the smaller number by the bigger number, and then multiplied the outcome by 100 to obtain percent agreement between researchers on script procedural integrity. Overall, percent accuracy of agreement was 99.71% across 100% of script procedural integrity recordings.

Experimental Design and Procedure

A concurrent, multiple baseline design across participants was utilized during this study. The staggered format of this type of design was used in order to limit the potential for extraneous variables (e.g., time of implementation, outside sources of education) that may explain any potential results aside from the independent variables. Participants did not receive sessions on the same days due to scheduling conflicts, although they started on the same week and ended on the same week in the study. Moreover, each participant received conditions in the same order during this study. Participants were assigned to a group. P1 and P2 were assigned to group 1, and P3 and P4 were assigned to group 2. Initially, participants were to receive one session instruction and a video model at the same time in a group-format; however, due to scheduling conflicts, group 1 received instruction in a group-format and group 2 received instruction and video model separately. All subsequent and preceding sessions were conducted in isolation for each participant. Data collection occurred for a total of 9 weeks. For all participants, sessions occurred 2 to 3 times each week across a period of 7 weeks. The two week follow-up occurred 2 weeks later.

Performance Probes

During all sessions, the researcher began by greeting participants and turning on the video camera. The camera was placed behind the researcher and participant approximately five feet off the ground. If participants reported hearing about functional analyses at all in their coursework or job, the researcher would record this on a separate sheet. A procedural checklist was used during all sessions to provide the researcher with the steps relevant to the condition, script, and session number. The researcher led sessions by stating that “remember, if you do not know how to proceed during each condition, you can state ‘I don’t know how to proceed, and we will end the condition and move to the next condition.” If at any point in time participants stated they were unsure following this statement, the researcher terminated that performance probe. Additionally, the researcher provided participants with demographic information regarding the hypothetical client, including the target behavior for the functional analysis, items ranging from mild to high preferred items, and magazines and math sheets. The preferred items were randomized each session. Then, the researcher stated “some of these items may be relevant to the condition you will be asked to run.”

Following this introduction, the primary researcher checked the procedural checklist to identify which condition to probe first, which corresponded to the script assigned to the session number and the order of conditions on the script. The primary researcher then led the probe with "the first condition we will ask you to run is the _____ condition. Which item or items do you need to run the ____ condition?" The participants then selected the item(s) necessary to run the condition, while the researcher prepared the script and put on headphones connected to the Android phone with the script. The participants did not receive feedback regarding their selection of items. The researcher told the participants when to "start" when participants indicated they were ready. The researcher simulating the client was complaint to instructions unless otherwise told to do so in the script (i.e., playing the game as instructed, answering questions). The researcher did not probe anything supplemental verbally that was not on the script. Sessions were terminated by the researcher if they ran the full time. Then, the video camera was re-started for the next condition to be probed. This process was repeated for the remaining three conditions. All four conditions were probed each session.

Baseline

The researchers conducted performance probes in the manner described in the prior section. They provided no feedback to the participants regarding their performance. If questions regarding functional analysis were asked, the primary researcher explained that she was unable to answer these questions. If data represented an ascending trend or if performance was greater than 50% across two consecutive sessions in baseline only, participants would meet exclusionary criteria. It should be noted that this did not occur at any point in this study.

Instruction and Video Model

When participants entered the Instruction and Video Model (IVM) condition of the experiment, they received instruction and a video model during the first session of the condition only. All subsequent sessions included only performance probes. During the first session, the researcher began the timer when the instruction began. The participants were provided with a paper copy of the methods section of the Iwata et al. (1982/1994) article and Day et al. (1988) article. Participants were then instructed to follow along as the researcher read through pages 201-203 of the Iwata et al. (1982/1994) article beginning at the section titled “social disapproval” and ending at the end of the “unstructured play” condition. Then, researcher read through the section titled “positive reinforcement condition” on page 569 of the Day et al. (1988) article as participants followed along. While reading through these sections of both articles, the researcher clarified which of these sections were similar to the performance probes (e.g., unstructured play is equivalent to the control condition) and any differences in the probes (e.g., length of consequences). The researcher also answered any questions posed by participants. Then, the researcher provided the participants a paper copy of the task analyses for each condition (see Appendix B for the specific criteria for each condition), and read through the sections as the participants followed along.

Next, the participants were shown a video model of a researcher performing each of the functional analysis conditions. The researcher answered any questions that the participants had during this time, and outlined each of the components of the task analysis throughout the video while it played. The researcher paused the video briefly to answer any questions, and resumed playing the video when finished. The video model occurred for roughly the same length of time for each participant, as it was a set length of 13 minutes and 37 seconds.

Following the instruction on the article and task analyses and video model, participants completed a 21-question written quiz, taken and adjusted from the article by Iwata et al. (2000), to 100% accuracy. The researchers added questions regarding the tangible condition to this quiz and eliminated redundant questions. Each time an error was emitted on the quiz, feedback was be provided in the form of the correct answer and an explanation for why this was the correct answer, and participants re-answer the questions. If participants missed a question again, the researcher provided similar feedback to before regarding the correct answer and an explanation, and required the participants to retake only the part of the quiz missed until 100% accuracy was achieved. After participants achieved 100% on the quiz, the total time was calculated by adding the time of the video model, the time of the instruction, and the time of the quiz. Following this first session, the researcher sent participants an electronic copy of the articles and the task analyses.

The researcher then conducted performance probes of the conditions with the participant. In the beginning of subsequent sessions to the first session of this condition, the researcher began the session by first asking participants if they had looked at the task analyses, articles, or any other items related to functional analysis since the last session and recorded their answers.

Video Self-Monitoring

Prior to the first Video Self-Monitoring (VSMN) session only, the researchers conducted a practice session with participants on how to video self-monitor. The researcher reviewed a written set of instructions regarding the purpose of the practice session and the target skill, providing descriptive praise, for the practice session. The target skill for the practice session was different than the functional analysis task analyses to simulate how to video self monitor with a simple skill, without providing participants experimenter feedback on the task analysis used in the study itself. The researcher instructed participants to tell her five things to do and that she would be compliant. Contingent on compliance, each participant was instructed to provide descriptive praise. The researcher went over the task analyses consisting of four components for descriptive praise with the participant and answered any questions. Following this instruction, the researcher videotaped the practice session while the participants practiced this skill. The video ended after praise was delivered contingent on compliance with five demands. Then, the participants transitioned to self-monitoring their performance alongside the researcher for each of the five probes. If their score matched the researcher's score, the they were eligible to move onto the VSMN condition. If their score did not match the researcher's score, then this process was repeated until the their score matched the researcher's score.

Next, the participants were asked to read through a specific section of the task analysis used to score in order to orient them to the task analysis. Then, participants were shown a video of their prior performance in each of the conditions after reading through the specific section of the task analysis. The order of the conditions shown to participants participant was randomized. Then, they were told they can score the behaviors however they would like, and they did not have to master scoring of their performance. Participants were not given the same scoring criteria as the researchers so as to simplify the sheet for scoring.

Immediately following the scoring of all four conditions, performance probes were conducted. The first condition probed in the performance probe was never the last one that participants self-monitored. No feedback or praise was provided to the participants for their performance unless feedback specific to the use of the checklist is requested (e.g., clarification of a specific section). All sessions were timed from the time participants started orienting to the first task analysis to self-monitor to the beginning of the performance probe.

VSMN and Experimenter Feedback

If participant performance did not reach mastery criteria during video self-monitoring, the researcher implemented VSMN plus experimenter feedback. The researcher scored the performance separate from participants before each session, using researcher scoring criteria. The researcher then had participants observe and score their performance using the task analyses identically to the VSMN condition. After they finished scoring, the researcher compared her scores to their scores, outlining any errors in recording after each condition. Moreover, the researcher provided detailed explanations of errors with the help of examples to outline the component missed. Occasionally, the researcher praised a specific component being scored correctly on the task analysis, especially if it was a more difficult component to score. The researchers did not provide feedback regarding performance on the videos (e.g., praise for correct performance, corrective feedback about inaccurate performance). Following feedback regarding a particular condition, this process was repeated for all other conditions from the prior session. If participant scores matched researcher scores, the researcher stated “our sheets match” then moved onto the next condition. Following scoring of the videos of all conditions, the researchers conducted performance probes with participants.

Maintenance

The researchers met with participants two weeks following the end of initial data collection to conduct a maintenance probe. The maintenance probe was conducted identical to baseline and consisted only of the performance probes.

RESULTS

Primary Dependent Variable Measure

Figure 2 below summarizes the results of each of the four participants on their percentage correct for the steps in each functional analysis test condition across the four intervention conditions. Averaged across all four participants, performance in baseline was low (6.9%) but was notably higher in each of the intervention conditions; IVM (75.76%), VSMN (84.84%), and VSMN plus feedback conditions (95.35%), respectively. These results maintained across all four participants at follow up. As demonstrated in Figure 2, all participants showed minimal accuracy during all FA conditions in baseline, and each participant demonstrated notably improvements in the accuracy of FA steps as the experimental interventions were implemented.

Although P1 did show an ascending trend in baseline in the control condition, this performance never exceeded 50%. P1 demonstrated mastery performance in the attention and tangible condition during the VSMN intervention. During baseline, P1 earned 0% for tangible condition throughout and in the attention condition one time. It should be noted that P1 did attempt to run these conditions when a score of 0% was achieved, with the exception of session 8. During session 8 within this condition, P1 stated inability to proceed in the tangible condition (i.e., this participant chose not to perform, resulting in no errors of omission or commission), but performance increased again thereafter. With the introduction of the VSMN plus feedback condition, P1's performance in the demand condition also reached mastery. P1 did not achieve mastery in the control condition. P1 reported looking at the task analyses one time on session 4 and never the articles during any sessions within any experimental condition.

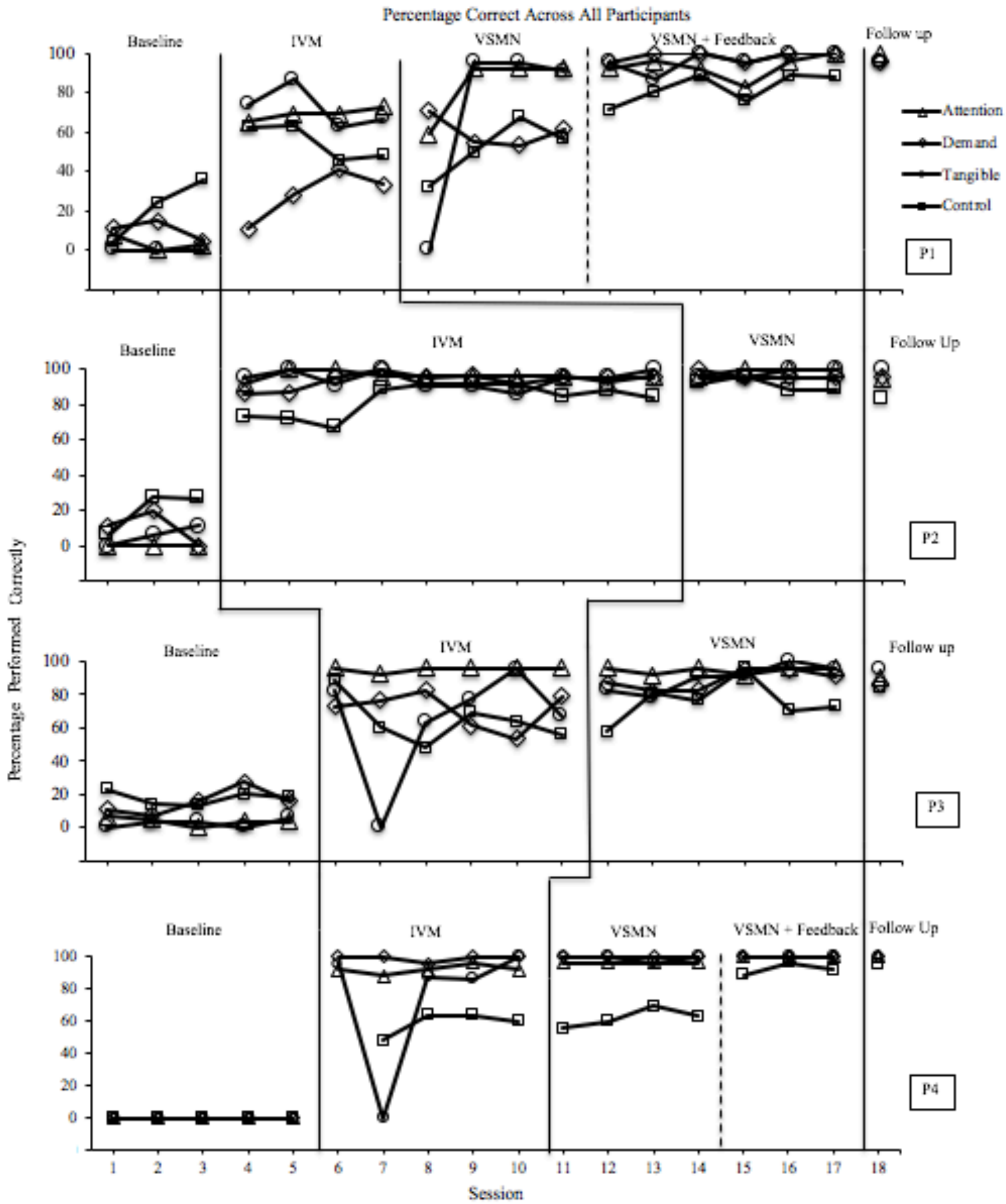

Figure 2. Graph of Participant Performance Within Sessions. This figure depicts a concurrent graph of the each participant's percentage correct performance in each of the functional analysis conditions.

P2 demonstrated mastery performance across the attention, demand, and tangible condition in the IVM condition. P2 did attempt all FA conditions across all experimental conditions, including baseline, although P2 did request to terminate some FA conditions during baseline. A level change is seen during session 7 as a result checking the task analysis before this session. This resulted in performance in the control condition increasing to mastery level for 3 sessions, but mastery criteria were not achieved as performance did not maintain until the end of this experimental condition. The researchers did not transition P2 into the VSMN plus feedback condition because there was not much room for improvement in data and due to time constraints. Although P2's performance improved slightly in the control condition with the addition of the VSMN, mastery criteria were still not achieved in this condition.

P3 demonstrated mastery performance across in the attention condition during the IVM condition. Performance dropped briefly to 0% during session 7 as a result of P3 confusing the tangible condition with the attention condition; however, P3 did still attempt to run this FA condition. P3 did attempt all FA conditions across all experimental conditions. P3 achieved mastery criteria in the demand and tangible condition with the addition of the VSMN condition. Similarly to P1 and P2, however, P3 did not achieve mastery criteria in the control condition. Prior to session 8, P3 reported looking at the task analysis 3 times, the articles once, and another item related to functional analysis 1 time. P3 reported no other instances throughout the duration of the study contacting any of these items.

P4 demonstrated mastery performance within the attention and demand conditions during the IVM condition. During the VSMN condition, P4 also achieved mastery during the tangible condition. Then, with the addition of VSMN plus feedback, P4 achieved mastery in the control condition. P4 was also only of the only participants to elect not to run certain FA conditions during baseline at all. Specifically, P4 did not attempt to run the demand condition during any baseline sessions, the tangible or attention condition following the first baseline session, and the control condition for the fourth and fifth baseline sessions. During session 7, P4 elected not to run the tangible condition. Additionally, P4 reported observing the task analyses the most out of all of the participants. P4 reported observing the task analyses 5 times during the IVM condition, 3 times during the VSMN condition, 2 times during VSMN plus feedback condition, and then 3 times prior to the follow up session.

The researchers ran a one-way between subjects ANOVA to compare the effect of the experimental condition on the average percentage correct. The null hypothesis of this ANOVA was that all means were equal. The F-Value resulting from the ANOVA was 112.25, and the P-Value was <0.0001, thus indicating a statistically significant effect of the experimental condition on the percentage correct across all participants and across all FA conditions.

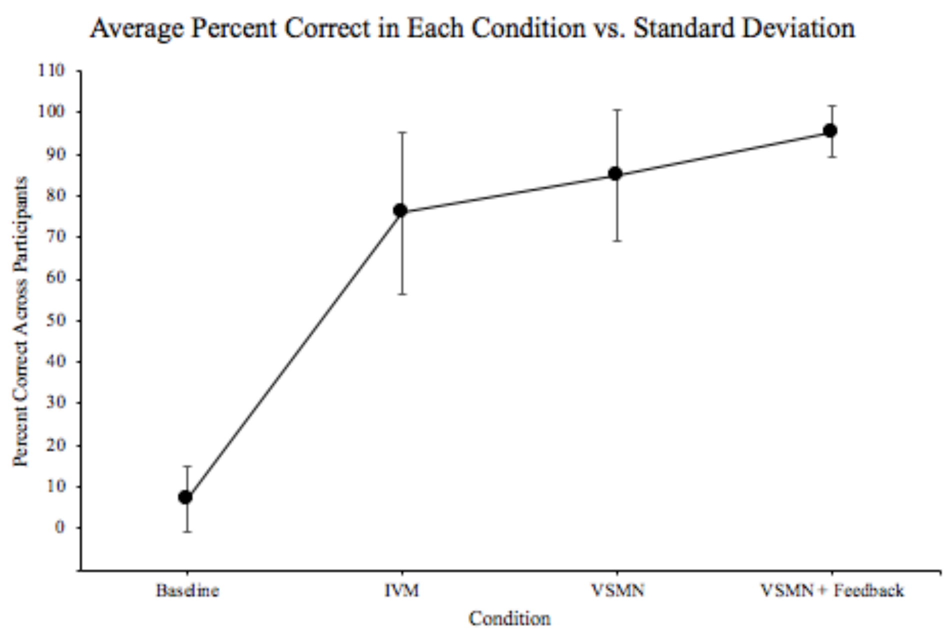

Figure 3. Statistical Analysis of FA Conditions. This graph depicts the mean percentage correct across all participants within each condition (represented by the data path with circles) and the standard deviation from the mean with 95% confidence interval (represented by the error bars).

The results for each experimental condition, summarized across all participants and FBA skills are depicted graphically in Figure 3. Overall, this graphic display documents a significant improvement in FA skills, averaged across all 4 participants with the most notable improvement occurring from Baseline to the IVM phase. Moreover, the standard deviation in mean scores across participants was higher in the IVM condition and VSMN condition, as demonstrated by the Figure 3, which would indicate more variability in performance among participants within these conditions than during Baseline or VSMN plus feedback. This variance tapered off in the VSMN plus feedback condition, although those data are obtained from a sample size of only 8 percentage scores and a ceiling effect was observed.

Secondary Dependent Variable Measures

Individual Component Data

Table 4 below represents the average percent correct across all participant data within all conditions with the exception of baseline. Baseline was excluded from this analysis as participants were not knowledgeable about the specific criteria yet. By excluding baseline data, these average percent correct provide some guidance on skill components that were easily mastered by the participants in this study and those components that proved more challenging. As indicated by data in bold within the table, there are a number of FA skill components that participants struggled with throughout the different conditions.

In particular, participants provided access to one moderate and one highly preferred toy in the attention condition only 78.57% of the time, and provided access to verbal attention prior to the beginning of the condition only 25% of the time. In the demand condition, participants began the condition with “it’s time to do some work,” 73.68% of the time, and they used the correct prompting hierarchy (least to most intrusive) 65.45% of the time. Within the tangible condition, participants struggled with the timing of consequences. In particular, they provided access to the toy for only 5 seconds approximately 64.81% of instances in the beginning of the condition, and provided the toy for only 5 seconds contingent on target behaviors only 76.60% of the instances. Finally, four criteria in the control condition fell below 80%. These task components of relatively low performance might be useful in prioritizing training interventions.

Table 4.

Participant Performance in Task Analysis Components

Condition	**Task Analysis Components**	**Average Score**
Attention	1. Provide access to at least 1 moderately and 1 highly preferred toy.	78.57%
	2. Gives verbal attention to client for 5 seconds prior to beginning condition.	25%
	3. Begins condition with statement similar to: "I'm sorry, but I cannot talk to you right now. I need to do some work."	100%
	4. Attends to some other item while "ignoring" client (not making eye contact, turning away, attending to another item).	100%
	5a. Provides verbal attention contingent on target behavior of interest, and attention must include related to problem behavior.	93.79%
	5b. If provided attention contingent on target behavior, was the attention high quality (e.g., greater than 3 words)?	92.42%
	6. Does not provide attention (no verbal statements, physical attn., or looking in their direction) for other inappropriate behaviors.	98.80%
	7. Does not provide attention (no verbal statements, physical attn., or looking in their direction) for appropriate behaviors (requests, greetings, invitations to play) or other verbal bids for attention (general statements).	100%
Demand	1. Selects demand-related items.	87.5%
	2. Begins condition with statement similar to "It's time to do some work now."	73.68%
	3. Demands are placed to complete work throughout session.	96.42%
	4. Delivers demands that can be physically prompted (math, one-step directions).	94.54%
	5. Continues prompting demands using correct prompting hierarchy if client stops completing work within 5 seconds.	65.45%
	6a. If the client engages in target behavior, provides a break and removes instructional items. No verbal statements are given.	88.39%
	6b. If provided break contingent on target behavior, did the break occur for approximately 5 seconds?	86.02%
	7. Ignores or neutrally redirects (gesturally or verbally) all inappropriate behaviors back to work without acknowledging them.	98.80%
	8. Ignores or neutrally redirects (gesturally or verbally) all appropriate behaviors back to work.	88.83%
Tangible	1. The most preferred item is selected for this condition	92.59%
	2. Begins by allowing the client to play with a highly preferred toy or item for at least 5 seconds.	64.81%
	3. Begins the condition with the statement, "My turn."	87.03%
	4. Immediately following the statement of "my turn," removes the item from client.	96.29%
	5. Pretends to play with item, eat snack, etc. after removing item from client.	98.14%
	6a. If the client engages in the target behavior at any time, provides the preferred item.	94.13%
	6b. if provided tangible item contingent on target behavior, did the access to tangible occur for approximately 5 seconds?	76.60%
	7. Does not deliver preferred item OR attention for appropriate behaviors.	97.68%
	8. The student does not deliver preferred item OR attention for inappropriate behaviors.	97.06%

Table 4 – Continued

Control	1. Selects at least 1 moderately and 1 highly preferred item for this condition.	81.81%
	2. Begins session with "here are some toys to play with."	63.63%
	3. Provides the client with the preferred toys to engage with.	100%
	4. Student plays with client (e.g., attempts to involve self in whichever free play activity client is engaging in) or engages in parallel play beside client.	100%
	5. Delivers high quality attention (not a simple "good" or "nice") on a fixed time schedule (30 seconds for probe).	59.33%
	6. If target behavior occurs when participant is scheduled to deliver attention (e.g., at 30 second interval), waits until there are 5 seconds free of target behavior before delivering scheduled attention, even if client engages in appropriate behaviors. OK if continues to engage in parallel play.	35.29%
	7. Does not provide attention (no verbal statements, physical attn., or looking in their direction) for target behaviors.	89.39%
	8. Does not provide attention (no verbal statements, physical attn., or looking in their direction) for other inappropriate behaviors.	98.78%
	9. Delivers high quality attention for all appropriate behaviors, requests, appropriate verbalizations, and answers appropriate questions.	65.51%

Note. The bold criteria indicate an average score of less than 80%.

Participants provided the beginning statement of "here are some toys to play with" 63.63% of the time. The fixed time schedule gave some participants some difficulty. In particular, they delivered high quality attention on a fixed time schedule 59.33% of instances, paused if target behavior occurred before delivering attention 35.29% of instances, and delivered high quality attention 65.51% of the instances. Otherwise, responses to target behaviors and other inappropriate behaviors remained above 88%.

Finally, there are other errors that are noteworthy that are not included in Table 4. For instance, participants often made errors of commission (i.e., performing a step that was not a part of the protocol) during the control condition. In particular, P1 and P3 consistently implemented social demands in the control condition during the IVM condition. Moreover, P1 and P3 confused the demand condition with another condition (e.g., the attention or tangible).

Self-Recording Percent Agreement

Average accuracy in self-recording across all participants and conditions was 88.72% in VSMN conditions. *Figure 4* depicts the average percent agreement between researcher recording and participant self-recording for participants. P1 and P4 received feedback regarding their accuracy of self-recording. A ceiling effect was seen in the tangible condition for P1, and in the demand and tangible conditions for P4. For all other conditions for both participants performances in accuracy were observed. P2 averaged at least 90%, and P3 averaged at least 83% across all FA conditions.

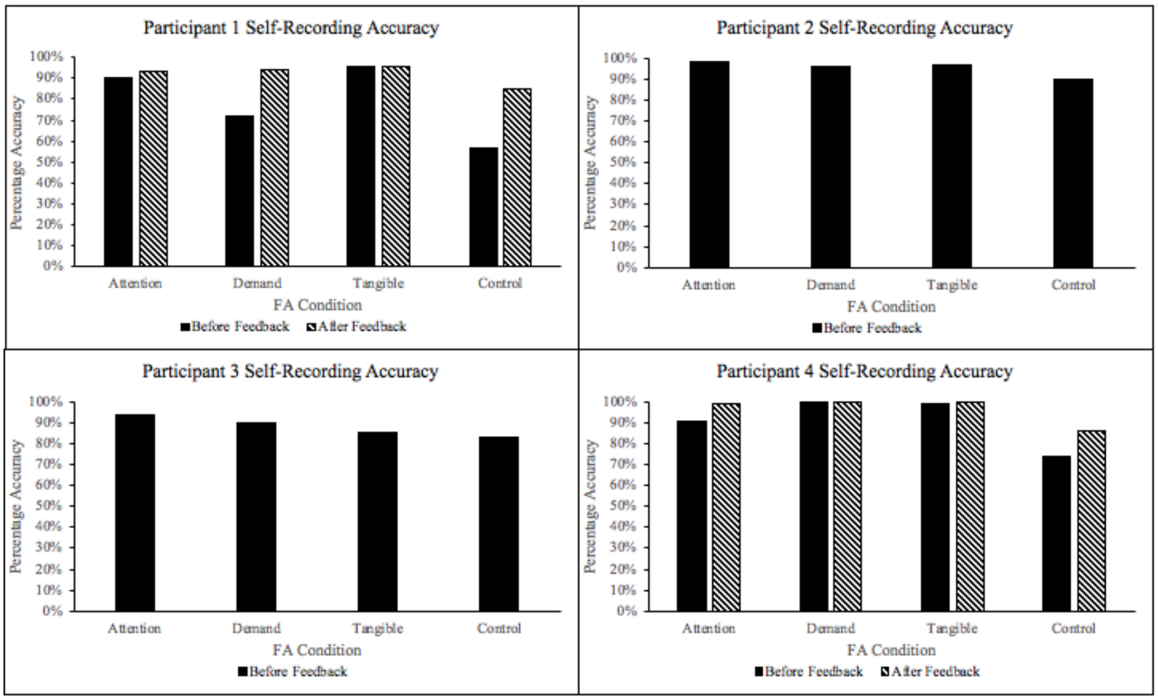

Figure 4. Participant Accuracy of Self-Recording. Within the graph, the black bars represent scoring during the VSMN condition, and the striped bars represent scoring during the VSMN plus feedback condition.

Average Length of Time

Figure 5 depicts the average time across all participants within each experimental condition. All performance probes occurred for approximately 14-17 minutes, depending on the time participants spent selecting items prior to beginning the condition and the introduction to the client in the beginning. Of that 14-17 minutes, 12 minutes was spent in a performance probe. The instruction and video model session averaged approximately 57.6 minutes (range: 44-75 minutes) from beginning to end, with the time spent taking the quiz included.

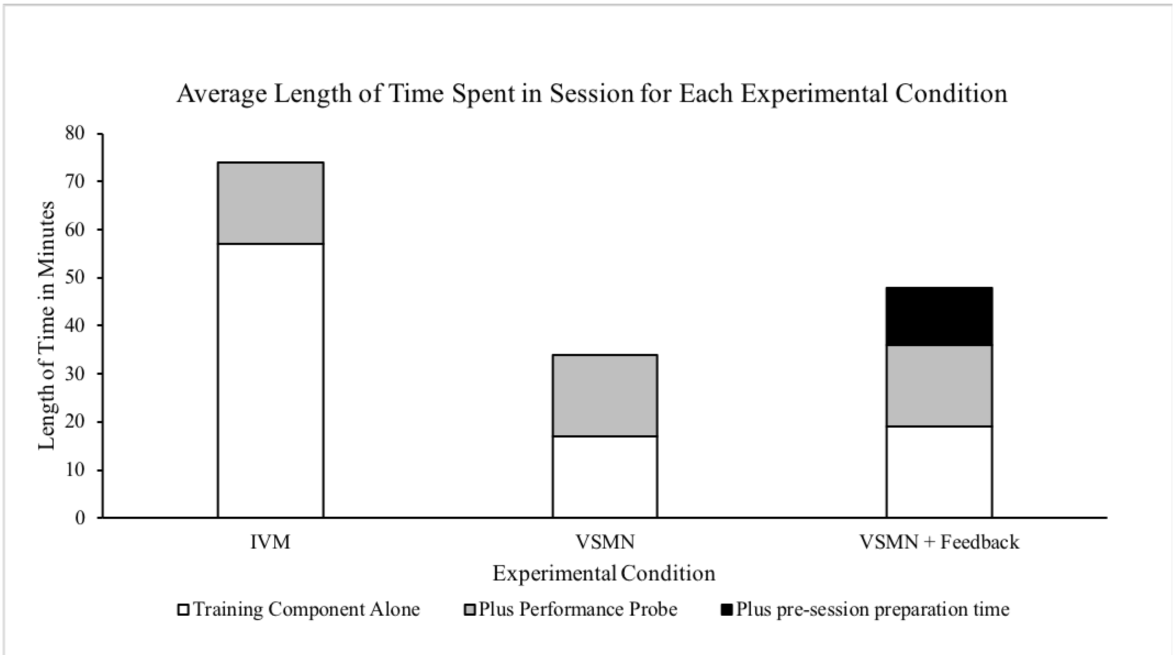

Figure 5. Average Session Lengths. In the following stacked bar graph, the white bars denote the length of time within the training component alone, the grey bars indicate the average length of time within a session with the addition of the performance probe to the training component, and the black bars indicate the average length of time preparing prior to the session.

For the following reported times, the researchers are excluding the performance probe time spent following the video self-monitoring and video self-monitoring plus feedback. It should be noted that an additional 14-17 minutes after self-monitoring (with or without feedback) was spent conducting a performance probe during each session. The time spent practicing video self-monitoring with a simple skill in the beginning of the first session within this condition ranged from 5-12 minutes. On average, the time spent video self-monitoring was approximately 17 minutes and 19 seconds (range: 13:57-22:18). On average, the time spent video self-monitoring plus feedback was generally 19 minutes and 51 seconds (range: 14:46-25:08). It should be noted that video self-monitoring plus feedback did require pre-session time on the part of the researcher that took approximately 12-13 minutes each session in order to score the session and compare researcher task analysis to the participant's.

Social Validity Measure

Table 5 below displays the participant's scores on the social validity questionnaire used to assess each participant's satisfaction with the intervention(s).

Table 5.

Social Validity Questionnaire Responses

Question	**Participant Responses**			
	P1	***P2***	***P3***	***P4***
"The video self-monitoring intervention was helpful to me in learning how to conduct a functional analysis."	Yes	Somewhat	Yes	Yes
"Do you feel like you learned more about functional analysis?"	Yes	Yes	Yes	Yes
"Was it aversive to observe yourself on the video?"	No	No	Yes	Somewhat
"How applicable do you feel that this will be for your future career?"	Very applicable	Somewhat applicable	Somewhat applicable	Very applicable
"How time-consuming was video self-monitoring?"	Not at all	Somewhat	Somewhat	Somewhat
"Was it more helpful or less helpful to receive researcher feedback than to provide yourself feedback?"	More helpful	N/A	N/A	About the same
"Which do you feel was the most helpful to you?"	Experimenter feedback	N/A	N/A	Video self-monitoring
Additional comments	N/A	N/A	N/A	N/A

As indicated by Table 5, 75% of participants reported that the video self-monitoring intervention as helpful to them. However, about 75% of participants also reported it was somewhat time consuming. 100% of participants stated that they felt as if they had learned more about functional analyses. 50% of participants reported that it wasn't aversive to view themselves on a video, while 25% reported it was aversive and 25% reported it was somewhat aversive. Overall, the results were favorable for the most part for video self-monitoring.

DISCUSSION

Overall, the results of this study demonstrate that IVM produced a notable performance in FA skills for all four of the participants, with 3 participants achieving mastery in at least one FA condition. However, this intervention alone was not sufficient to produce mastery levels of performance across all of the FA skill components. All participants showed further performance improvements with the sequential introduction of VSMN and subsequently with VSMN plus feedback. Three participants never achieved mastery in the control condition despite transitioning to a more intensive experimental condition; however, mastery was achieved in the attention, demand, and tangible condition across all participants. These results should be interpreted cautiously because of the obvious sequence effects that are apparent in the implementation of the same sequence of interventions for all four participants. However, each successive intervention represented an increment in the intensiveness of training condition and also an implementation sequence that is often seen in practical training setting.

Participant's accuracy in recording did improve with feedback during self-monitoring, resulting in a decreased length of time receiving feedback during subsequent sessions during this condition. The decreased length of time is notable because the less feedback participants received during this condition, the more similar this condition became to the preceding condition of just video self-monitoring on its own. This indicates that some participants may only need feedback regarding their scores on a few occasions in order to video self-monitor independently.

The results of this study indicate that all four participants benefited with the addition of the VSMN condition. This finding is similar to prior research, where it has been demonstrated across a range of skills that VSMN is effective for teaching simple skills (Belfiore et al., 2008; Bishop et al., 2015; Wright et al., 2012). Moreover, the results of the present study further demonstrated the effectiveness of VSM in teaching individuals to implement functional analyses to mastery levels of performance, similarly to Field et al. (2015). In the present study, each experimental condition represented an increase in the intensiveness of the independent variable and subsequent increases in performance. This same effect was observed in the study conducted by Field et al. (2015). As such, the interventions used in the present study could be seen as "additive" in that the researchers moved from the most "lean" method of instruction to the "fattiest" method (Tiemann & Markle, 1985) in terms of resources and time demands of each successive intervention.

The results of the present study have some notable differences from those found by Field et al. (2015). In the present study, IVM demonstrated more success than the instruction component within the Field et al. (2015) study. This difference may have been a result researchers in the present study structuring the instructional section more similarly to that described by Iwata et al. (2000), including task analyses, articles, a video model, and mastery performance on a quiz. In contrast, Field et al. (2015) provided instruction on the task analyses used in the study. While the present study included a longer instructional component, the researchers implemented this due to the success found in the study by Iwata et al. (2000), the lack of efficacy of instruction alone (Ward-Horner & Sturmey, 2012), and to more closely simulate a classroom environment.

One other difference in the findings in the present study from those found by Field et al. (2015) was the lack of mastery performance in the control condition within the present study. Field et al. (2015) demonstrated relatively stable, high performance across many participants within the control condition. By contrast, only one participant in the present study achieved mastery in the control condition. This result may have been due to a number of potential reasons. To start, mastery criteria differed between both studies, which may have led to differences in achievement of mastery between both studies. The scripts used in the present study may have been more difficult than those used in the Field et al. (2015) study, as Field et al. (2015) integrated only five target behaviors and no other appropriate or inappropriate behaviors whereas the present study probed more behaviors that participants had to respond to.

Moreover, participant behavior was assessed using different task analyses in both studies, with the present study requiring participants engage in different behaviors than those assessed in the Field et al. (2015) study. This may have led to differences in percentages as some of these component skills may have been more challenging. Finally, participants in the Field et al. (2015) study were graduate students and participants in the current study were undergraduate students; therefore, their entering repertoires may have been different.

The present study extends the findings of the study conducted by Field et al. (2015) in that the present study implemented VSMN instead of VPMN when mastery was not achieved during the first experimental condition, which included non-examples of performance. As noted by Field et al. (2015) in their discussion, the use of the non-examples, as opposed to perfect performance, may be a reason why the VSMN condition is effective. The results of the present study demonstrate that VSM was effective without the use of VPMN prior. One additional extension from the prior research was the inclusion of experimenter feedback if VSMN did not result in mastery performance. This led to further performance improvements, and did not take much additional experimenter time to integrate into protocol. As a result, this may be a viable alternative to BST for participants that need the supplemental feedback.

Limitations

The results should be approached with caution as the present study has several limitations. The results should be approached with caution as the present study has several limitations. While visual and statistical analysis suggest a significant treatment effect, the sample size was small and all of the participants had similar demographics. As such, it is not clear whether these results generalize to the general population. It would be helpful to replicate this study with additional participants and with participants with a wider range of backgrounds and entering skills (e.g., special education teachers, social work students, practicing BCBAs) for whom high level implementation of functional analysis assessments would be important.

A second, related limitation pertains to the simulated nature of the performance probes, which are likely different from functional analyses with real clients. As such, it is not clear how much these results generalize to the general population or with actual clients. Ultimately, the utility of any training interventions, including those that rely on simulations, will need to be assessed by an evaluation of how well those skills are applied in practice settings over extended periods of time. However, these simulations were based on real-life scenarios and proved challenging to the participants (as evidenced by the baseline levels of performance). Eventually, it would be helpful to identify the range of FA skill applications, possibly from performance assessments in real world applications, that would need to be included in any simulation-based training so as to increase the validity of the training simulations and improve the applicability of the resulting skills to real world settings. For example, expanding the simulations to include more challenging behaviors (e.g., aggression or severe dangerous behavior).

Additionally, the manner in which inter-observer agreement was calculated is a limitation as secondary observers recorded at a separate time from the primary observer. As such, although roughly 93% of the time the frequency scores were the same, however, it is not clear that the observers counted the same instances. Moreover, there is a potential for sequence effects as the conditions occurred in the same order for all participants and were not counter-balanced. The researchers implemented conditions in this fashion because the conditions were least to most intrusive across all participants and some conditions (e.g., instruction) may be a prerequisite for others (e.g., video self-monitoring). Finally, the lead researcher played the role of the client during all simulations, which may have aided in the acquisition of FA skills, as the learning environments contained similar stimuli during all probes. While this was helpful in ensuring consistency of difficulty in probing performance skills with participants, it may be the case that had another individual been used performance may not have improved as significantly among participants due to common stimuli instead of the experimental conditions alone.

Future Directions

Future research should include a component analysis of the specific variables that make up VSMN in order to identify key components. For example, variations on the amount of feedback to provide participants, as well as whether or not this feedback is necessary to foster results would be useful. Additionally, the training time for VSMN may still not be feasible in a classroom setting without modifying VSMN so that a researcher is not present during all observations. For the purposes of this study, the presence of a researcher was necessary in order to ensure integrity of the independent variables. Future research could explore variations on the frequency of probes to alleviate researcher time allotment, or could focus on VPMN with the integration of examples and non-examples due to its efficacy with only examples in the Field et al. (2015) study.

Interestingly enough, some participants achieved mastery during the IVM condition, which indicates that some participants may benefit from a less intrusive intervention. The two participants in the present study that showed the most significant improvement showed mastery on the quiz during the IVM condition during their first attempt. As such, the researchers speculate that although the participants were similar, they entered with different repertoires. Future research should assess the necessary entering repertoires among participants that would benefit the most from an intervention such as VSMN or IVM. Then, if an entering repertoire could be identified, interventions could be targeted in a tiered fashion to audiences for which they would be the most helpful.

Moreover, these results suggest that some training components were more challenging than others and it might be helpful to identify those conditions in simulation based training (and in real world applications) that merit supplemental training to insure high level performance of that component. It may also be interesting to see if preference has any impact on the efficacy of self-monitoring in a choice paradigm arrangement, as some participants reported preference for video self-monitoring, others for feedback, and others appeared to prefer neither.

Finally, the applicability of VSMN and VSMN plus feedback should be assessed with regard to conditions other than the standard FA conditions. To the author's knowledge, no study to date has assessed the utility of either of these interventions with other variations of FA conditions (e.g., social avoidance) or FA procedures (trial-based FAs). Depending on the setting, it may not be often in real world practice that the standard FA conditions are ran with clients. If there was more of a data base from real world practice, it might help us focus on more challenging conditions that could more closely approximate those found in the real world.

REFERENCES

Allinder, R. M., Bolling, R. M., Oats, R. G., & Gagnon, W. A. (2000). Effects of teacher self-monitoring on implementation of curriculum-based measurement and mathematics computation achievement of students with disabilities. *Remedial and Special Education, 21,* 219-226. DOI:10.1177/074193250002100403.

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.

Bailey, J. S., & Burch, M. R. (2016). *Ethics for behavior analysts (3rd edition).* New York, NY: Routledge.

Barnes, C. S., Mellor, J. R., & Rehfeldt, R. A. (2014). Implementing the verbal behavior milestones assessment an placement program (VB-MAPP): Teaching assessment techniques. *Analysis of Verbal Behavior, 30,* 36-47. DOI:10.1007/s40616-013-0004-5.

Behavior Analysis Certification Board. (n.d.). *Verified Course Sequence (VCS) directory.* Retrieved from http://info.bacb.com/o.php?page=100358.

Behavior Analysis Certification Board. (2016a). *BCBA examination pass rates for verified course sequences.* Retrieved from https://www.bacb.com/wp-content/uploads/BCBA-pass-rates-percent_180814.pdf.

Behavior Analysis Certification Board. (2016b). *Professional and ethical compliance code for behavior analysts.* Retrieved from https://www.bacb.com/wp-content/uploads/2017/09/170706-compliance-code-english.pdf.

Belfiore, P. J., Fritts, K. M., & Herman, B. C. (2008). The role of procedural integrity: Using self-monitoring to enhance discrete trial instruction (DTI). *Focus on Autism and Other Developmental Disabilities, 23,* 95-102. DOI:10.1177/1088357607311445

Bellini, S., & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional Children, 73,* 264-287. DOI:10.1177/001440290707300301.

Bishop, C. D., Snyder, P. A., & Crow, R. E. (2015). Impact of video self-monitoring with graduated training on implementation of embedded instructional learning trials. *Topics in Early Childhood Special Education, 35,* 170-182. DOI:10.1177/0271121415594797

Catania, C. N., Almeida, D., Liu-Constant, B., & Reed, F. D. D. (2009). Video modeling to train staff to implement discrete-trial instruction. *Journal of Applied Behavior Analysis, 42,* 387-392. DOI:10.1901/jaba.2009.42-387.

Chok, J. T., Shlesinger, A., Studer, L., & Bird, F. L. (2012). Description of a practitioner training program on functional analysis and treatment development. *Behavior Analysis in Practice, 5,* 25-36. DOI:10.1007/BF03391821.

Cihak, D. F., & Schrader, L. (2008). Does the model matter? Comparing video self-modeling and video adult modeling for task acquisition and maintenance by adolescents with autism spectrum disorders. *Journal of Special Education Technology, 23,* 9-20. DOI:10.1177/016264340802300302.

Conway, A. A. (2018). *Evaluating recommendations versus utilization of essential components of a functional analysis procedure (Published doctoral dissertation).* Western Michigan University, Kalamazoo, MI.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis (2nd edition).* Upper Saddle River, NJ: Pearson.

Day, R. M., Rea, J. A., Schussler, N. G., Larsen, S. E., & Johnson, W. L. (1988). A functionally based approach to the treatment of self-injurious behavior. *Behavior Modification, 12,* 565-589. DOI:10.1177/01454455880124005.

Deochand, N., & Fuqua, R. W. (2016). BACB certification trends: State of the states (1999 to 2014). *Behavior Analysis & Practice, 9,* 243-252. DOI:10.1007/s40617-016-0118-z.

Downs, H. E., Miltenberger, R., Biedronski, J., & Witherspoon, L. (2015). The effects of video self-evaluation on skill acquisition with yoga postures. *Journal of Applied Behavior Analysis, 48,* 930-935. DOI:10.1002/jaba.248.

Fetherston, A. M., & Sturmey, P. (2014). The effects of behavioral skills training on instructor and learner behavior across responses and skill sets. *Research in Developmental Disabilities, 35,* 541-562. DOI:10.1016/j.ridd.2013.11.006.

Field, S. P., Frieder, J. E., Mcgee, H. M., Peterson, S. M., & Duinkerken, A. (2015). Assessing observer effects on the fidelity of implementation of functional analysis procedures. *Journal of Organizational Behavior Management, 35,* 259-295. DOI:10.1080/01608061.205.1093058.

Filipiak, S. N., Rehfeldt, R. A., Heal, N. A., & Baker, J. C. (2010). The effects of points for preparation guides in interteaching procedures. *European Journal of Behavior Analysis, 12,* 115-132. DOI:10.1080/15021149.2010.11434338.

Gelbar, N. W., Anderson, C., McCarthy, S., & Buggey, T. (2012). Video self-modeling as an intervention strategy for individuals with autism spectrum disorders. *Psychology in the Schools, 49,* 15-22. DOI:10.1002/pits.20628.

Graudins, M. M., Rehfeldt, R. A., DeMattei, R., Baker, J. C., & Scaglia, F. (2011). Exploring the efficacy of behavioral skills training to teach basic behavior analytic techniques to oral care providers. *Research in Autism Spectrum Disorders, 6,* 978-987. DOI:10.1016/j.rasd.2011.12.010.

Hogan, A., Knez, N., & Kahng, S. (2015). Evaluating the use of behavioral skills training to improve school staffs' implementation of behavior intervention plans. *Journal of Behavioral Education, 24,* 242-254. DOI:10.1007/s10864-014-9213-9.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982/1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197-209. DOI:10.1901/jaba.1994.27-197.

Iwata, B. A., & Dozier, C. L. (2008). Clinical application of functional analysis methodology. *Behavior Analysis in Practice, 1,* 3-9. DOI:10.1007/BF03391714.

Iwata, B. A., Wallace, M. D., Kahng, S. W., Lindberg, J. S., Roscoe, E. M., Conners, J., Hanley, G. P., Thompson, R. H., & Worsdell, A. S. (2000). Skill acquisition in the implementation of functional analysis methodology. *Journal of Applied Behavior Analysis, 33,* 181-194. DOI:10.1901/jaba.2000.33-181.

Markle, S. M. (1967). *Empirical texting of programs.* In Phil C. Lange (Ed.) Programmed Instruction. The sixty-sixth Yearbook of the National Society for the Study of Education, Part II, Chicago: The University of Chicago Press.

Palmen, A., Didden, R., & Korzilius, H. (2010). Effectiveness of behavioral skills training on staff performance in a job training setting for high-functioning adolescents with autism spectrum disorders. *Research in Autism Spectrum Disorders, 4,* 731-740. DOI:10.1016/j.rasd.2010.01.012.

Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-based staff training: A guide for practitioners. *Behavior Analysis in Practice, 5,* 2-11. DOI:10.1007/BF03391819.

Sellers, T. P., Valentino, A. L., & LeBlanc, L. A. (2016). Recommended practices for individual supervision of aspiring behavior analysts. *Behavior Analysis in Practice. 4,* 274-286. DOI:10.1007/s40617-016-0110-7.

Tiemann, P. W., & Markle, S. M. (1990). *Analyzing instructional content: A guide to instruction and evaluation (4th edition).* Champaign, IL: Stripes Publishing LLC.

Van Houten, R. V., Axelrod, S., Bailey, J. S., Favell, J. E., Foxx, R. M., Iwata, B. A., & Lovaas, O. I. (1988). Right to effective behavioral treatment. *Journal of Applied Behavior Analysis, 21,* 381-384. DOI:10.1901/jaba.1988.21-381.

Ward-Horner, J., & Sturmey, P. (2012). Component analysis of behavioral skills training in functional analysis. *Behavioral Interventions, 27,* 75-92. Doi:10.1002/bin.1339.

Wright, M. R., Ellis, D. N., & Baxter, A. (2012). The effect of immediate or delayed video-based teacher self-evaluation on head start teachers' use of praise. *Journal of Research in Childhood Education, 26,* 187-198. DOI:10.1080/02568543.2012.657745.

APPENDIX A

HSIRB Approval Letter

WESTERN MICHIGAN UNIVERSITY

Institutional Review Board
FWA00007042
IRB00000254

Date: February 19, 2019

To: Wayne Fuqua, Principal Investigator
Haley Hughes, Student Investigator for Thesis
Student Investigators: Angel Liceaga, Shanise Carlson, Drew Farrell, Jessica Fett, Patrick Wieszciecinski, Jennifer Trobaugh, Meredith Holland

From: Amy Naugle, Ph.D., Chair Amy Naugle

Re: IRB Project Number 18-10-24

This letter will serve as confirmation that the changes to your research project titled "An Analysis of Video Self-Monitoring as a method to Train Students to Implement Functional Analysis" requested in your memo received February 18, 2019 (to revise follow up assessments from two to one; revise compensation distribution with total remaining $25; revise consent document to reflect this change) have been approved by the WMU Institutional Review Board.

The conditions and the duration of this approval are specified in the Policies of Western Michigan University.

Please note that you may **only** conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the IRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: **October 29, 2019**

Office of the Vice President for Research
Research Compliance Office
1903 W. Michigan Ave., Kalamazoo, MI 49008-5456
PHONE: (269) 387-8293 FAX: (269) 387-8276
WEBSITE: wmich.edu/research/compliance/hsirb

CAMPUS SITE: Room 251 W. Walwood Hall

APPENDIX B

IOA Calculations and Average Scores

IOA Calculations and Average Scores

The criteria listed for each FA condition were edited based on another research colleague's dissertation (Conway, 2018).

Attention Condition				
Criteria	**IOA calculation for each participant**		**Average score per each section across all participants**	
For any of the following criteria, if the researchers both score the box as N/A, then this is considered 100% IOA for that respective box. If one researcher scores it as N/A and the other does not, then this is considered 0% IOA.				
1. **Provide access to at least 1 moderately and 1 highly preferred toy.** (	Agreement/(agreement + disagreement) X100 = total % agreement for criteria 1-4.		99%	
2**. Gives verbal attention to client for 5 seconds prior to beginning condition.**				
3. **Begins condition with statement similar to: "I'm sorry, but I cannot talk to you right now. I need to do some work."**				
4. **Attends to some other item while "ignoring" client (not making eye contact, turning away, attending to another item).** (Tally as 1 correct or incorrect. Must do for entire time (except following TB) to be correct.				
5a. **Provides verbal attention contingent on target behavior of interest, and attention must include related to problem behavior.**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	98%	98%
5b. If provided attention contingent on target behavior, was the attention high quality (e.g., greater than 3 words)?	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	100%	98%
6. **Does not provide attention (no verbal statements, physical attn., or looking in their direction) for other inappropriate behaviors.**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	98%	98%
7. **Does not provide attention (no verbal statements, physical attn., or looking in their direction) for appropriate behaviors (requests, greetings, invitations to play) or other verbal bids for attention (general statements).**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	96%	98.66%
8. **Does not deliver any demands to the participant**.	Smaller count/ larger count X 100 = total count IOA%		93.68%	
9. **Does not provide attention for any other reasons aside from following target behavior.**	Smaller count/ larger count X 100 = total count IOA%		93.78%	
10. **States that he or she is unsure how to proceed before condition is over.**	Smaller count/ larger count X 100 = total count IOA%		100%	
Total:	Average IOA= Individual IOA scores (per incorrect and incorrect sections above)/ Total number of sections (12)		97.59%	

Demand Condition		
Criteria	**IOA calculation for each participant**	**Average score per each section across all participants**
For any of the following criteria, if the researchers both score the box as N/A, then this is considered 100% IOA for that respective box. If one researcher scores it as N/A and the other does not, then this is considered 0% IOA.		
1. **Selects demand-related items**.	Agreement/(agreement + disagreement) X100 = total % agreement for criteria 1-5.	94.4%
2. **Begins condition with statement similar**		

to "It's time to do some work now."				
3. **Demands are placed to complete work throughout session.**				
4. **Delivers demands that can be physically prompted (math, one-step directions).**				
5. **Continues prompting demands using correct prompting hierarchy *if client stops completing work* within 5 seconds.**				
6a. **If the client engages in target behavior, provides a break and removes instructional items. No verbal statements are given**.	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	96%	96%
6b. If provided break contingent on target behavior, did the break occur for approximately 5 seconds?	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	95.2%	89.33%
7. **Ignores or neutrally redirects (gesturally or verbally) all inappropriate behaviors back to work without acknowledging them.**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	92%	96.66%
8. **Ignores or neutrally redirects (gesturally or verbally) all appropriate behaviors back to work.**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	100%	100%
9. **Does not provide access to preferred games or toys**.	Smaller count/ larger count X 100 = total count IOA%		100%	
10. **Does not provide reprimands or other attention for any other reasons throughout condition.**	Smaller count/ larger count X 100 = total count IOA%		81.96%	
11. **Does not promise a break upon completion of work**.	Smaller count/ larger count X 100 = total count IOA%		100%	
12. **States that he or she is unsure how to proceed before condition is over.**	Smaller count/ larger count X 100 = total count IOA%		100%	
Total:	Average IOA= Individual IOA scores (per incorrect and incorrect sections above)/ Total number of sections (13)		95.50%	

Tangible Condition				
Criteria	**IOA calculation for each participant**		**Average score per each section across all participants**	
For any of the following criteria, if the researchers both score the box as N/A, then this is considered 100% IOA for that respective box. If one researcher scores it as N/A and the other does not, then this is considered 0% IOA.				
1. **The most preferred item is selected for this condition**	Agreement/(agreement + disagreement) X100 = total % agreement for criteria 1-5.		94.4%	
2. **Begins by allowing the client to play with a highly preferred toy or item for at least 5 seconds.**				
3. **Begins the condition with the statement, "My turn."**				
4. **Immediately following the statement of "my turn," removes the item from client**.				
5. **Pretends to play with item, eat snack, etc. after removing item from client.**				
6a**. If the client engages in the target behavior at any time, provides the preferred item.**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	97.86%	88.53%
6b. if provided tangible item *contingent on target behavior, did the access to tangible occur for approximately 5 seconds?*	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	92.26%	82%
7. **Does not deliver preferred item OR attention for appropriate behaviors.**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	99%	91.64%
8. **The student does not deliver preferred item OR attention for inappropriate behaviors**.	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	96.33%	88.66%
9. **The tangible item is NOT provided for any other reasons aside from target behavior.**	Smaller count/ larger count X 100 = total count IOA%		96%	
10. **No other demands OR attention other**	Smaller count/ larger count X 100 = total count IOA%		90.15%	

than “my turn” are placed on client (social demands or other).		
11. **States that he or she is unsure how to proceed before condition is over.**	Smaller count/ larger count X 100 = total count IOA%	100%
Total:	Average IOA= Individual IOA scores (per incorrect and incorrect sections above)/ Total number of sections (12)	93.07%

Control Condition				
Criteria	**IOA calculation for each participant**		**Average score per each section across all participants**	
For any of the following criteria, if the researchers both score the box as N/A, then this is considered 100% IOA for that respective box. If one researcher scores it as N/A and the other does not, then this is considered 0% IOA.				
1. **Selects at least 1 moderately and 1 highly preferred item for this condition**.	Agreement/(agreement + disagreement) X100 = total % agreement for criteria 1-4.		93%	
2. **Begins session with “here are some toys to play with.”**				
3. **Provides the client with the preferred toys to engage with.**				
4. **Student plays with client (e.g., attempts to involve self in whichever free play activity client is engaging in) or engages in parallel play beside client.**				
5. **Delivers high quality attention (not a simple “good” or “nice”) on a fixed time schedule (30 seconds for probe).**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	84.39%	77.19%
6. **If target behavior occurs when participant is scheduled to deliver attention (e.g., at 30 second interval), waits until there are 5 seconds free of target behavior before delivering scheduled attention, even if client engages in appropriate behaviors. OK if continues to engage in parallel play.**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	96%	90%
7. **Does not provide attention (no verbal statements, physical attn., or looking in their direction) for target behaviors.**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	94.66%	91%
8. **Does not provide attention (no verbal statements, physical attn., or looking in their direction) for other inappropriate behaviors.**	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	95.33%	92%
9. **Delivers high quality attention for all appropriate behaviors, requests, appropriate verbalizations, and answers appropriate questions**.	Smaller count/ larger count X 100 = total count IOA%	Smaller count/ larger count X 100 = total count IOA%	95.34%	79.33%
10. **Student does not place any demands or asking questions throughout the session.**	Smaller count/ larger count X 100 = total count IOA%		79.64%	
11. **Student does not take away toys at any time.**	Smaller count/ larger count X 100 = total count IOA%		88%	
12. **States that he or she is unsure how to proceed before condition is over.**	Smaller count/ larger count X 100 = total count IOA%		96%	
Total:	Average IOA= Individual IOA scores (per incorrect and incorrect sections above)/ Total number of sections (14)		89.42%	

APPENDIX C

Copyright Approval

Re: Copyright Request

Alissa Anne Conway

Wed 5/22/2019 4:13 PM

To:Haley Ciara Hughes <haley.c.hughes@wmich.edu>;

Hello Haley,

You have my consent to utilize these materials in your thesis document and for publication purposes.

Thank you,
Alissa Conway, Ph.D., BCBA-D
Drexel University

On May 22, 2019, at 9:17 AM, Haley Ciara Hughes <haley.c.hughes@wmich.edu> wrote:

Dear Dr. Conway:

I would like to request your permission to include an excerpt from the following item in my dissertation:
Conway, A. A. (2018). Evaluating recommendations versus utilization of essential components of a functional analysis procedure (Unpublished doctoral dissertation.) Western Michigan University, Kalamazoo, MI.

I would like permission to include information listed in your dissertation, specifically component steps in functional analyses that I have used to create data sheets for my master's thesis.

The source will receive full credit in the manuscript.

By agreeing to the use of the item in my dissertation, you give ProQuest Information and Learning (PQIL) the right to supply copies of this material on demand as part of my doctoral dissertation. Please attach any other terms and conditions for the proposed use of this item.

If you no longer hold the copyright to this work, please indicate to whom I should direct my request.

Thank you for your time and attention to this matter.

Sincerely,

Haley Hughes, B.S.

WMU/Summit Pointe Behavioral Consultation Services

Program and Speaker Coordinator for the Michigan Autism Conference

Behavior Analysis Master's Student

Western Michigan University

APPENDIX D

Social Validity Questionnaire

Participant ID ________

Date ___________

Evaluation of Video Self Monitoring

Directions: Please provide the corresponding answer that best matches your response to each of the questions below.

1. **The video self monitoring intervention was helpful to me in learning how to conduct a functional analysis.**

No Somewhat Yes

If there was any specific skill, or skills, that you feel video self-monitoring was not helpful in learning, please list: ____.

2. **Do you feel like you learned more about functional analysis?**

No Somewhat Yes

3. **Was it aversive to observe yourself on the video?**

No Somewhat Yes

4. **How applicable do you feel that this will be for your future career?**

Not at all Applicable Somewhat Applicable Very Applicable

5. **How time-consuming was video self-monitoring?**

Not at all Somewhat Very

<u>Only for participants that received more than just video self-monitoring:</u>

6. **Was it more helpful or less helpful to receive researcher feedback than to provide yourself feedback?**

Less Helpful About the Same More Helpful

7. **Which do you feel was the most helpful to you?**

Video self-monitoring Experimenter Feedback

Any additional comments, questions, concerns, or thoughts that you would like to share with the researcher: