EXAMINING THE ACCURACY OF THE SERIAL RECORDING FORM WHEN USED BY SCHOOL PSYCHOLOGISTS

By

LILIANA WAGNER

(Under the Direction of Scott Ardoin)

ABSTRACT

Insufficient research exist evaluating the quality of behavioral assessment methods in a way that takes into account a principle tenant of behavioral assessment--that consistency in behavior across time or settings cannot be assumed. This dissertation first examined the quality of different direct FBA methods, specifically their accuracy, sensitivity, and treatment utility, with a focus on A-B-C recording. The second aim of this dissertation was to replicate and extend research by evaluating the accuracy of the Serial Recording form, a hybrid form of A-B-C recording that combines the practitioner-friendly format of structured recording with the superior analytic properties of continuous recording. The study also examined a range of potential factors that can impact accuracy when conducting direct observations. Forty-four school psychologists and four advanced graduate students in school psychology were randomly assigned to one of two conditions: *Informed*, in which participants were provided with a hypothesized function of the target student's inappropriate behavior and *Neutral*, in which participants were not provided with a hypothesized function. Participants were trained to use the Serial Recording form and then watched a 10-min video during which they coded environmental events and target behaviors in the sequence in which they were observed. Participants then identified the function of the target behavior. Results showed unacceptable levels of accuracy in

recording the occurrence and sequence of environmental events and student behaviors. Despite the poor accuracy, the majority of participants selected the correct function of behavior. A secondary aim of this study was to gain a broader understanding of how school psychologists are conducting FBAs in schools, as there is evidence to suggest that a failure to conduct a thorough FBA can have legal implications. This sample of participants are not including many important components (identification of target behavior, hypothesis of function) in their FBAs. Further, many participants are not regularly conducting FBAs in their schools. Findings from this study have important implications for continuing to evaluate the quality of behavioral assessment methods and training practitioners in function-based assessment and treatment.

INDEX WORDS: Functional Behavior Assessment, Direct assessment methods, A-B-C recording, Continuous recording, School psychologists

EXAMINING THE ACCURACY OF THE SERIAL RECORDING FORM WHEN USED BY SCHOOL PSYCHOLOGISTS

by

LILIANA WAGNER

B.A., Emerson College, 2007

M.Ed., Vanderbilt University, 2013

A Dissertation submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSPHY

ATHENS, GEORGIA

2018

EXAMINING THE ACCURACY OF THE SERIAL RECORDING FORM WHEN USED BY SCHOOL PSYCHOLOGISTS

by

LILIANA WAGNER

Major Professor:	Scott Ardoin
Committee:	Kevin Ayres Stacey Neuharth-Pritchett Kristin Sayeski

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2018

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Functional behavior assessment (FBA) is a set of information-gathering techniques, arranged on a continuum from indirect to direct methods, with the goal of identifying environmental variables and conditions associated with the presence of a target behavior (Alter, Conroy, Mancil, & Haydon, 2008; Cooper, Heron, & Heward, 2007; Gresham, Watson, & Skinner, 2001). FBAs employ many of the same tools utilized in traditional psychological assessments (e.g., rating scales, direct observation); however, they are unlike traditional assessments in that the focus is on identifying alterable environmental variables as opposed to intrinsic qualities of the individual (Shriver, Anderson, & Proctor, 2001).

With a foundation in applied behavior analysis, central to FBA is the idea that behavior cannot be examined in isolation from the environment (Shriver et al., 2001). FBAs are used to identify the events that reliably precede (antecedents) and follow (consequences) problem behavior, ideally using direct observation and other low-inference techniques (McIntosh, Borgmeier, Anderson, Horner, Rodriguez, & Tobin, 2008). As the same behavior may be controlled by multiple variables, or functions, the behavior should be evaluated in multiple environments using a variety of assessments and informants (Alter et al., 2008, Shriver et al., 2001).

When the function of, or reason why, a behavior that occurs in a specific context is identified, individualized, function-based treatments can be created that have a higher probability of reducing problem behavior and increasing appropriate alternative behavior (Campbell, 2003).

Such interventions may involve teaching a new skill or altering some aspect of the environment to facilitate a change in behavior (Cooper et al., 2007; Shriver et al., 2001). Function-based treatments are tailored to the individual and are more effective in reducing problem behavior than universal behavior modification techniques (Cooper et al., 2007; Gable, Park, & Scott, 2014; Hanley, Jin, Vanselow, & Hanratty, 2014; Herzinger & Campbell, 2007; Newcomer & Lewis, 2004). There is no one form of treatment that is effective across every behavior or situation (VanDerHeyden, Witt, & Gatti, 2001). Research suggests that most classroom interventions are developed to override or compete with the problem behavior using highly desirable reinforcers or overly aversive punishers (Couvillon, Bullock, & Gable, 2009; Newcomer & Lewis, 2004). Although delivery of punishment may temporarily reduce a child's problem behavior, it is unlikely to address the underlying reason for the behavior. A one-size-fits-all approach to behavior reduction increases the chance that a child will engage in another equally undesirable behavior in an attempt to access the same reinforcer. Function-based treatments are therefore more likely to prevent the occurrence of future problem behavior (Cooper et al., 2007) and to incorporate positive behavior change strategies over punishment-based techniques (Cooper et al., 2007; IDEA, 2004; Iwata, Dorsey, Slifer, Bauman, & Richman, 1982).

Although schools have employed function-based techniques for decades, including FBA as part of the development of behavioral interventions was not mandated in schools until the 1997 amendment to the Individuals with Disabilities Education Act (IDEA; Gresham et al., 2001). IDEA (1997, 2004) requires that an FBA be created and employed by a child's individualized education plan (IEP) team in response to disciplinary actions taken by school personnel. Specifically, an FBA and behavior intervention plan (BIP) are required in cases of a

discipline-related change in educational placement (i.e., removal for more than 10 school days) for a child with a disability when the conduct of the child is determined to be a manifestation of the child's disability (IDEA, 2004). The data collected through the FBA should be used for developing a new BIP for the student or making modifications to an existing plan. IDEA (1997, 2004) also stipulates that the IEP team consider employing positive behavioral interventions, which may include an FBA and/or BIP, when a child's problem behavior negatively affects their learning, or the learning of others (Zirkel, 2011).

Despite the legal mandate for its use, there is little direction on how the FBA process should be implemented, and the specific assessment methods that should be included (Asmus, Vollmer, & Borrero, 2002; Ervin, Ehrhardt, & Poling, 2001; Stage et al., 2006). FBA methods typically include indirect assessment methods, experimental manipulations, such as a structural or functional analysis, and/or direct or descriptive assessment methods (Alberto & Troutman, 2006; Cooper et al., 2007). Due to the lack of specificity provided by the law, many researchers advocate a combination of assessment methods to assess behavioral function (Cooper et al., 2007; Thompson & Borrero, 2011; Weber, Killu, Derby, & Barretto, 2005).

Indirect. Indirect FBA methods are used to gather information about behavior using methods that are removed in time and place from the behavior of interest such as interviews, rating scales, and record reviews. As users of such assessment methods rely on informant recall to generate a hypothesis of function and do not involve direct observation of the behavior, they are sometimes criticized for generating biased data (Cooper et al., 2007; Rooker, DeLeon, Borrero, Frank-Crawford, & Roscoe, 2015). Thus, it is often necessary to incorporate experimental or direct methods of assessment into the FBA.

Functional analysis. Experimental functional analysis (EFA) involves the systematic manipulation of variables hypothesized to be controlling the target behavior such as access to attention or escape from a task (Fisher, Adelinis, Thompson, Worsdell, & Zarcone, 1998). Due to the level of control necessary for EFAs, they are typically conducted by trained professionals in an analog, as opposed to applied, setting (Cooper et al., 2007). An EFA is the only method that allows practitioners to confirm hypotheses (Asmus et al., 2002; Cooper et al., 2007) and research has demonstrated that the largest reductions in problem behavior are observed when an EFA is included in the assessment process (Campbell, J. M., 2003; Hanley et al., 2014).

An EFA is considered the "gold standard" for identifying behavioral function (Lanovaz, Argumedes, Roy, Duquette, & Watkins, 2013; Shriver et al., 2001), but there are many limitations to its use in applied settings including difficulty, time constraints, and limited generalizability (Gresham et al., 2001). For example, a reinforcer identified in an analog assessment may lose its value when applied in the classroom (Shriver et al., 2001). To obtain a complete picture of the contingencies operating in the natural environment one may turn to direct FBA methods[1].

Direct. Direct FBA methods involve direct observation of the behavior and the surrounding environmental variables in the natural environment (Cooper et al., 2007; Shapiro & Kratochwill, 2000). These data can be collected using different formats, ranging from narrative recording to more structured interval recording procedures. Because direct FBA methods are used in the natural environment they are more flexible, do not disrupt the subject's normal routine, and provide unique information compared to experimental methods. For example, an

[1] Although EFAs involve direct observation of the target behavior, a quality of all direct FBA methods, they are distinct from other direct FBA methods in that they are the only method to experimentally manipulate variables (Hanley, 2012). Thus, in this paper, the discussion of direct FBA methods will exclude EFAs.

uninterrupted picture of the subject's environment may allow an observer to see if the existing contingencies are currently supporting or suppressing the problem behavior (Pence, Roscoe, Bourret, & Ahearn, 2009). Direct FBA methods are also more objective than indirect FBA methods as observers do not have to rely on informant recall.

Direct FBA methods are often employed by practitioners (Kern, Hilt, & Gresham, 2004; Love, Carr, Almason, & Pettursdottir, 2009), which can likely be attributed to the qualities listed above as well as the information provided by federal and state education agencies outlining the components of a FBA (Georgia Department of Education, 2016; U.S. Department of Education, 1999). In a review of 20 studies conducted in applied settings between 1991-2002, one of the most frequently employed FBA methods was direct observation in the natural environment, a necessary component of direct assessment (Kern et al., 2004). The majority of reviewed studies used a combination of interviews and direct observation strategies, with EFAs employed far less frequently. In a more recent survey of professionals employed at early and intensive behavioral intervention programs, Love et al. (2009) reported a similar practitioner preference for direct FBA methods over experimental procedures. Unfortunately, there are problems associated with data collected from direct FBA methods.

Although the purpose of direct FBA methods, and all FBA methods, is to obtain data from which an accurate hypothesis of function can be made and a function-based treatment can be created, there is considerable research to suggest that findings from direct FBA methods do not always converge with those of an EFA which is recognized as the gold standard for identifying behavioral function (Camp, Iwata, Hammond, & Bloom, 2009; Lerman & Iwata, 1993; Pence et al., 2009; Thompson & Iwata, 2007).There are several possibilities as to why the use of direct FBA methods might fail to result in an accurate hypothesis of behavioral function.

First, guidelines for the implementation and interpretation of direct FBA methods are lacking (Anderson, English, & Hedrick, 2006). Thus, the extent to which they might generate data that accurately reflect the relationship between behavior and the environment is questionable. Second, data from direct FBA methods might also fail to result in an accurate hypothesis of function if decisions are based on an inadequate sample of data (Rooker et al., 2015). This may be attributed to the limited time and resources afforded to observation in applied settings (such as a classroom), a lack of qualified personnel, or the fact that collecting data in an applied setting undoubtedly results in a high level of variability as it lacks the control of an analog setting. In an effort to obtain an adequate sample several researchers have attempted to standardize the observation procedures, proposing criteria for a minimum number of observations and/or recordings of the target behavior (McKerchar & Thompson, 2004; Thompson & Iwata, 2007). Another option for obtaining an adequate sample may be to collect data during times identified by a teacher as problematic, rather than randomly selecting an observation time or scheduling a time based on convenience (Tiger et al., 2013).

A third possibility as to why the use of direct FBA methods might fail to result in an accurate hypothesis of behavioral function is the tendency of such methods to yield false positive results when the behavior is maintained by intermittent reinforcement. For example, the assessment methods are designed to identify the variables which most frequently follow the target behavior, even if they have no reinforcing value (Cooper et al., 2007). Also, if a behavior is maintained by intermittent reinforcement in the natural environment, the correlation between the occurrence of the behavior and the delivery of the reinforcer maintaining it may be low (Camp et al., 2009; Lerman & Iwata, 1993). Lanovaz et al. (2013) suggest that individuals with extensive training in behavioral observation may be able to overcome this limitation of direct

FBA methods by taking into consideration context and other variables when analyzing the data (e.g., effects of intermittent reinforcement, attributing more weight to certain events).

Inadequate data analysis is yet another problem associated with direct FBA methods that might prevent the identification of behavioral function (Rooker et al., 2015). Data from direct FBA methods do not necessarily allow for analyses of the environmental conditions that are most often associated with problem behavior. Researchers have proposed a variety of analytic strategies in the hopes of better identifying contingent relations between behavior and its maintaining variables. One of the most common strategies is the conditional probability analysis, a strategy that provides the probability of an event (e.g., consequence such as attention) given the occurrence of another event (e.g., the target behavior) (Eckert, Martens, & DiGennaro, 2005; Thompson & Borrero, 2011). Although conditional probability analyses, as well as other advanced analyses, have been proposed in the literature (Martens, DiGennaro, Reed, Szczech, & Rosenthal, 2008; Thompson & Borrero, 2011), there are limitations to their use in applied settings. For example, Eckert et al. (2005) noted the difficulty in making predictions with low rate behaviors and the time added to the assessment process. The complexity of such analyses may also limit their use in applied settings (Thompson & Borrero, 2011).

Despite the limitations associated with direct FBA methods, there is considerable research highlighting their practical value when used in combination with other assessment methods such as EFAs (Fisher et al., 1998; Mace, Lalli & Lalli, 1991; McKerchar & Thompson, 2004; Tiger, Hanley, & Bessette, 2006) and as an alternative to experimental procedures when an EFA is contraindicated or impractical (Camp et al., 2009). For instance, in a set of best practice guidelines for conducting EFAs, Rooker et al. (2015) described the necessity of collecting data from direct FBA methods to inform the development of EFA conditions. One way this might be

accomplished is by using direct FBA data to identify naturally occurring schedules of reinforcement which can be incorporated into traditional EFAs (Mace et al., 1991). When an EFA produces inconclusive results it is possible that the contingency responsible for maintaining the problem behavior was not present in one of the test conditions. Therefore, it may be necessary to use direct FBA methods to identify idiosyncratic variables existing in the natural environment that are reinforcing the problem behavior (Fisher et al., 1998; Thompson & Borrero, 2011; Tiger et al., 2006).

Fisher, Adelinis, Thompson, Worsdell, and Zarcone (1998) attempted to identify the function of the problem behavior exhibited by two adolescents using an analog EFA. When the EFA resulted in near zero rates of target behavior across conditions the researchers employed A-B-C recording, a form of direct FBA method. Subsequent, more focused EFAs were then conducted that confirmed the hypothesis suggested by the data from the A-B-C recording. The treatment plan developed based on these findings successfully reduced the problem behavior of both participants. This study adds support to the utility of direct FBA methods in identifying unique antecedents and consequences that maintain problem behavior and using that information to create specific, more targeted experimental analyses.

Similarly, Tiger, Hanley, and Bessette (2006) conducted direct FBA methods after their initial EFA yielded inconclusive results. The researchers used a narrative A-B-C recording procedure to examine the participant's self-injurious behavior across three activities. Given that problem behavior occurred most often during naptime, stimuli from this event were incorporated into a second EFA. Elevated levels of problem behavior were observed across conditions after this modification, indicating the behavior was maintained by automatic reinforcement. Tiger et al. (2006) recommend an initial well-controlled EFA, followed by open-ended direct FBA

methods in the natural setting if the EFA yields undifferentiated results. Direct FBA methods can also be used as an alternative to EFAs in certain situations. When target behaviors are too challenging to examine using traditional EFAs (e.g., dangerous behaviors or behaviors that are difficult to control, low frequency behaviors), direct FBA methods may be the best alternative (Lerman, Hovanetz, Strobel, & Tetreault, 2009) as they can be used to identify "precursor behavior," or behavior that reliably precedes the occurrence of the problem behavior targeted for intervention (Camp et al., 2009; Sloman, 2010).

Other practical advantages of data from direct FBA methods include its ability to be used both as a baseline to measure treatment effects and as a way to generate effective treatments (Ellingson, Miltenberger, Galensky, & Garlinghouse, 2000; Mace et al., 1991; VanDerHeyden et al., 2001). VanDerHeyden, Witt, and Gatti (2001) used a direct FBA method to analyze the disruptive behavior exhibited by two early childhood classrooms. Data were recorded on two target child behaviors, one peer behavior, and 5 to 8 teacher behaviors using an interval recording procedure. Conditional probabilities were calculated and the data served as a baseline from which to measure the intervention effects. A treatment plan created from the results of the data from the direct FBA method effectively reduced the disruptive behavior in both classrooms. This study suggests that assessments conducted in the natural environment, utilizing more flexible procedures than EFAs, are capable of informing effective interventions. More research is needed to establish the treatment utility of direct FBA methods.

Out of the possible assessment methods included in FBAs, direct FBA methods are unique in that they provide the benefits of direct observation of the behavior and the possibility of identifying idiosyncratic variables maintaining behavior, all while remaining practical for use in an applied setting. As EFAs are not explicitly required by the law and have limitations when

implemented in schools (Blood & Neel, 2007; Bloom, Iwata, Fritz, Roscoe, & Carreau, 2011; Solnick & Ardoin, 2010), more research is needed to examine the accuracy of the assessment methods that are both legally required and feasible for school-based practitioners.

The first part of chapter 2 will provide an overview of FBAs conducted in school settings, touching upon the federal and state requirements and the ambiguity surrounding the process, as well as introduce some of the more common assessment methods (experimental, indirect, and direct) and examine their strengths and limitations in applied settings. The second half of the chapter will delve into the different types of direct FBA methods, describing research evaluating their quality using the criteria proposed by Hayes, Nelson, and Jarrett (1986). This section will focus on A-B-C recording. A-B-C recording is considered the most common direct FBA method (Anderson et al., 2006; Asmus et al., 2002) and one that is often employed in school settings (Eckert et al., 2005; VanDerHeyden et al., 2011) and school-based research (Anderson, Rodriguez, & Campbell, 2015). It is also the only direct FBA method with any empirical research evaluating its quality.

Chapter 3 will attempt to address several unanswered research questions uncovered in Chapter 2. As discussed above, there is a concern that the data obtained through direct FBA methods do not always agree with those obtained through EFAs. Before the conclusion is made that findings from direct FBA methods are not consistent with those of EFAs, research must first establish the accuracy of direct FBA methods. As A-B-C recording is the only direct FBA method with any research evaluating its quality, this method will be selected for the current study.

There are multiple A-B-C recording formats, each with their own benefits and disadvantages. A-B-C structured recording is a more accurate and preferred format relative to A-

B-C narrative recording (Lerman et al., 2009). With structured recording, observers record the occurrence of the problem behavior and place checkmarks next to a list of pre-identified antecedent and consequent variables occurring contiguous to the behavior. Despite its advantages, A-B-C structured recording is limited in that data are only collected when the target behavior occurs which restricts possible analyses and prevents an examination of the ecology of the classroom, in comparison to data collected in a continuous manner (regardless of behavior occurrence). As relevant antecedents and/or consequences may not be immediately apparent to the observer (Shapiro & Kratochwill, 2000), continuous recording ensures that an adequate representation of the environment is captured. Research is needed to explore the possibility of using a structured form of A-B-C recording in a continuous manner.

Studies have examined the accuracy of A-B-C recording methods collected by teachers (Lerman et al., 2009; Mayer & Reed, 2013); however, there are problems associated with teacher-collected direct FBA data. Not only is it unlikely that teachers have received adequate training in FBA (Couvillon et al., 2009), it is likely that the teachers themselves are providing the antecedents and/or consequences that they are responsible for recording. As self-recording of behavior has a reactive effect on behavior (Lipinski & Nelson, 1974), it is unlikely that the teacher-collected data using direct FBA methods yields an accurate picture of the environment. The primary purpose of Chapter 3 is to examine the accuracy of a structured form of A-B-C recording used in a continuous manner (the Serial Recording form), when collected by school psychologists currently practicing in the schools.

Indirect FBA methods are typically employed at the beginning of the FBA process (Floyd, Phaneuf, & Wilczynski, 2005). This creates the potential that an observer conducting the direct FBA methods is biased from their knowledge of the results from the indirect FBA

methods. For example, knowing that a rating scale completed by a classroom teacher suggests attention as a possible function of problem behavior may result in the observer looking for evidence to support that function and ignoring information that refutes the hypothesis, a phenomenon known as confirmation bias. The secondary purpose of Chapter 3 will be to examine the effect that an a priori hypothesis of behavioral function has on the A-B-C recordings of school psychologists, as well as on their selection of behavioral function.

CHAPTER 2

LITERATURE REVIEW

A functional behavior assessment (FBA) is a collection of assessment methods designed to identify the function of a behavior, or the reason why it occurs, by examining the surrounding environmental conditions (Gresham et al.,2001). FBA methods range from indirect methods, removed in time and place from the behavior, to direct methods that involve direct observation of the behavior in the natural environment (Cooper et al., 2007). In some instances, experimental manipulations of environmental variables (experimental functional analysis) are employed. Ideally, FBAs lead to individualized, function-based interventions that effectively decrease problem behavior and increase appropriate alternative behavior(s) (Cooper et al., 2007; Shriver et al., 2001).

Although FBAs have been considered best practice for decades (Gresham et al., 2001), prior to the 1990s FBAs were conducted almost exclusively in clinical settings with individuals with developmental disabilities (Anderson et al., 2015). FBAs are now mandated in certain circumstances under the Individuals with Disabilities Education Act (IDEA, 1997; 2004). Specifically, FBAs and behavior intervention plans (BIPs) are required when a child with a disability is removed from school for more than 10 days as a result of misconduct that was determined to be a manifestation of the student's disability (IDEA, 2004). Although the exact circumstance under which an FBA is required is narrow, it may be appropriate, or even best practice, to conduct an FBA if a child's behavior is impacting his or her ability to learn, regardless of disability status (von Ravensburg & Tobin, 2008).

Furthermore, a failure to address students' behavioral needs is a denial of their access to a free and appropriate public education, which is the principal tenant of IDEA (Poucher, 2016). The inclusion of FBAs in IDEA has increased their use in schools (Anderson et al., 2015) and has expanded their use with children with or at risk for emotional and behavioral disorders (Kern et al., 2004), as well as with typically developing children (Anderson et al., 2006). The widespread adoption of FBAs is not without problems. Despite the mandate, IDEA fails to provide a definition of FBA or a list of necessary components for its implementation in the school setting (Asmus et al., 2002; Stage et al., 2006; Zirkel, 2011). Without clear guidelines for implementation there is great potential for FBAs to be implemented incorrectly, thus decreasing the probability that resulting data will lead to effective, function-based interventions.

Although IDEA contains little information on the components that should be included as part of an FBA, the federal government provides some guidance on the essential components (U.S. Department of Education, 1999; National Association of State Directors of Special Education, 1998). In a synthesis of that guidance, Weber, Killu, Derby, and Barretto (2005) created a list of components that should be considered standard practice for FBA completion. These components include but are not limited to: a clear definition of the target behavior, interviews with individuals familiar with the target behavior, direct observation and behavioral recording techniques, and the development of a hypothesis on the function of the target behavior. Although Weber et al. (2005) did not anticipate that each component be included in every FBA, they stressed that state education agencies should, at a minimum, provide information to school districts on how to implement all of these components A more recent paper echoed the guidance provided by Weber and colleagues, stating that in the absence of a clear set of criteria special education experts recommend FBAs include an operational definition of the problem

behavior(s), descriptions of the environmental events occurring before and after the problem behavior, and direct observation (Zirkel, 2011).

In addition to identifying standard practice criteria for FBA completion, Weber et al. (2005) examined the FBA-related resources that are distributed to school districts through state education agencies and compared them with the information provided by the federal government. The authors reported that seven states failed to provide any information to their school districts on FBAs and many more failed to provide information on all of the standard practice components. When states provided schools with materials for conducting components of the FBA (e.g., data collection sheets, rating scales), the materials often lacked guidance on how to use them. One examination of the state statutes and regulations specific to FBAs found that only 17 state laws define FBAs, and those definitions lack specificity (Zirkel, 2011). Weber et al. (2005) suggested that most school districts are following a "cookbook approach" to conducting FBAs. That is, districts are instructed to complete certain steps in a particular sequence; however, state education agencies have failed to provide them with a scientific basis for conducting the steps. Without a scientific understanding of why each of the FBA components are necessary and how they work together there is little likelihood that schools will correctly conduct their FBAs, thus decreasing the possibility that FBA results will lead to an effective intervention.

An example of one state's failure to provide sufficient FBA-related information is found on the Georgia Department of Education (DOE) website. The Georgia DOE provides schools with basic information on FBAs and BIPs; however, the information is buried within a 228-page implementation manual ("Special Education Rules Implementation Manual," 2016). The language is surface-level and non-specific, leaving room for misinterpretation. The implementation manual provides a list of components an FBA *may* include (record reviews,

interviews, observations), but fails to describe (a) by whom and how often observational data should be collected, (b) exactly how data should be collected and analyzed or (c) how much data are necessary for generating a hypothesis regarding the function of a student's behavior. The implementation manual also fails to provide necessary tools for implementation (e.g., data sheets, training videos).

Further complicating the issue of insufficient information provided by states to their school districts is a finding from Couvillon, Bullock, and Gable (2009) that many teachers lack the necessary training in FBA-related procedures. After surveying 134 special education service providers, they found that a majority of respondents did not receive FBA-related training until after their fifth year of teaching. With evidence to suggest that information provided to school personnel is often insufficient, combined with the lack of training provided to school personnel, it seems unlikely that FBAs are being implemented as intended in schools.

Findings from the limited studies examining FBAs in schools suggest that there are multiple flaws with their current implementation (Blood & Neel, 2007; Van Acker, Boreson, Gable, & Potterton, 2005). Blood and Neel (2007) examined the files of children with emotional and behavioral difficulties in a large school district in Washington. They found that the majority of children who should have had an FBA did not have one on file. Furthermore, the FBAs that were conducted were seemingly inadequate, relying heavily on one type of assessment method that did not involve direct observation of the behavior. This research, although limited in scope, provides some evidence that educators are not implementing FBAs as intended.

In a similar study, Van Acker et al. (2005) reviewed FBAs submitted by service providers from 70 schools distributed across 21 school districts. A team of experts in behavior analysis compared the FBAs against what the authors deemed "best practice" according to the research

literature on FBAs. Approximately half of the submitted FBAs contained multiple, critical flaws—such as a failure to adequately define the target behavior and a failure to identify the function. As a result, Van Acker et al. (2005) concluded that schools are failing to conduct "legally defensible and technically adequate" FBAs (p. 54). Failure to conduct a thorough FBA can have serious legal implications, as evidenced by the number of due process hearings schools have been involved in over absent or inadequate FBAs since the IDEA amendments (Drasgow & Yell, 2001; Poucher, 2015).

Drasgow and Yell (2001) investigated FBA-related court cases between 1997 and 2000 and found that an inadequate FBA (one containing limited data sources) resulted in school districts losing 94% of the parent-initiated, state-level due process hearings. Other research has found inconsistency in the court rulings in FBA-related cases between parents and school districts (Poucher, 2016), likely due to the limited direction provided by the federal government and state departments of education as to what should be included. In the absence of a legal standard it is up to the courts to decide how to measure the adequacy of an FBA, who, in turn, often recruit persons with expert knowledge in the subject (Poucher, 2016). For example, a hearing officer may decide to contract an independent evaluator to conduct a thorough FBA and then compare the FBA in question to the FBA created by the expert evaluator (Drasgow & Yell, 2001).

Although there are no legal guidelines that specify the components which comprise a thorough FBA, Drasgow and Yell (2001) submitted a checklist for conducting legally defensible and educationally appropriate FBAs. The checklist specifies that an FBA should consist of: (a) interviews of individuals who are familiar with the behavior, (b) direct observations of the student in a variety of settings, (c) an experimental manipulation of environmental variables, if

necessary, and (d) a hypothesis of behavioral function. These checklist items are consistent with the standard practice components provided by Weber et al. (2005) and Zirkel (2011). Therefore, it is prudent to suggest that FBAs comprised of these components are reflective of best practice and are comprehensive and individualized enough to yield information that can be translated into effective interventions (Drasgow & Yell, 2001).

The remainder of the paper will discuss components of FBAs beginning with a brief discussion of experimental functional analyses (EFA) and the potential difficulties with conducting EFAs in an applied setting. This will be followed by a discussion of indirect and direct FBA methods, which are typically conducted in school settings.

Experimental Functional Analyses (EFA)

EFA involves the systematic manipulation of variables (e.g., attention, access to tangibles) hypothesized to be controlling the target behavior, while observing and recording the subsequent effect on behavior (Fisher et al., 1998; Iwata et al., 1982; Rooker, Iwata, Harper, Fahmie, & Camp, 2011). EFAs are most often implemented in analog settings where the manipulation of environmental variables can be tightly controlled (Alberto & Troutman, 2006; Anderson et al., 2015; Cooper et al., 2007; Scott et al., 2004). Experimental functional analysis is the only method that allows practitioners to confirm hypotheses regarding the function of behavior (Asmus et al., 2002; Cooper et al., 2007) and research has demonstrated that the largest reductions in problem behavior are observed when an EFA is included in the FBA process (Hanley et al., 2014). Researchers have proposed several reasons for the lack of EFAs in applied settings (Asmus et al., 2002; Calloway & Simpson, 1998; Gable et al., 2014).

One explanation for the lack of EFAs in applied settings is the dearth of individuals trained in data collection, implementation, and analysis (Asmus et al., 2002). A second

explanation for their absence in applied settings is teacher perception that the process is time-consuming and difficult (Asmus et al., 2002; Calloway & Simpson, 1998; Hanley, 2012). Time is a valuable and limited resource in schools and EFAs require considerable time for training, implementation, and analysis (Anderson et al., 2015). If an assessment method is not practical or acceptable it is unlikely to be implemented as intended (Floyd et al., 2005; Scott et al., 2004). Thus, FBA methods must be weighed not just for their efficacy, but also for their practicality. A third possibility as to why EFAs are not conducted more frequently in applied settings is that teachers may be hesitant to intentionally reinforce the inappropriate behavior of children who are high-functioning (Kern et al., 2004), and unwilling to evoke potentially dangerous or highly disruptive behaviors in the classroom (Hanley, 2012; Lanovaz et al.ns, 2013).

In addition to the difficulties associated with conducting EFAs within schools, which leads to their infrequent implementation, the lack of control provides yet another limitation to their use in applied settings (Bloom et al., 2011; Tiger et al., 2006). There are many variables within a classroom that are difficult to control (e.g., attention from peers, interruptions in the form of announcements or fire drills). If the presentation of the environmental variables is not tightly controlled the data will be invalid (Tiger et al., 2006). A method of assessment may be appropriate in a laboratory setting when implemented by trained professionals, but ineffective when introduced in a classroom by a teacher who has limited to no knowledge regarding behavioral assessment. Despite being the only FBA method that can validate the function of behavior, concerns over the feasibility of implementing EFAs in applied settings have schools turning to indirect and direct FBA methods.

Indirect FBA Methods

Indirect FBA methods are removed in time and place from the behavior of interest and are typically conducted at the outset of the assessment process (Floyd et al., 2005). Generally, these FBA methods are employed for the purpose of gathering information about the student behavior and the environmental events (antecedents and consequences) surrounding the behavior (Alberto & Troutman, 2006; Cooper et al., 2007; Gresham et al., 2001). Examples of indirect FBA methods include behavioral interviews, behavioral rating scales such as the Motivation Assessment Scale (Durand & Crimmins, 1988) and the Questions About Behavioral Function (Matson & Vollmer, 1995), and record reviews. Both behavioral interviews and rating scales are generally administered by a practitioner to a caregiver or teacher who is highly familiar with the student and the behavior of interest.

Behavioral interviews are typically semi-structured and are intended to identify the target behavior, obtain comprehensive information about the settings and events surrounding the target behavior, and identify an appropriate behavior to replace the inappropriate behavior (Alberto & Troutman, 2006; Gresham et al., 2001). Behavior rating scales commonly involve ranking behaviors on a Likert-type scale according to their frequency of occurrence and/or severity, with the purpose of identifying the function of the target behavior. Each item on the rating scale corresponds to a particular function (e.g., attention, escape, tangible, automatic) and the function with the highest rating after the items are summed is hypothesized to be the controlling variable. Finally, record reviews involve a comprehensive review of the student's school records (e.g., attendance, work samples, test scores, office referrals, previous FBAs). These three assessment methods are considered "indirect" because the information gathered is based solely on informant recollection and opinion and does not involve direct observation of student behavior.

FBA models proposed by researchers generally recommend indirect FBA methods be used early in the assessment process to: (a) help formulate hypotheses regarding behavioral function, (b) create operational definitions of behavior, and (c) identify appropriate times for observations (Floyd et al., 2005; Stage et al., 2006). Due to their ease of use, relatively short administration time, and the lack of training required, schools tend to rely heavily on indirect FBA methods (Floyd et al., 2005; Gable et al., 2014). In a review of the existing research and instructional materials, Allday, Nelson, and Russel (2011) found that structured interviews were the most common form of data collection conducted and recommended out of the various indirect and direct FBA methods.

Despite their widespread use and potential to assist in identifying the cause(s) of problem behavior, there are problems associated with indirect FBA methods that must be considered. First, data collected through indirect FBA methods are based upon subjective recall by caregivers that is prone to bias and relies heavily on the informant's ability to accurately identify and describe the behavior and the environmental events occurring contiguous to the behavior (Alberto & Troutman, 2006; McIntosh et al., 2008). Another problem with indirect FBA methods is that behavior can vary as a function of the setting or the individual(s) present, which can lead to discrepant reports across informants (Floyd et al., 2005).

A third, significant problem associated with indirect FBA methods is with their psychometric properties (McIntosh et al., 2008; Matson & Minshawi, 2007; Stage et al., 2006). The research reports low levels of reliability and validity (Alberto & Troutman, 2006; Barton-Arwood, Wehby, Gunter, & Lane, 2003; Koristas & Iacono, 2013) and has generally involved small sample sizes, which limits the legitimacy of the findings (Kozlowski & Matson, 2012). Furthermore, the research mostly employs research designs and statistical analyses that ignore

the principle tenant of behavioral assessment--that consistency in behavior across time or settings cannot be assumed. Researchers have, unfortunately, ignored this tenant of behavioral assessment in an attempt to evaluate their measures from a traditional psychometric perspective (Koristas & Iacono, 2013; Paclawskyj, Matson, Rush, Smalls, & Vollmer, 2000; Shogren & Rojahn, 2003). Traditional psychometric theory assumes the causes of behavior reside within the individual and are stable over time and across contexts; whereas in behavioral assessment, behavior is viewed as a function of the interaction between a person and his current environment (Hayes et al., 1986). Although there are exceptions, in general it is not appropriate to apply traditional psychometric procedures to behavioral assessment data (Gresham et al., 2001) as variability may be a function of true changes in the organism's behavior, not a flaw in the instrument itself.

Another problem with indirect FBA methods is the lack of data supporting the treatment utility of indirect FBA methods. In a comprehensive review of indirect FBA methods, Floyd et al. (2005) found that only one study examined the isolated effects of these methods on treatment outcomes. This is a significant problem as the degree to which assessment methods inform treatment is one of the primary ways to evaluate the efficacy of behavioral assessment (Hayes et al., 1986). A final problem associated with indirect FBA methods is that research suggests low correspondence between the functions identified through indirect FBA methods and those identified using EFA (Rooker et al., 2015). This is problematic as EFAs are the only assessment method that can claim to identify the true cause of behavior (Cooper et al., 2007; Pence et al., 2009; Rooker et al., 2015).

Due to the problems noted above, the literature has cautioned against exclusive reliance on indirect FBA methods for identifying the function of behavior (Gresham et al., 2001; Iwata,

DeLeon, & Roscoe, 2013; Stage et al., 2006). Evidence suggests indirect FBA methods are most effective when using multiple methods and multiple informants who are highly familiar with the student and behavior (Gable et al., 2014; Kozlowski & Matson, 2012). Given the questionable reliability and validity of these data, indirect FBA methods also have the potential to bias an individual collecting these data prior to conducting a direct FBA method. As data resulting from indirect FBA methods are based on informant recall rather than direct observation they may be most appropriate for less severe and less complex forms of behavior.

Direct FBA Methods

Due to the disputed feasibility of EFAs in applied settings and the numerous limitations associated with indirect FBA methods, direct, or descriptive, assessments provide a possible compromise[2]. These assessment methods employ unobtrusive, direct observation of student behavior in the environment in which it typically occurs (Cooper et al., 2007; Shapiro & Kratochwill, 2000). This allows one to obtain a complete, objective sample of the contingencies operating in the natural environment with more flexibility than experimental methods allow (Floyd et al., 2005; Gresham et al., 2001). Direct FBA methods are frequently employed to create operational definitions of behavior and to generate hypotheses of behavioral function (Shapiro & Kratochwill, 2000; Sloman, 2010). They are used to identify naturally occurring schedules of reinforcement and to determine if the existing environment is supporting or suppressing problem behavior—both of which can assist with treatment planning (Mace et al., 1991; Pence et al., 2009). The quality of behavioral assessment methods is largely determined by their ability to generate effective treatments (Gresham et al., 2001; Hayes et al., 1986; Nelson &

[2] Although EFAs involve direct observation of the target behavior, a quality of all direct FBA methods, they are distinct from other direct FBA methods in that they are the only method to experimentally manipulate variables (Hanley, 2012). Thus, in this paper, the discussion of direct FBA methods will exclude EFAs.

Hayes, 1979) and data from direct FBA methods can be used to reduce problem behavior through the development of function-based treatment plans (Anderson et al., 2006; Arndorfer, Miltenberger, Woster, Rortvedt, & Gaffaney, 1994; Ellingson et al., 2000; VanDerHeyden et al., 2001).

Another use for direct FBA methods is to inform EFAs (Fisher et al., 1998; Mace & Lalli, 1991; McKerchar & Thompson, 2004; Tiger et al., 2006). For example, direct FBA methods are often used to identify idiosyncratic variables operating in the natural environment that might not be identified by traditional experimental analysis procedures (Fisher et al., 1998; Thompson & Borrero, 2011; Tiger et al., 2006). These variables may then be incorporated into EFAs. For example, Tiger, Hanley, and Bessette (2006) conducted A-B-C narrative recording, a form of direct FBA method, following inconclusive results obtained from the original EFA. The researchers examined the participant's severe self-injurious behavior across three different activities: centers, free choice, and naptime. The problem behavior (hand-mouthing) occurred most often during naptime, so stimuli from this activity were used in the subsequent EFA. The researchers observed elevated levels of hand-mouthing across conditions after this modification, indicating the behavior was maintained by automatic reinforcement.

Direct FBA methods can also serve to validate EFAs by examining whether traditional EFA conditions accurately simulate the contingencies present in the natural environment (McKerchar & Thompson, 2004; Thompson & Iwata, 2007). For example, McKerchar and Thompson (2004) used direct FBA methods to determine whether the consequences typically presented in EFAs were present in the preschool environment. The researchers conducted observations across a range of activities, employing an interval recording procedure to capture

the occurrences of the target behavior and environmental events. Attention was found to be the most common consequence following problem behavior.

When target behaviors are dangerous or occur infrequently, direct FBA methods may serve as a good alternative to EFAs (Camp et al., 2009; Lerman et al., 2009). Direct FBA methods can be used to identify a precursor behavior to be reinforced in an EFA instead of a more severe topography, or behavior that reliably precedes the occurrence of the target behavior (Camp et al., 2009; Sloman, 2010). Due to their considerable utility, both alone and in combination with EFAs, direct FBA methods enjoy much popularity amongst school-based researchers (Anderson et al., 2015) and practitioners (Kern et al., 2004; Love et al., 2009). In a survey of 211 professional supervisors of early and intensive behavioral intervention programs, Love et al. (2009) found that 96% of professionals used direct FBA methods when assessing problem behavior, whereas only 56% used experimental procedures.

Despite the popularity and utility of direct FBA methods, there are problems surrounding their use. First, direct FBA methods are only capable of identifying correlations between behavior and environmental variables, not causal relationships (Rooker et al., 2015; Sloman, 2010). A second problem is the considerable research suggesting a discrepancy between the functions identified through EFAs and direct FBA methods (Camp et al., 2009; Lerman & Iwata, 1993; Pence et al., 2009; Thompson & Borrero, 2011; Thompson & Iwata, 2007). There are several possibilities as to why direct FBA methods may fail to produce an accurate hypothesis of the function of behavior.

One possibility is the lack of clear guidelines for their implementation and interpretation (Anderson et al., 2006). An inaccurate hypothesis of behavioral function may also result from basing decisions on an inadequate sample of data (Rooker et al., 2015). This may be attributed to

the limited time and resources allotted for observations in applied settings, the variability that comes with collecting data in uncontrolled settings, or simply a lack of consensus around the amount of data necessary for decision-making (Rooker et al., 2015). In an attempt to increase the adequacy of the data samples, some researchers recommend collecting data during teacher-identified times rather than scheduling observations based on convenience (Tiger et al., 2013). Others have proposed guidelines for a minimum number of observations (Asmus et al., 2002) and have imposed stringent requirements for data collection in their own research studies (Thompson & Iwata, 2007) that are referenced as examples to strive towards. Although research suggests that increasing both the length of observations and the number of observations per week can help ensure an adequate behavioral sample (Hintze & Matthews, 2004; Tiger et al., 2013), these are not always viable solutions for busy practitioners.

Discrepant findings between EFAs and direct FBA methods may also occur if the environmental variables maintaining the behavior are absent across the observation periods. This may be due to the uncontrolled nature of direct observations or the lean schedules of reinforcement that typically operate in the natural environment (Rooker et al., 2015; Sloman, 2010). It is equally possible that the specific antecedent event or consequence maintaining the behavior in the natural environment (e.g., the presence of a particular peer or removal of a specific task) is not present in the EFA test condition. Structured direct assessments (SDAs) attempt to address some of these flaws inherent in direct FBA methods by systematically presenting specific antecedent events that are similar to those seen in EFAs (e.g., presentations of demands, withholding of attention; Anderson & Long, 2002; Anderson et al., 2006). This activity increases the possibility of a relevant contingency being presented within an observation period. Consequences are not manipulated—that is, caregivers are encouraged to respond to the behavior

as they would under normal circumstances. The assessment is conducted in the natural environment and natural intervention agents (e.g., teachers, parents) interact with the students. Interestingly, in many of the studies that reported discrepant results between EFAs and direct FBA methods, the EFA was conducted in a therapy room outside of the child's natural environment with a novel therapist and/or materials (Camp et al., 2009; Lerman & Iwata, 1993; Pence et al., 2009; Thompson & Iwata, 2007). In contrast, many of the studies reporting correspondence between the two methods conducted the EFA in the child's classroom or home, with a familiar intervention agent running the sessions (Alter et al., 2008; Lanovaz et al., 2013; Sasso et al., 1992).

The lack of sophisticated data analysis associated with direct FBA methods might also contribute to the misidentification of behavioral function (Rooker et al., 2015). Direct FBA data do not necessarily allow for analyses of the environmental conditions that are most often associated with problem behavior. The calculation of conditional probabilities has been proposed as a possible solution for better identifying functional relations using direct FBA data (Martens et al., 2008; Rooker et al., 2015). Conditional probabilities provide the probability of an event (e.g., consequence such as attention) given the occurrence of another event (e.g., the target behavior) (Thompson & Borrero, 2011). To increase the likelihood of identifying a contingency between environment and behavior, the conditional probability can be compared to the background probability of the same event (Rooker et al., 2015). If the conditional probability of an event is higher than the background probability of the same event, a positive contingency exists, and a potential function is identified. If the conditional probability of an event is not higher than the background probability it should not be assumed to be maintaining the behavior--it is simply presented freely and often in the natural environment. In practice, this means the continuous

recording of the presence of environmental variables, along with instances of both target and appropriate behavior. There are limitations associated with the use of conditional probability analysis and its various iterations in applied settings, such as their complexity and the added time to the assessment process (Eckert et al., 2005; Martens et al., 2008; Thompson & Borrero, 2011).

The remainder of this chapter will review the research on the most common direct FBA methods, addressing their strengths and limitations in regards to research support and feasibility in an applied setting. The focus will be on direct FBA methods as they have the benefit of direct observation of the behavior and the possibility of identifying idiosyncratic variables maintaining behavior, all while still being practical for use in an applied setting and frequently employed in schools (Anderson et al., 2015; Kern et al., 2004; Love et al., 2009). As EFAs are not explicitly required by the law and are rarely conducted in the school setting (Blood & Neel, 2007; Doggett, Edwards, Moore, Tingstrom, & Wilczynski, 2001), more research is needed to examine the accuracy of the assessment methods that are both legally required and frequently used in applied settings.

A second aim of this chapter is to examine research evaluating the quality of some of the most popular direct FBA methods. Hayes et al. (1986) assert that it is inappropriate to evaluate the quality of behavioral assessments the same way that traditional assessments are evaluated. Psychometric theory assumes stability in the quality or trait being measured, so measures used to evaluate traditional psychometric assessments seek consistency in behavior. This violates one of the major assumptions inherent in behavioral assessment—namely that behavior is variable, and inconsistency across time, setting, and people should be expected, rather than viewed as measurement error. As an alternative to traditional methods of evaluation, Hayes et al. (1986)

propose that behavioral assessments should be evaluated against three criteria: treatment utility, accuracy, and sensitivity.

The most commonly cited direct FBA methods include anecdotal reports (Alberto & Troutman, 2006), scatterplot analysis (Touchette, MacDonald, & Langer, 1985), and Antecedent-Behavior-Consequence recording (or A-B-C recording) (Bijou, Peterson, & Ault, 1968), of which there are different formats. Although these behavioral assessment methods are included in practitioner-oriented textbooks (Alberto & Troutman, 2006; Chandler & Dahlquist, 2010; Cooper et al., 2007) and used in applied settings (Allday et al. 2011; Anderson et al., 2015), there is limited research evaluating their treatment utility, accuracy, and sensitivity.

Anecdotal reports

Anecdotal reports involve an observer recording, in prose, a complete description of a student's behavior and the context in which it occurs (Alberto & Troutman, 2006). Contrary to other direct FBA methods, a target behavior is not identified before the observation. Rather, a target behavior is thought to emerge out of this data collection process. For example, a teacher may identify math class as a particularly challenging time for a student. An observer then writes a running record of all events associated with the target student occurring during that class period, in the hopes of identifying a specific behavior to be operationalized and targeted for further data collection and intervention. Observers note the time of the observation, and ideally collect data across several days.

A variation of anecdotal reporting is behavioral stream data collection, which involves the observer recording a complete sequence of events as the behavioral incident unfolds (Watson & Steege, 2003). The limitations associated with this method include the considerable amount of irrelevant information observers are required to sift through when searching for relevant

information and the difficulty in organizing the data into a structure amenable to analysis. Anecdotal recording has a lack of empirical support, with no evidence that this behavioral assessment method has treatment utility, or is accurate or sensitive.

Scatterplot

Another commonly employed and recommended direct FBA method is the scatterplot (Touchette et al., 1985). Scatterplots are designed to detect temporal patterns of behavior and then examine the environmental conditions occurring during those periods of behavioral escalation (Cooper et al., 2007). First described by Touchette, MacDonald, and Langer (1985), an observer creates a grid with units of time (usually 30 min) listed vertically and days of the week (or activities within one day) listed horizontally (Kahng, Iwata, Fisher, Williams, & Smith, 1998; Sloman, 2010). When the target behavior occurs, the observer marks the relevant square on the grid. In some instances, codes are created to denote whether the behavior did not occur, or occurred with high or low frequency (Thompson & Borrero, 2011). The data produced from scatterplots may be useful in isolating times of day to conduct more systematic observations, or for identifying a time period for intervention activities. In addition, the method is popular with practitioners due to the limited training, time, and resources involved (Thompson & Borrero, 2011).

Scatterplots are used more frequently in applied settings than in the research and there is limited research examining their ability to identify behavioral function (Kahng et al., 1998; Sloman, 2010). Kahng et al. (1998) investigated the utility of scatterplots in identifying the function of behavior using data gathered from 15 participants. Employing visual analysis, trained behavior analysts were unable to detect any reliable distribution of behavior across time. Interestingly, when the same data were re-examined using a statistical procedure, temporal

patterns of behavior were identified for 12 of the participants. Although it is possible advanced statistical methods might make it easier to detect temporal patterns of behavior using scatterplots, these are unlikely to be employed in applied settings. This limits their utility. Scatterplots are also limited in that they do not provide any information on the specific environmental variables that may evoke or maintain the problem behavior and there could be multiple variables contributing to an increase in behavior during a specific time interval (Gresham et al., 2001; Thompson & Borrero, 2011). A search of the literature did not reveal any empirical research evaluating the treatment utility, accuracy, or sensitivity of scatterplots.

A-B-C Recording

As early as 1968, Bijou, Peterson, and Ault described a recording system during which an individual observes behavior as it occurs in the natural environment and records the sequence of events surrounding the behavior. This seminal paper set the stage for what is now referred to as A-B-C recording (Sloman, 2010). In their book, Shapiro and Kratochwill (2000) recommend also gathering information on the relevant ecological conditions when collecting A-B-C datathat may include the time the behavior occurred, the assigned tasks or activities during the observation period, the contingencies in place (classroom rules, expectations, etc.) when the behavior occurs, the academic content area, and the format of instructional delivery (individual, small group, whole class).

A-B-C recording is identified as the most common form of direct FBA method (Anderson et al., 2006; Asmus et al., 2002) and one which is often employed in school settings (Eckert et al., 2005; VanDerHeyden et al., 2011) and school-based research (Anderson et al., 2015). In their review of 233 articles describing school-based FBAs Anderson et al. (2015) found that of the articles including direct FBA methods, 89% included A-B-C recording. A-B-C

recording has many benefits which contribute to its popularity in the school. For one, a complete sample of the environment may help to identify contingencies operating under intermittent or delayed reinforcement that are often difficult to detect using direct FBA methods (Shapiro & Kratochwill, 2000). A-B-C recording also provides the benefit of real-time recording of the behavior as it occurs in nature, with no disruption to the subject's normal routine (Cooper et al., 2007). Another benefit is the variety of formats in which the data can be collected. After recording behavioral sequences the relationships between the target behavior and the events surrounding it can be analyzed in a variety of ways, depending upon the format in which the data are collected. There are three primary formats for collecting A-B-C data: narrative, structured or checklist, and continuous.

Narrative. A-B-C narrative recording requires that observers write down the behaviors in narrative form, as well as the relevant antecedents and consequences, providing both topographical and sequential information (Shapiro & Kratochwill, 2000; Thompson & Borrero, 2011). In narrative recording data are typically only collected when the target behavior occurs (Cooper et al., 2007; Lerman & Iwata, 1993; Mayer & Reed, 2013). Data collected using this method are often used to: (a) confirm the existence of a problem, (b) create operational definitions of the target behaviors, (c) develop empirical recording systems, (d) develop other direct observation procedures for future recording, and (e) identify any environmental variables that may be maintaining the target behaviors (Shapiro & Kratochwill, 2000).

There are many benefits associated with A-B-C narrative recording. The data can provide detail and rich contextual information (Thompson & Borrero, 2011) often needed when writing operational definitions, problem analysis, and generating treatment plans. The method can be especially valuable for novel behavior that is difficult to define through caregiver report and

rating scale items prior to observation and for detecting shifts in behavior over time, such as changes in intensity or topograpy (Cooper et al., 2007; Thompson & Borrero, 2011; Watson & Steege, 2003). By writing down observations, as opposed to using predetermined codes (e.g., A-B-C continuous recording), one can provide more detailed information on various dimensions of the behavior including its frequency, duration, and intensity. Narrative recording is also ideally suited for detecting idiosyncratic variables (Lerman et al., 2009; Thompson & Borrero, 2011). For example, one might find that only certain types of demands are likely to evoke a behavior, or that attention from a peer may be more desirable than from a teacher. Furthermore, its open-ended nature allows one to create a highly individualized EFA (Tiger et al., 2006).

A-B-C narrative recording is popular in school settings, with 50% of practitioners reporting that they use it with most of their cases (Thompson & Borrero, 2011). The procedure is relatively easy to learn and implement (Lerman & Iwata, 1993; Thompson & Borrero, 2011) as less training is required and there are fewer recording rules. It is resource-efficient and yields an extensive amount of specific, contextual data. Although it may be more time-consuming to complete A-B-C narrative recording in the moment, there is less set up beforehand as observers do not need to create codes for all possible antecedents and consequences. These benefits likely contribute to its popularity in applied settings. Examples of narrative recording forms are described (Alberto & Troutman, 2006; Thompson & Borrero, 2011) and provided in many practitioner-oriented textbooks (Chandler & Dahlquist, 2010; Cipani & Schock, 2007; McDougal, Chafouleas, & Waterman, 2006; Watson & Steege, 2003). Even though the literature suggests that many professionals use A-B-C narrative recording in their practice (Lerman et al., 2009; Thompson & Borrero, 2011), limited research exists examining the validity of the method in identifying behavioral function or the accuracy of the recording procedure. Some studies have

found agreement between the results of A-B-C narrative recording and EFA results (Alter et al., 2008; Lanovaz et al., 2013; Sasso et al., 1992); however, these studies should be interpreted with caution due to the small number of participants and procedural variations which may affect the validity of the results. For example, Alter et al. (2008) did not calculate interobserver agreement on the narrative recordings and the participants' target behavior was broadly defined in the hopes of capturing more occurrences.

Despite the positive aspects, there are problems associated with the use of narrative recording. First, the data are more susceptible to the subjective impressions of observers and untrained professionals may use emotional, inferential language that lacks precision and objectivity (Sloman, 2010; Thompson, & Borrero, 2011). Second, the output (detailed, written observations) is difficult to transform into quantifiable units of analysis, which limits the possible analyses (Lerman et al., 2009;Thompson & Borrero, 2011). Third, with this type of recording process environmental events are traditionally only recorded in relation to the target behavior (Cooper et al., 2007; Lanovaz et al., 2013; Lerman et al., 2009). This suggests there is no way to report the frequency with which the antecedents and consequences occur independent of target behavior (Pence et al., 2009), preventing more complex analyses such as the calculation of conditional probabilities (Cooper et al., 2007; Sloman, 2010). For example, one would be unable to report the probability of the target behavior given the occurrence of some antecedent event using this method. Fourth, the recording procedure is discontinuous, as observers must pause to write down their observations, which increases the chance that an important event be missed while the observer looks down to write. Finally, the writing skills of the observer also have the potential to impact the quality of the observations.

Structured/Checklist. A second type of A-B-C recording is called checklist (Thompson & Borrero, 2011) or structured (Lerman et al., 2009) recording. Observers record the occurrence of the problem behavior and place checkmarks next to a list of pre-identified antecedent and consequent variables occurring contiguous to the behavior. As with A-B-C narrative recording, structured recording can be implemented with little training and in the initial phases of hypothesis development (Thompson & Borrero, 2011). Given that this method requires a fixed list of antecedents and consequences prior to observations it may be most useful when measuring predictable, uncomplicated behavior. A-B-C structured recording does not provide the level of detail obtained through narrative recording; however, the data are considered more objective (Lerman et al., 2009; Thompson & Borrero, 2011). It seems likely that interobserver agreement is more easily calculated with this method as the data are in quantifiable units. As with A-B-C narrative recording, A-B-C structured recording is limited in that it is traditionally collected in a discontinuous manner (Lerman et al., 2009). Data are also collected contingent on the occurrence of the target behavior, which limits data analysis (Thompson & Borrero, 2011).

Continuous. The final type of A-B-C recording is called continuous recording (Cooper et al., 2007), sometimes referred to as empirical recording, frequency, interval, or time-sample recording (Thompson & Borrero, 2011) or descriptive analysis (Camp et al., 2009; McKerchar & Thompson, 2004; Thompson & Iwata, 2007). A-B-C continuous recording uses codes for behavior, antecedents, and consequences, the definitions of which are created from data collected through indirect FBA methods or A-B-C narrative recordings (Alberto & Troutman, 2006; Cooper et al., 2007). Data collectors employ an interval (time-sampling) or event recording procedure to gather a sample of the events occurring in the natural environment within a set timeframe. When conducting A-B-C continuous recording the observer does not need to pause

his observation to record an event. The observer has the option of collecting data independent of the target behavior or only recording when the behavior occurs. The former is used frequently in research (Eckert et al., 2005; Lerman et al., 1993; Thompson & Iwata, 2007), as recording the independent occurrence of events and behavior allows for more sophisticated analyses to determine the probability of the behavior given a specific antecedent, consequence, or both. The data produced from A-B-C continuous recordings are easily quantified, lending themselves to various analytical procedures and the calculation of interobserver agreement (Skinner, Dittmer, & Howell, 2000; Thompson & Borrero, 2011). The data are also well-suited for monitoring intervention effects over time (McDougal et al., 2006).

A-B-C continuous recording possesses several limitations for use in applied settings. First, the procedure can be difficult to implement (Thompson & Borrero, 2011) as observers must be capable of quickly identifying and classifying observed events, which can be a challenge when a specific event may take many different forms (Lerman et al., 2009). For example, the consequence of "attention" may be both a cross look from a teacher and laughter from a peer. Second, the recording codes are typically developed from data gleaned from indirect FBA methods or direct observations (Cooper et al., 2007), which can be time-intensive. It should be noted that these two limitations are also associated with A-B-C structured recording. Third, the recording process often requires extensive training and the use of specialized equipment not always available in schools (Tarbox et al., 2009). Fourth, although it is possible to conduct various analyses with the data produced from these assessments (e.g., conditional probability analysis, lag sequential analysis, contingency space analysis), such complex analyses are not always feasible for busy practitioners with limited training in behavioral assessment (Thompson & Borrero, 2011). Finally, even with systematic data collection and various analytical procedures

A-B-C continuous recording data do not consistently identify the same function as EFAs, which is commonly accepted as the gold standard for identifying the function of behavior (Lerman & Iwata, 1993; Thompson & Iwata, 2007). Therefore, A-B-C continuous recording should not be used as a standalone method for determining behavioral function.

Terminology. It should be noted that there is inconsistency in the terminology associated with A-B-C recording. Although there is considerable variation in the different formats, the research does not consistently use the specifiers (e.g., narrative, continuous). In some of the literature on FBAs, A-B-C recording is only defined by the narrative format (Scott, Liaupsin, Nelson, & Jolivette, 2003; Shippen, Simpson, & Crites, 2003). In the empirical literature, A-B-C continuous recording may be referred to as descriptive assessment or descriptive analysis (Tarbox et al., 2009; Vollmer, Borrero, Wright, Van Camp, & Lalli, 2001) or conditional probability descriptive analysis (Pence et al., 2009). Search terms of "A-B-C recording," both alone and with the format specifiers, yield limited results, suggesting the assessment method is either rarely used in research, or falls under another descriptor (descriptive analysis, conditional probability, etc.). Inconsistent terminology creates a challenge for practitioners who are attempting to learn more about the assessment methods and for researchers seeking to evaluate the quality of different methods of behavioral assessment.

Evaluating the Quality of Direct FBA Methods

Out of the direct FBA methods discussed, A-B-C recording is the only method with any empirical research evaluating its quality using the criteria proposed by Hayes et al. (1986): Treatment utility, accuracy, and sensitivity.

Treatment utility. Treatment utility is the extent to which an assessment contributes to treatment effectiveness and/or efficiency. Research used to evaluate treatment utility should

examine whether the findings produced by the assessment method in question leads to a more effective treatment than the data produced by another method, or by a contraindicated treatment. A review of the research yields two studies examining the treatment utility of A-B-C recording procedures (Ellingson et al., 2000; VanDerHeyden et al., 2001). VanDerHeyden et al. (2001) employed A-B-C continuous recordings to examine the disruptive behavior exhibited in two early childhood classrooms and identify variables maintaining the behavior. Two treatment plans were created and compared using an alternating treatments design: one of which was based on the hypothesized function identified by A-B-C recording (i.e., withdrawing attention for disruptive behavior and praising appropriate behavior), the other of which was a contraindicated treatment based on the hypothesized function (i.e., providing attention following instances of disruptive behavior). The treatment plan based on the results of the A-B-C continuous recording data produced the greatest reductions in behavior across both classrooms.

Ellingson et al. (2000) used A-B-C structured recording to develop hypotheses about the functions maintaining the problem behavior of three students with intellectual disabilities. The authors used a brief reversal experimental design to validate the hypotheses generated by the data. Each student was exposed to two treatment conditions: (a) an intervention developed based on the hypothesized function and (b) a typical classroom intervention for problem behavior that was not based on the hypothesized function. The results suggested that the function-based intervention was more effective than the non-function-based intervention for two of the three participants. These two studies suggest that assessments conducted in the natural environment utilizing more flexible procedures than EFAs, are capable of informing effective interventions. Although the results from these studies are promising, more research is needed to support the treatment utility of direct FBA methods in general and A-B-C recording specifically.

Accuracy. The accuracy of an assessment method reflects the extent to which the data obtained from the method reflect the true state of nature. This requires a standard that data must be compared against, which does not usually exist outside of the research setting. Interobserver agreement is often used as a proxy for accuracy in research based on the argument that if data obtained from two independent observers agree, then the data are accurate. However, even if two data sets are identical there is no way to be certain that these data reflect what actually occurred.

Research has examined the accuracy of A-B-C recording (Lerman et al. 2009; Mayer & Reed, 2013). Lerman et al. (2009) compared the accuracy of teacher-collected A-B-C data using both narrative and structured recording methods. The authors created two versions of a 15-min video that had actors engage in a set number of problem behaviors, with clear antecedents and consequences accompanying each instance of behavior. As the videos were scripted by the authors, there was a gold standard document to which all observations could be compared. Prior to watching the videos, the teachers were exposed to an hour-long lecture on behavioral functions and data collection procedures and interpretation typical of the professional development provided by school systems. The participants were then split into two groups to watch the videos. The first group used A-B-C narrative recording with the first video, then structured recording with the second video. The second group completed the procedures in the reverse order.

To evaluate the accuracy of the structured recording format, teacher-completed forms were compared to the gold standard form created from the scripts. Three types of agreement were calculated to ensure the events immediately preceding and following each instance of the problem behavior were scored accurately: occurrence agreement of the antecedents and consequences, nonoccurrence agreement of the antecedents and consequences and the percentage

of agreement on antecedents and consequences for each instance of problem behavior. The authors also attempted to evaluate the accuracy of A-B-C narrative recording; however, the nature of the format (freehand, written descriptions) prevented the possibility of creating the same gold standard document generated for the checklist format. Instead, Lerman and colleagues recruited functional assessment experts to code participant responses into the categories that matched those on the checklist recording sheets. This procedure allowed the data to be quantified and more easily analyzed. Out of the two formats, A-B-C structured recording was found to be more accurate than narrative recording; however, neither format resulted in high levels of accuracy when using the most stringent method: percentage of problem behavior with correctly scored antecedents and consequences.

Mayer and Reed (2013) attempted to improve the accuracy of A-B-C narrative recording collected by direct service personnel. The authors created 5-min video clips of naturally occurring interactions between students and their caregivers. Some of the clips included problem behavior as the target behavior to be recorded; others included appropriate behavior such as functional communication. Accuracy in this study was measured by having five experts in behavior analysis watch all of the videos and complete an A-B-C narrative recording sheet. The experts' data were used to create a template against which the participants' responses could be compared. Agreement was calculated for the occurrence of setting information, antecedents, consequences, and behavior. It should be noted that the participants' responses did not have to match the exact phrasing of the experts'. This study only used occurrence agreement to measure accuracy, which is less stringent than the methods used by Lerman et al. (2009). The data do not indicate, for example, the extent to which the environmental variables surrounding each instance of the target behavior were correctly scored. In general, the authors found low levels of accuracy

during baseline measurement that improved following the introduction of an intervention combining task clarification and feedback.

Some researchers suggest that the first step to determining the accuracy of a direct FBA method is to examine whether the method's findings are consistent with another, well-established method for determining behavioral function (e.g., EFA; Lanovaz et al., 2013). There are several studies comparing the function identified through direct FBA methods to the function identified through an EFA (Camp et al., 2009; Lerman & Iwata, 1993; Thompson & Iwata, 2007). For example, Camp et al. (2009) used both direct FBA methods and EFA to identify the function maintaining the problem behavior of seven individuals with developmental disabilities, aged 16-54. They used A-B-C continuous recording to record the occurrence of problem behavior and antecedent and consequent events during continuous 10-s intervals. The authors recruited a team of behavior analysts to view the descriptive and functional analytic data and reach a consensus on function. The results indicate relatively low correspondence between the two methods, a finding consistent with results from similar studies (Lerman & Iwata, 1993; Thompson & Iwata, 2007).

Measuring the degree to which the results from direct FBA methods converge with those of an EFA does not answer the question of whether or not an assessment method is accurate (as defined by Hayes et al., 1986). When evaluating the accuracy of an assessment method, one is seeking to uncover how closely the assessment method produces data that approximate reality by comparing it to an indisputable standard (Hayes et al., 1986). None of the studies examining the convergence between EFA results and the results from A-B-C recording included such a standard, so there can be no confirmation of accuracy.

Sensitivity. Finally, sensitivity refers to an assessment's ability to detect changes in behavior that are produced by changes in the environment. Sensitivity does not assume accuracy. Rather, if one assessment method shows a change in behavior following an environmental manipulation a sensitive measure would reflect a similar change in behavior. A search of the literature revealed no empirical studies evaluating the sensitivity of direct FBA methods.

Research Implications

The first part of this chapter provided an overview of FBAs conducted in school settings touching upon the federal and state requirements, the ambiguity surrounding the process, and introduced some of the more common assessment methods (EFA, indirect, and direct) along with their strengths and limitations. The second half of the chapter described the different types of direct FBA methods, with a focus on A-B-C recording. A-B-C recording is the most commonly used method by practitioners and is the only direct FBA method with any empirical research evaluating its quality. The final section of this paper will pose several questions for the research community to address:

First, research is needed to examine the accuracy of behavioral assessment methods collected by school psychologists, who have more training in behavioral assessment and are less likely to be a part of the contingency they are responsible for recording. There are several reviews of FBAs conducted in applied settings (e.g., schools; Anderson et al., 2015; Stichter & Conroy, 2005). These reviews report that, within the existing research base, the researchers assume the largest responsibility for the FBA process—often selecting the methods of assessment to be used, conducting the observations, and analyzing the data. Anderson et al. (2015) suggest that this arrangement is likely to produce different outcomes than FBAs led by school-based practitioners. Furthermore, most of the research on FBAs that choose to incorporate

school-based personnel use classroom teachers as participants (Anderson et al., 2006; Lerman et al., 2009; Sasso et al., 1992). This is problematic, as it is likely that the teachers are providing the antecedents and/or consequences that they are responsible for recording (e.g., attention). As self-recording of behavior has been shown to have a reactive effect on behavior (Lipinski & Nelson, 1974), it is unlikely that the teacher-collected data using behavioral assessment methods provides an objective picture of the environment.

Second, more research is needed to establish the accuracy of direct FBA methods. Although an EFA is the only method that can be used to identify causal relationships between environment and behavior (Cooper et al., 2007), they are often infeasible in applied settings. Thus, direct FBA methods, with their flexibility, ecological validity, and potential for detailed analysis, provide a promising compromise. However, there is limited research evaluating the quality of the direct FBA methods currently available. The majority of the research evaluating the quality of direct FBA methods has focused on A-B-C recording; however, a search of the literature only yielded two studies examining the accuracy of A-B-C recording (Lerman et al., 2009; Mayer & Reed, 2013). Although a number of studies examine the degree to which the findings of A-B-C recordings converge with EFAs (Camp et al.,2009; Lerman & Iwata, 1993; Thompson & Iwata, 2007), these studies are not evaluating accuracy as there is no indisputable standard against which the data can be compared. Before the conclusion is made that the findings of A-B-C recording do not support the findings of an EFA, research must first establish the accuracy of A-B-C recording as a standalone method for determining behavioral function.

A third research idea is to examine the effect, if any, a priori knowledge of behavioral function (as hypothesized from an indirect assessment method) has on data gathered through subsequent direct FBA methods. Indirect FBA methods are traditionally employed at the

beginning of the FBA process (Floyd et al., 2005; Hanley, 2012). This creates the possibility that an observer conducting direct FBA methods is biased from their knowledge of the indirect assessment results. For example, knowing that a rating scale completed by a classroom teacher suggests attention as a possible function of problem behavior may result in the observer looking for evidence to support that function and ignoring evidence that refutes the hypothesis, a phenomenon known as confirmation bias.

Finally, research is needed to examine the feasibility of using a structured form of A-B-C recording in a continuous manner. There is evidence to suggest A-B-C structured recording is a more accurate and preferred assessment method relative to narrative recording (Lerman et al., 2009). Structured recording is also less complex, and therefore more feasible for use in classrooms, than A-B-C continuous recording. Despite its attributes, A-B-C structured recording is limited in that data are only collected when the target behavior occurs, restricting the analyses possible and preventing a complete sample of the ecology of the classroom in comparison to data collected in a continuous manner (regardless of behavior occurrence). As relevant antecedents and/or consequences may not be immediately apparent to the observer (Shapiro & Kratochwill, 2000), continuous recording ensures that a complete picture of the environment is captured.

CHAPTER 3

EXPERIMENTAL STUDY

In an effort to give every child access to a free and appropriate public education the Individuals with Disabilities Education Act (IDEA, 1997; 2004) was amended in 1997, mandating that schools conduct a functional behavior assessment (FBA) to address the needs of children with significant behavior problems under certain circumstances. Specifically, an FBA must be completed when a student with a disability is removed from school for more than 10 days (constituting a change in educational placement) as a result of misconduct determined to be a manifestation of his or her disability (IDEA, 2004). Although the circumstance in which an FBA is explicitly required under federal law is narrow, it may be appropriate for districts to conduct an FBA if that student's behavior is having a significant impact on his or her ability to learn, regardless of disability status (von Ravensburg & Tobin, 2008).

An FBA is a set of information-gathering techniques designed to identify environmental variables that may be maintaining a child's problem behavior (Alter et al., 2008; Cooper et al., 2007; Gresham et al., 2001). FBAs are comprised of methods that range from indirect assessment methods such as caregiver interviews and behavior rating scales, to direct or descriptive assessment methods to experimental manipulations, such as a structural or functional analysis (Alberto & Troutman, 2006; Cooper et al., 2007). Assessment methods are typically employed in order of least to most intrusive with indirect assessment methods employed first, then direct, and finally experimental (Floyd et al., 2005). Once the relevant environmental variables are identified, function-based treatments can be created to reduce undesirable behavior

and increase appropriate alternatives. Function-based treatments are individualized and are proven to be more effective in reducing problem behavior than universal behavior modification techniques (Cooper et al., 2007; Gable et al., 2014; Hanley et al., 2014; Herzinger & Campbell, 2007; Newcomer & Lewis, 2004).

Prior to the amendment within IDEA (1997; 2004), FBAs were rarely conducted outside of clinical settings (Anderson et al., 2015). Although now regularly conducted in schools across the country, the quality with which FBAs are being implemented in these applied settings is lacking. One major obstacle to their successful implementation is that IDEA fails to provide a definition of FBA or a list of required components (Asmus et al., 2002; Stage et al., 2006). Without clear direction there is great potential for FBAs to be conducted incorrectly, which jeopardizes the likelihood that resulting data will lead to effective, function-based interventions.

Another problem affecting the quality of FBA implementation in schools is the amount of time, resources, and training that are necessary to implement all of the components. An experimental functional analysis (EFA) involves the systematic manipulation of variables (e.g., attention, access to tangibles) hypothesized to be controlling the target behavior, while observing and recording the subsequent effect on behavior (Fisher et al., 1998; Rooker et al., 2011). EFAs are rarely conducted in applied settings, despite being the only FBA method that can validate the function of behavior (Cooper et al., 2007). The exclusion of this method in applied settings can be attributed to the lack of individuals trained in data collection, implementation, and analysis (Asmus et al., 2002), teacher perception that the process is time-consuming and difficult (Asmus et al., 2002; Calloway & Simpson, 1998; Hanley, 2012), and the lack of control inherent in a typical classroom setting (Bloom et al., 2011; Tiger et al., 2006). Due to these challenges with EFA implementation, schools often choose to employ direct FBA methods as a compromise

(Love et al., 2009). These assessment methods, if used correctly, allow one to obtain a complete picture of the contingencies operating in the natural environment with more flexibility than experimental methods allow (Floyd et al., 2005; Gresham et al., 2001) and with more objectivity than indirect FBA methods.

Despite the popularity of direct FBA methods within schools (i.e., A-B-C recording procedures), a thorough review of the literature suggests there is limited research evaluating the quality of these assessment methods. This is surprising considering the number of published reports calling for more evidence supporting the technical adequacy of FBA methods and the conditions under which they are most appropriate (Gable et al., 2014; Gresham et al., 2001; McIntosh et al., 2008). The majority of researchers that have attempted to evaluate the quality of FBA methods examined indirect FBA methods, particularly rating scales (Matson & Minshawi, 2007). This research is flawed in that the authors attempted to evaluate their measures from a traditional psychometric perspective (Koristas & Iacono, 2013; Shogren & Rojahn, 2003). This perspective assumes the causes of behavior reside within the individual and are stable over time and across contexts, which is in contrast to the purpose of conducting an FBA. In alignment with the science behind conducting an FBA, behavior is viewed as a function of the interaction between a person and his/her current environment (Hayes et al., 1986). Hayes et al. (1986) propose that behavioral assessments should not be evaluated using traditional reliability and validity standards but rather should be evaluated against three criteria: treatment utility, accuracy, and sensitivity. Out of the direct FBA methods, A-B-C recording is the only method with any empirical research evaluating its quality against any of these three criteria.

A-B-C recording is a recording method in which an individual observes behavior as it occurs in the natural environment and records the sequence of events surrounding the behavior

(the antecedents and consequences). There are multiple A-B-C recording formats (narrative, structured or checklist, and continuous). A-B-C structured recording is a more accurate and preferred assessment method relative to narrative recording (Lerman et al., 2009). With structured recording observers record the occurrence of the problem behavior and place checkmarks next to a list of pre-identified antecedent and consequent variables occurring contiguous to the behavior, whereas in narrative recording an observer must write down the behaviors and environmental events in narrative form as they are observed (Shapiro & Kratochwill, 2000; Thompson & Borrero, 2011). Lerman et al. (2009) compared the accuracy of teacher-collected A-B-C data using both narrative and structured recording methods. Two scripted videos were created by the authors that included a set number of problem behaviors with clear antecedents and consequences accompanying each instance of behavior. A scoring template was created for each video and was used as a gold standard against which the participants' recordings could be compared. To evaluate the accuracy of the structured recording method, teacher-completed forms were compared to the gold standard form created from the scripts. For narrative recording, the authors recruited functional assessment experts to code participant responses into the categories that matched those on the structured recording sheets. A-B-C structured recording was found to be more accurate than narrative recording; however, neither format resulted in high levels of accuracy when using the most stringent scoring method: percentage of problem behavior with correctly scored antecedents and consequences.

In addition to improved accuracy over narrative recording, A-B-C structured recording is less complex and time-consuming, and therefore more feasible for use in classrooms, than A-B-C continuous recording. A-B-C continuous recording uses codes for behavior, antecedents, and consequences, the definitions of which are often created from data collected through indirect

FBA methods or A-B-C narrative recordings (Alberto & Troutman, 2006; Cooper et al., 2007). Data collectors employ an interval (time-sampling) or event recording procedure to gather a sample of the events occurring in the natural environment within a set timeframe. Continuous A-B-C recording often requires extensive training and the use of specialized equipment not always available in schools (Tarbox et al., 2009); however, this form of recording allows for more complex analyses and a more complete sample of the environment in which the behavior occurs (Skinner et al., 2000; Thompson & Borrero, 2011). Research is needed to determine if one of these A-B-C recording methods, or a combination of these recording methods, can be considered an appropriate method for identifying behavioral function.

Before any format of A-B-C recording can be considered appropriate for use within an FBA, additional research is needed to examine the accuracy of the recording method. Some researchers suggest that the first step to determining the accuracy of a direct FBA method is to examine whether the method's findings are consistent with another well-established method for determining behavioral function, such as an EFA (Lanovaz et al., 2013). There are several studies comparing the function identified through direct FBA methods to the function identified through an EFA (Camp et al., 2009; Lerman & Iwata, 1993; Thompson & Iwata, 2007). Camp et al. (2009) used both direct FBA methods and EFA to identify the function maintaining the problem behavior of seven individuals with developmental disabilities, aged 16-54. They used A-B-C continuous recording, recording the occurrence of problem behavior and antecedent and consequent events during continuous 10-s intervals. The authors recruited a team of behavior analysts to view the descriptive and functional analytic data and reach a consensus on function. Results indicated relatively low correspondence between the two methods, a finding consistent with results from similar studies (Lerman & Iwata, 1993; Thompson & Iwata, 2007). Measuring

the degree to which the results from direct FBA methods converge with those of an EFA is addressing a separate research question. These studies do not answer the question of whether or not direct FBA methods are accurate when using the definition of accuracy provided by Hayes et al. (1986).

Although research suggests that data collected through direct FBA methods often do not correspond to data collected through EFAs (Camp et al., 2009; Lerman & Iwata, 1993), it is unclear as to whether this lack of correspondence is due to a flawed measurement procedure or to a failure of data collectors to collect data accurately. To date, few researchers (Lerman et al., 2009; Mayer & Reed, 2013) have examined the accuracy of data collectors using the definition provided by Hayes et al. (1986). Before the conclusion is made that the findings of direct FBA methods such as A-B-C recording do not support the findings of an EFA, research must first establish the accuracy of A-B-C recording as a standalone method. Accuracy of recording and accuracy of outcome must be measured separately.

A second problem with research examining the accuracy of A-B-C recording is the use classroom teachers as the participants in the studies (Lerman et al., 2009; Mayer & Reed, 2013). This is problematic for two reasons. First, research suggest that teachers rarely receive adequate training in FBA theory and implementation (Couvillon et al., 2009) and second, it is likely that the teachers themselves are providing the antecedents and/or consequences that they are responsible for recording. Teachers, therefore, should not be the individuals conducting the FBAs. Research is needed to examine the accuracy of behavioral assessment methods collected by school psychologists who have more training in behavioral assessment and are less likely to be a part of the behavioral contingency. Research shows that school psychologists, along with

special education teachers and behavior specialists, are most likely the ones responsible for conducting FBAs in schools (Scott & Kamps, 2007; Van Acker et al., 2005).

In addition to insufficient research examining its accuracy, there is an inherent flaw with the A-B-C narrative and structured recording methods which must be addressed before accepted as an alternative to experimental methods. These forms of A-B-C recording are discontinuous. This means that data are only collected when the target behavior occurs which restricts possible analyses and prevents collection of a true sample of the ecology of the classroom in comparison to data collected in a continuous manner (regardless of behavior occurrence). Conditional probabilities can be calculated when data are collected in a continuous manner. Conditional probabilities provide the probability of an event (e.g., consequence such as attention) given the occurrence of another event (e.g., the target behavior) (Thompson & Borrero, 2011). To increase the likelihood of identifying a contingency between environment and behavior, the conditional probability can be compared to the background probability of the same event (Rooker et al., 2011). In practice this means the continuous recording of environmental variables as they occur, along with instances of both inappropriate behavior and appropriate behavior. As relevant antecedents and/or consequences may not be immediately apparent to the observer (Shapiro & Kratochwill, 2000), data collected in a continuous manner ensures that an accurate sample of the environment is captured. Research is needed to explore the accuracy of an A-B-C recording method that combines the practitioner-friendly format of structured recording with the capability of being collected in a continuous manner.

Purpose

Due to the practical barriers associated with the implementation of EFAs in schools, direct FBA methods are frequently employed in an attempt to determine the function of problem

behavior. Despite the popularity of direct FBA methods in schools there is limited research evaluating their accuracy, which has serious implications for the functional relevance of the interventions designed from the resulting data. In addition to a lack of research examining the accuracy of the recording method itself, there is limited research evaluating other potential factors impacting accuracy. For example, as indirect FBA methods are traditionally employed at the beginning of the FBA process in practice (Floyd et al., 2005; Hanley, 2012), it is possible that an observer conducting direct FBA methods is biased from their knowledge of the indirect assessment results. Knowing that a rating scale completed by a classroom teacher suggests attention as a possible function of problem behavior may result in the observer looking for evidence to support that function and ignoring information to refute it, a phenomenon known as confirmation bias. Such bias could potentially impact the accuracy of observers' data recording when conducting direct FBA methods and impact their ability to identify the function of a student's problem behavior.

In addition to the limited research evaluating the accuracy of A-B-C recording (Lerman et al., 2009; Mayer & Reed, 2013), the research that does exist primarily uses classroom teachers as participants which may jeopardize the validity of the resulting data as teachers are often a part of the behavioral contingency. As school psychologists presumably have training in behavioral observation and are removed from the behavior of interest, they are ideal participants for a study that evaluates the accuracy of a direct FBA method. Although school psychologists appear ideally suited for the task of conducting FBAs, it is unclear the extent to which they are implementing them in practice and whether or not more experience with FBAs in general, and direct observation methods specifically, results in more accurate data collection. In an attempt to

address some of these limitations in the current research and practice, this study sought to answer the following research questions:

1. Can school psychologists accurately collect data on student behavior using a structured form of A-B-C recording in a continuous manner (the Serial Recording form)?
2. Do participants become more accurate in their data collection over time?
3. Is there a significant difference in the accuracy of recording based upon a priori knowledge of the hypothesized function of behavior?
4. Is there a significant relationship between participant accuracy of recording and accurate identification of the function of the student's target behavior?
5. Does a priori knowledge of hypothesized behavioral function bias participants' selection of behavioral function?
6. Do school psychologists incorporate standard practice components for FBA completion into their FBAs?
7. Is there a significant relationship between participant experience with FBA and accuracy of recording?

Methods

Participants and Settings

Participant characteristics. Participants included 44 school psychologists from 13 school districts and 4 advanced graduate students in school psychology (8 males and 39 females). The majority of the participants were Caucasian (83.3%) with a terminal degree of Education Specialist (54.2%; see Table 1). Note that demographic data were missing for one school psychologist. As an a priori power analysis indicated the need for 24 participants in each of the two conditions, advanced Ed.S. and Ph.D. level school psychology students were recruited

in order to achieve an adequate sample size. School psychology students who were behavior analysts or who were actively pursuing certification as a behavior analyst were excluded, as not all school psychologists are required to have this level of behavioral training. Data were either collected during regularly scheduled monthly professional development meetings or at a district building before or after work.

Participants were randomly assigned to one of two conditions upon signing the consent form to participate. The two conditions differed in that in one condition (Informed) participants were provided with a hypothesized function of the target students' inappropriate behavior. Participants assigned to the second condition (Neutral) were not provided with a hypothesized function. Sessions occurred in a group setting to facilitate administration; however, each participant watched the videos on their individual laptops, using their personal headphones to listen to the audio. On average, administration procedures took between 45-60 minutes. Each participant had a unique code, which was written on the top right corner of each page in their packet. The code had a number (e.g., 01, 02, etc.) and a letter (A or B). The letter indicated the experimental condition to which the participant was exposed (e.g. 'A' for Informed and 'B' for Neutral).

Table 1

Demographics of a Sample of School Psychologists and Advanced School Psychology Graduate Students

Characteristic	*n*	%
Participants	48	
School Psychologists	44	91.67
Graduate Students	4	8.3
Gender	47	
Male	8	16.7
Female	39	81.3
Ethnicity		
African-American	4	8.3
Asian	1	2.1
Caucasian	40	83.3
Hispanic	1	2.1
Other	1	2.1
Degree		
Ed.S	26	54.2
Ph.D.	12	25.0
M.A.	1	2.1
Other	8	16.7

Materials

Participant packets. Each participant received the same recording packet, regardless of the condition to which they were assigned. Each packet began with a cover sheet, consent form and background questionnaire (Appendices A and B). Background questionnaires were used to collect (a) basic demographic information, (b) the number of years participants had served as a practicing school psychologist or as a school psychology graduate student, (c) their terminal degree and any additional certifications or licenses, (d) information on prior coursework and training experiences, (e) the number of FBAs conducted in the past 3 months, and (f) the percentage of their most recent FBAs that involved an A-B-C data collection procedure. Finally, participants were asked to reflect upon the past five FBAs they had conducted and indicate through a checklist which of the standard practice components, as identified by Weber et al. (2005), they included in their FBAs.

Next, sheets with operational definitions, examples, and non-examples of the setting events and environmental events (Appendix C) and target student behaviors (Appendix D) were provided. Following those sheets were two Serial Recording forms. One form was intended to be completed while participants watched a 2-min training video (Appendix F) and one form was intended for a 10-min test video (Appendix I). They were labeled "Training Sheet" and "Test Sheet." The back of the training video recording form provided participants with an answer key so that participants could check their accuracy. An answer key was only provided for the training video. There were two, multiple-choice questions on the back of the test video recording form. The first asked participants to select the function they believed to be maintaining the student's target behavior (TB) out of the following response options: attention (ATTN) or escape (ESC). The second multiple-choice question asked participants to select the next most likely function to

be maintaining the TB if they were not certain about their first selection. The same response options were available as provided on the first question (ATTN or ESC), with the addition of the response option, "No second function."

***Hypothesized function*.** Before viewing the test video, participants were directed to read a word document containing information about the target student on whom participants were to collect data (Appendices G and H). The document included the student's name, grade, and age, and provided a brief summary of his behavior. Participants who were assigned to the Informed condition had an additional piece of information in their statements. For this condition, the statement included an excerpt from the child's classroom teacher that provided the teacher's opinion of the function of the student's problem behavior (ATTN):

> Timothy is all over the place and cannot seem to focus on the lesson. I am constantly redirecting him. If I had to guess, I would say Timothy is acting out to get attention from a couple of the other children in the class and to see if he can get a rise out of me (Appendix G).

The hypothesized function was presented in this manner because it reflects standard practice in the school system. School-based practitioners typically conduct observations following communication with the child's teacher during which a hypothesized function is provided, either in the form of a rating scale, interview, or informal conversation. In this case the hypothesized function was incorrect.

The final page of each packet contained a follow-up questionnaire used to assess participants' perceptions of the Serial Recording form (Appendix F). Using a 5-point Likert-type scale, participants were asked to provide information on the perceived ease of using the Serial Recording form, the relevance of the form for developing a hypothesis of the function of

behavior, the likelihood of the participants using the Serial Recording form in their practice, and their confidence in the function of the problem behavior they selected.

Video recordings. Two videos simulating behavioral observations in a classroom were created for the study. The first was a 2-min training video that allowed participants to practice using the Serial Recording form. The second video was a 10-min test video. The videos were divided into 1-min segments that corresponded to the recording forms. To ensure participants were recording their observations in the appropriate segment, each minute was separated by a 3-s long title card with the upcoming minute written on the card (e.g., "Minute 1. Start…Minute 2. Start"). When the first minute elapsed, the participant was instructed to move down the recording form to the second row to complete the behavioral sequence for the second minute, and so on. A screenshot of the target student was displayed for 5-s prior to the start of the training and test videos. There was also be a brief sentence stating the behavioral expectations for the classroom, so the participant was able to accurately code the target student's behavior when the video began.

Video development. Both the training and test videos were developed by first generating a list of environmental events that are present in classrooms (e.g., teacher attention, delivery of instructions). Next, scripts were developed (Appendix K) that incorporated these environmental events within the context of group instruction in an elementary school classroom. Child and adult actors acted out the scripts while being filmed. The videos were taped in an elementary school classroom located within a school in the southeast. Participants were instructed to collect data on the behavior of the child within the video designated as the target student. The problem behaviors selected for the videos were discrete behaviors defined for participants prior to data collection. Problem behaviors were classified either as TB or as "other inappropriate behavior"

(OIB). There were three codes of problem behavior (a combination of the TB and OIB) in the training video, distributed across two, 1-min segments. For the test video, there were 22 codes of problem behavior distributed across 10, 1-min segments. There was an interval of at least 20-s between codes of problem behavior to provide enough time for the participants to mark the relevant environmental events surrounding the behavior. Therefore, there were never more than three codes of problem behavior within each 1-min segment. The test video script was written to indicate ESC as the function of the TB. In the absence of a standard for validating the function of a behavior, the script was written so that ESC was the only consequence (out of the list of environmental events provided to the participants) that occurred within 10-s of the TB for at least 80% of the TB occurrences. This standard was selected as there is research to suggest that the most direct effects of reinforcement decrease as the delay between the behavior and reinforcer is increased (Cooper et al., 2007).

After the scripts were written and the videos were filmed, the primary researcher filled out a Serial Recording form that served as the gold standard form against which participants' data were compared. An expert in behavioral assessment who was unaffiliated with the current study watched both the training and test videos to verify that all problem behavior and environmental events were portrayed as intended. The expert verified that every event on both video scripts actually occurred in the videos, in the correct order. Next, the expert was provided with an abbreviated version of the training and test video scripts with only the behaviors and environmental events that were coded on the gold standard form. The expert then coded each of these behaviors and events.

Serial recording form. Instructions on how to complete the Serial Recording form were presented via a 12-min video (see Appendix E) allowing for standardization of instructions

across participants within and between conditions. Participants were encouraged to reference their packets while attending to instructions. During the introduction, the primary researcher first discussed the purpose of A-B-C recording and then provided a brief explanation of the rationale for recording the environmental events surrounding both inappropriate and appropriate behavior. Next, operational definitions of the environmental events and the three student behaviors were provided. There was a brief, video example of a child actor engaging in the TB and OIB. Next, there was a thorough discussion of the recording method, and the importance of recording each event in chronological order. Participants then watched a brief, videotaped example of a behavioral sequence occurring in a classroom. Finally, participants had the opportunity to watch as that sequence was recorded on the Serial Recording form.

After watching the instructional video, participants watched the 2-min training video and had the opportunity to practice collecting data using the Serial Recording form. Participants had access to the answer key for the training video, so they were able to compare their responses to the gold standard. Participants were allowed to watch the training video and practice scoring as many times as they liked before starting the 10-min test video. Each participant was sent an individual email with a link to view the instructional video, training video, and test video.

Coding. Participants coded student behavior as either the TB, OIB, or appropriate behavior (AB). These definitions were provided on a sheet of paper distributed with the Serial Recording forms (Appendix D), and were reviewed in the instructions delivered at the beginning of the video. Definitions for OIB and AB were loosely adapted from the behavioral definitions provided on the Behavioral Observation of Students in Schools (BOSS) (Shapiro, 2003). The TB selected for this observation was leaving the instructional area. This was defined as any instance when the target student left his assigned area during instruction and/or at a time when the

directions were to remain in his area. OIB was defined as any active off-task behaviors (motor activity or audible verbalizations) that were not directly associated with an assigned task, were not permitted, and prevented the child from receiving instruction. The TB (leaving the instructional area) was not to be coded in this category. If the target student was out of the instructional area when he engaged in OIB, the participant was instructed to only code the student as engaging in the TB. Finally, AB was defined as any time the student was actively or passively attending to the assigned work and was not engaging in the TB or OIB. Examples and non-examples for all student behaviors were provided with the operational definitions that the participants could reference at any time.

Participants were instructed to indicate the presence of each environmental event by selecting from a list of objective descriptions of events that typically occur in the classroom environment. The Serial Recording form had antecedents and consequences grouped together under the heading of environmental events for ease of recording. The environmental events included: (a) teacher directive to group (TD-G), (b) individual teacher attention- negative (ITA-N), (c) individual teacher attention- other (ITA-O), (d) peer attention (PA), and (e) escape from work (ESC). Operational definitions for the environmental events, as well as examples and non-examples, were provided to the participants (Appendix C). Participants were asked to indicate which of the following setting events took place within each 1-min interval. The setting events included: (a) whole group instruction, (b) small group instruction, (c) independent work, and (d) transition. Operational definitions of the setting events were also provided to participants in their packets (Appendix C).

The environmental events and student behaviors were in columns listed across the Serial Recording form. The rows listed the setting events and the time intervals in 1-min increments.

The videos clearly stated the start time for each minute, so participants were cued to the correct row. Participants first recorded the student behavior at the beginning of each interval and then proceeded with writing numbers, in chronological order, under each student behavior or environmental event as it was observed. For example, when the video began the participant was to place a '1' in the first box under the student behavior occurring at that moment. The behavior or environmental event that immediately followed it was assigned a '2', and so on. Participants circled setting events whenever they occurred, and occasionally multiple setting events occurred within the same minute. Once a new minute began, participants shifted their recording one row down, and began recording a new sequence starting with number '1' again. Participants were instructed to always record environmental events as they were observed, regardless of whether the child was engaging in a problem behavior (TB/OIB) or AB.

Analyses

Total accuracy. To answer the research question-- Can school psychologists accurately collect data on student behavior using a structured form of A-B-C recording in a continuous manner (the Serial Recording form)?— two types of agreement were calculated: (b) *occurrence agreement,* the degree to which participants identified the same number of environmental events and behaviors in each interval, regardless of their place within the sequence and (b) *sequential agreement,* the accuracy with which participants captured the behavioral sequence. To determine occurrence agreement for the environmental events and the student behaviors, the primary researcher calculated the percentage of events and behaviors that the participants recorded out of those present on the gold standard Serial Recording form, for each sequence. The order in which participants recorded environmental events and behaviors was not factored into this calculation; if the participant scored the same number and type of events and behavior within each 1 min

interval, it was scored as an agreement. The primary researcher further analyzed the data by examining any differences in accuracy across the individual environmental events and student behaviors to see if participants were more or less accurate with scoring the different events and behaviors.

The primary researcher calculated sequential agreement by comparing the behavioral sequences scored by the participant to the gold standard sequence. For example, if a participant recorded the same antecedent and behavior but a different consequence than a behavioral sequence on the gold standard form, the participant accurately scored two out of three parts of that behavioral sequence, or 66.7%. The primary researcher obtained a mean accuracy score for each participant across all behavioral sequences, and this number was referred to as their unadjusted accuracy score. After the initial analysis, the researcher performed a second analysis that did not penalize the participant for insertions. With the first analysis one error (e.g., an insertion) penalized the remainder of the sequence, resulting in a low accuracy score even if participants accurately scored the remainder of the sequence. Accuracy scores obtained through this adjusted scoring method are referred to as adjusted accuracy scores. Both the unadjusted and adjusted scores were used as each participant's accuracy scores for the remaining analyses. In applied behavior analysis research, 80% is the conventional level for what is considered an acceptable level of agreement (Cooper et al., 2007), therefore this is the standard used for acceptable agreement in this study.

Accuracy across time. To answer the research question—Do participants become more accurate in their recording over time?—the primary researcher conducted a repeated measures ANOVA and reported coefficients for each time point (minutes 1-10).

Accuracy across condition. To answer the research question—Is there a significant difference in the accuracy of recording based upon a priori knowledge of the hypothesized function of behavior?— the primary researcher conducted an independent samples t-test to determine whether the means of the Informed and Neutral conditions differed significantly.

Accuracy of function. To answer the research question—Is there a significant relationship between participant accuracy of recording and accurate identification of the function of the student's target behavior?—the primary researcher conducted a one-way ANOVA with accuracy of recording as the dependent variable and accuracy of function as the independent variable. Each participant was re-coded into one of three groups based on their responses to the two multiple-choice questions regarding the function of the student's TB. The first group was comprised of participants who chose ESC as the function of the behavior, with no second function (ESC only). The second group included participants who chose ESC as the primary function, with ATTN as the second function (ESC/ ATTN). The third group included participants who chose ATTN as the function of the student's TB, with either no second function or ESC as the second function (ATTN).

To answer the research question—Does a priori knowledge of hypothesized behavioral function bias participants' selection of behavioral function?— the primary researcher used Fisher's Exact Probability test to determine whether experimental condition was significantly related to participants' responses regarding the function of the target behavior. The primary researcher calculated odds ratios for the different response options (ESC, ESC/ATTN, ATTN) to further examine the associations between condition and selection of behavioral function. Resulting data provided an indication of whether participants provided with the hypothesized,

incorrect function (ATTN) were more likely to select ATTN as either the primary or secondary function of the TB than participants without access to the hypothesized function.

To answer the research question—Do school psychologists incorporate standard practice components for FBA completion into their FBAs?—the primary researcher conducted descriptive statistics on participants' responses to item number 10 of the background questionnaire. This item asked participants to refer back to the past five FBAs they completed and estimate out of those five FBAs, how many incorporated each one of the components. Participants had the opportunity to view the standard practice components and record their number next to each item on the questionnaire.

To answer the research question—Is there a significant relationship between participant experience with FBA and accuracy of recording?— the primary researcher ran correlations with accuracy as the dependent variable and the participants' self-reported experience with FBA and years of experience as a school psychologist as the independent variables. As the independent variables were skewed, the researcher used Spearman rho statistics to examine the relationship between experience and accuracy.

Results

Total Accuracy

Occurrence agreement. Mean occurrence agreement of environmental events and student behaviors was calculated (see Table 2) and a repeated measures ANOVA was used to examine differences in accuracy within and across the environmental events and student behaviors. As the assumption of sphericity was violated when examining the accuracy within environmental events, a Greenhouse-Geisser correction was used and it was determined that mean accuracy levels for occurrence agreement differed significantly between environmental events ($F(3.217, 151.191) = 114.448$, $p < 0.001$). Post-hoc procedures using the Bonferroni

correction revealed school psychology participants were significantly less accurate in their recording of escape (ESC) than any of the other environmental events ($p < .001$). This was followed by their recording of teacher-directed instruction to the group (TD-G), which was significantly less accurate than recordings of the other teacher and peer-related environmental events ($p < .001$). The majority of participants underreported instances of ESC and over-reported TD-G. Participants were most accurate in recording individual teacher attention directed to the target student which was neutral (ITA-O; $p < .01$), and participants were significantly more accurate in their recording of negative individual teacher attention (ITA-N) than recordings of ESC and TD-G ($p < .01$).

Results from the repeated measures ANOVA also indicated that mean accuracy levels for occurrence agreement differed significantly for student behaviors ($F(2, 94) = 21.762$, $p < .001$). Post-hoc analyses showed school psychology participants were significantly more accurate in scoring occurrences of the target behavior (TB) compared to other inappropriate behavior and appropriate behavior (OIB, AB; $p < .001$). OIB and AB did not differ significantly from one another. Figure 1 displays participant occurrence agreement across each minute of the video for both environmental events and behaviors.

Table 2

Mean Numbers, Ranges, and Accuracy for Total Instances of Environmental Events and Student Behaviors

	True instances	Participant responses	Mean occurrence agreement	Range
		Environmental Events		
TD-G	13	16.42 (3-30)	.40	.10-.70
ITA-N	11	9.6 (6-14)	.67	.30-.90
ITA-O	2	3.77 (1-9)	.75	.40-1.0
ESC	24	12.75 (0-27)	.19	.00-.70
PA	9	4.1 (0-10)	.61	.30-.90
Total			.525	
		Student Behaviors		
TB	10	9.94 (4-24)	.67	.10-1.0
OIB	12	15.48 (0-30)	.43	.00-.80
AB	10	7.56 (1-15)	.48	.10-.80
Total			.529	

Note. TDG = Teacher directive to group; ITA-N = Individual teacher attention-Negative; ITA-O = Individual teacher attention- Other; ESC = Escape from work; PA = Peer attention; TB = Target behavior; OIB = Other inappropriate behavior; AB = Appropriate behavior

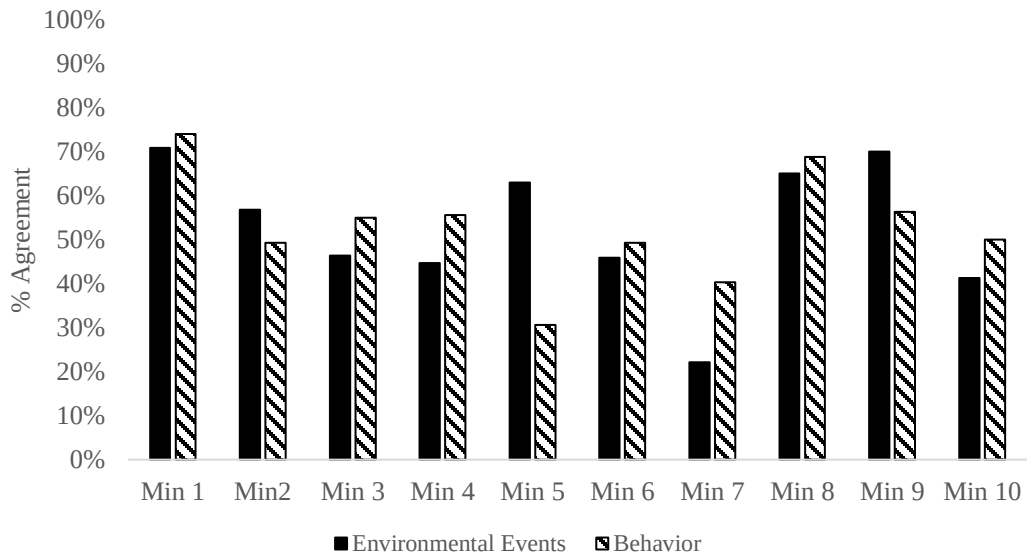

Figure 3.1. Occurrence Agreement across Time

Sequential agreement. All environmental events and student behaviors from each minute of the video were recorded onto a Serial Recording form, which was used as a gold standard against which the participants' recordings could be compared. Sequential agreement was calculated by comparing the behavioral sequences scored by each participant to the gold standard sequence, and a mean accuracy score was obtained for each participant across all sequences. This number was referred to as their unadjusted accuracy score. After the initial analysis, the primary researcher performed a second analysis that did not penalize the participant for insertions. With the first analysis one error (e.g., an insertion) penalized the remainder of the sequence, resulting in a low accuracy score even if participants accurately scored the remainder of the sequence. Accuracy scores obtained through this adjusted scoring method are referred to as adjusted accuracy scores. Both the unadjusted and adjusted scores were used as each participant's accuracy scores for the remaining analyses. Tables 3 and 4 provide the means, standard deviations, and ranges for the unadjusted and adjusted accuracy scores. On average, participants were more accurate when using the adjusted scoring procedure (see Table 4),

compared to the unadjusted scoring procedure (see Table 3). A visual display of participant accuracy across each minute of the video, using both adjusted and unadjusted accuracy scores, is shown in Figure 2.

Table 3

Means, Standard Deviations, and Ranges for Unadjusted Accuracy Scores

Accuracy	Length of Sequence	*M*	*SD*	Minimum	Maximum
Min 1	8	.58	.28	.00	1.0
Min 2	10	.31	.196	.00	.70
Min 3	9	.30	.18	.00	.67
Min 4	7	.24	.20	.00	.71
Min 5	7	.29	.21	.00	.86
Min 6	9	.22	.198	.00	.78
Min 7	14	.15	.12	.00	.50
Min 8	9	.34	.24	.00	1.0
Min 9	8	.35	.19	.00	.75
Min 10	9	.13	.12	.00	.44
Total		.29	.10	.08	.50

Table 4

Means, Standard Deviations, and Ranges for Adjusted Accuracy Scores

Accuracy	Length of Sequence	*M*	*SD*	Minimum	Maximum
Min 1	8	.81	.18	.13	1.0
Min 2	10	.64	.21	.10	1.0
Min 3	9	.60	.197	.11	.89
Min 4	7	.48	.20	.14	.86
Min 5	7	.63	.17	.29	1.0
Min 6	9	.58	.19	.11	1.0
Min 7	14	.46	.15	.21	.79
Min 8	9	.66	.22	.11	1.0
Min 9	8	.62	.19	.25	1.0
Min 10	9	.39	.14	.11	.78
Total		.59	.13	.30	.85

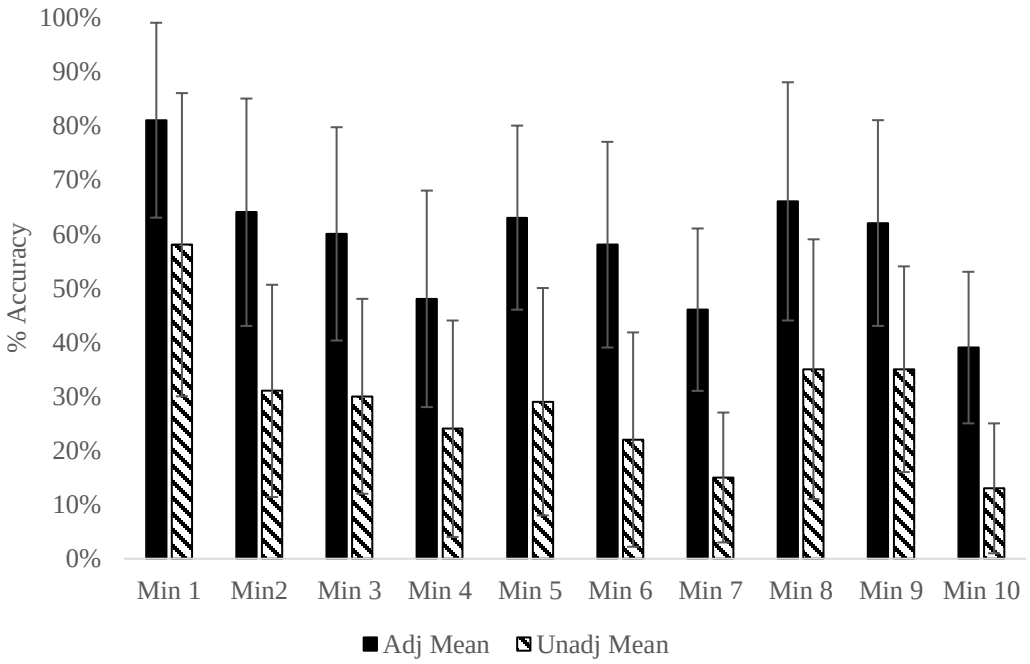

Figure 3.2. Accuracy across Time

Interobserver agreement

An independent observer scored 20% of the participants' responses using both occurrence agreement and the unadjusted and adjusted scoring methods for sequential agreement. The independent observer was able to achieve acceptable levels of agreement with the gold standard Serial Recording form using occurrence agreement and unadjusted and adjusted methods of sequential agreement. Mean agreement across minutes for occurrence agreement was 100%. Mean agreement across minutes using the unadjusted scoring method was 98% (range, 90-100%) and mean agreement using the adjusted scoring method was 94% (range, 80-100%).

Accuracy across Time

Adjusted accuracy. Mauchly's Test of Sphericity indicated that the data did not violate the assumption of sphericity (p = .199). There were significant differences in participant

accuracy across time, *F* (9, 44). Post-hoc tests using the Bonferroni correction revealed that participants did not become more accurate over time; rather, they were most accurate in their scoring during the first minute of the training video, with accuracy of min 1 significantly higher than all other minutes ($p < .001$). Participants were least accurate in their scoring of minutes 7 and 10, with accuracy for each of those minutes significantly lower than the majority of the remaining minutes ($p < .001$; min 7 less than minutes 1, 2, 3, 5, 6, 8, 9; min 10 less than minutes 1, 2, 3, 5, 6, 8, 9).

Unadjusted accuracy. Using the unadjusted accuracy scores, Mauchly's Test of Sphericity indicated that the data violated the assumption of sphericity ($p = .025$), and therefore Greenhouse-Geisser correction procedures were used. As with the adjusted accuracy scores, there was a significant effect of time on participant accuracy, $F(6.852, 44) = 23.445$, $p < .001$. Post-hoc tests using the Bonferroni correction again revealed the same pattern of participant accuracy. Participants were most accurate during min 1 ($p < .001$), and least accurate during minutes 7 and 10 ($p < .001$; min 7 less than minutes 1, 2, 3, 5, 8, 9; min 10 less than minutes 1, 2, 3, 5, 8, 9).

Accuracy across Conditions

Independent samples t-tests were used to examine differences in accuracy across conditions. The analyses indicate that those in the Informed condition were not significantly different from those in the Neutral condition when using both the adjusted ($p = .109$) and unadjusted ($p = .108$) accuracy scores (see Tables 5 and 6). These results suggest that participants' accuracy was not influenced by a-priori knowledge of the hypothesized function of behavior. For all remaining analyses, only adjusted accuracy scores were analyzed, as there was so little variability in the unadjusted scores due to much lower accuracy scores across minutes.

There was a greater amount of variability within the adjusted accuracy scores, making them a more appropriate measure of accuracy for the remaining analyses.

Table 5

Comparison of Informed and Neutral Conditions on Adjusted Accuracy Scores

Variable	*M*	*SD*	*t*	*df*	*p*
Accuracy			1.634	46	.109
Informed	61.64	12.25			
Neutral	55.64	13.18			

Table 6

Comparison of Informed and Neutral Conditions on Unadjusted Accuracy Scores

Variable	*M*	*SD*	*t*	*df*	*p*
Accuracy			1.641	46	.108
Informed	31.53	8.43			
Neutral	26.80	11.32			

Accuracy of Function

A one-way ANOVA was conducted to determine whether participants differed in their identification of the function of behavior based upon the accuracy of their recordings. Prior to conducting the analysis, each participant was re-coded into one of three groups based on their responses to the two multiple-choice questions regarding the function of the student's TB. The

first group included participants who chose ESC as the function of the behavior, with no second function (ESC). The second group included participants who chose ESC as the primary function, with attention (ATTN) as the second function (ESC/ ATTN). The third group included participants who chose ATTN as the function of the student's TB, with either no second function or ESC as the second function (ATTN). The analysis indicated that accuracy in recording did not have a significant effect on the accuracy with which participants identified the function of behavior, $F(2, 45) = 1.933$, $p = .157$ (see Table 7).

Table 7

One-Way Analysis of Variance Summary Table Comparing Response Groups on Total Adjusted Accuracy Scores

Source	*df*	*SS*	*MS*	*F*	*p*
Accuracy					
Between Groups	2	.062	.031	1.933	.157
Within Groups	45	.725	.016		
Total	47	.788			

Table 8

Means and Standard Deviations Comparing Total Adjusted Accuracy for Three Response Groups

Function groups	*n*	*M*	*SD*
ESC	23	.584	.133
ESC/ATTN	19	.616	.121
ATTN	6	.499	.119
Total	48	.586	.129

Fisher's Exact Probability test indicated a statistically significant difference between condition and selection of behavioral function, $p = .024$ (see Table 9). The odds of those in the Informed condition choosing ESC as the only possible function of the TB was 1.18 times as likely as the probability of those in the Neutral condition choosing ESC as the sole function of the TB, indicating a relatively small difference between conditions. However, the odds of participants in the Informed condition choosing ATTN as the second function, after ESC, was 2.43 times as high as the probability of those in the Neutral condition choosing ATTN as the second function. Odds ratios for the third response option (ATTN) could not be interpreted because the probability of those in the Informed condition choosing ATTN as the primary function of the TB was zero.

Table 9

Fisher's Exact Probability Test of Selection of Behavioral Function across Groups

		Function group			
Variable	*n*	ESC	ESC/ATTN	ATTN	*p*
Condition					.024
Informed	24	12	12	0	
Neutral	24	11	7	6	
Total	48	23	19	6	

Standard Practice Components

Descriptive statistics were run to examine which standard practice components participants included in their FBAs (see Table 10). Percentages reflect the percentage of participants' most recent FBAs that included each item from the list of standard practice components provided on their background questionnaire. The data indicated that FBAs completed by participants were most likely to include a definition of the TB (87%), a review of records (85%), and a hypothesis regarding the function of the problem behavior (75%). Conversely, participants were unlikely to incorporate scatterplots and experimental functional analyses (EFAs) into their FBAs, with only 20% of the most recent FBAs containing scatterplots and 13% containing an EFA. Participants reported using direct observation with no manipulation in 66% of their most recent FBAs and reported collecting some form of A-B-C recording in 69% of their most recent FBAs.

Table 10

Number, Means, and Standard Deviations of Standard Practice Components for FBAs Used by Participants

Criteria	*n*	*M*	*SD*
Definition/ identification of target behavior	45	.87	.34
Review of records	45	.85	.34
Checklist data	44	.63	.43
Student interview	44	.41	.40
Other interview types	44	.59	.45
Team meetings	43	.74	.43
Direct observation with no manipulation	45	.66	.41
Scatterplot	43	.20	.38
A-B-C recording	44	.69	.40
Functional analysis observation form	43	.48	.46
Reinforcer identification	43	.64	.43
Ecological context	42	.42	.49
Development of hypothesis	44	.75	.44
Experimental manipulation (EFA)	44	.13	.33

Experience

FBAs completed in the past three months. Given skewness (2.68) in the distribution of data for the number of FBAs participants reported having conducted in the past three months, a

Spearman rho statistic was calculated to examine the relationship between number of FBAs reported in the past three months and total accuracy scores. Results indicated no significant relationship between the two variables $r = .17$, $p = .266$.

Years of experience as a school psychologist. A Spearman rho statistic was used to determine the relationship between years of experience practicing as a school psychologist and total accuracy, as this variable was also skewed (skewness = 1.22). This analysis indicated a significant correlation between the two variables $r = -.30$, $p = .039$. The direction of the correlation was negative, suggesting that participants with more experience tended to have lower accuracy scores.

Social Validity

In general, participants did not believe the Serial Recording form was easy to use (2.4 out of 5), nor were they likely to use it in the future (2.9). On average, participants reported that the Serial Recording form was a helpful tool for determining behavioral function (3.9) and they were moderately confident in their selection of the function of the student's TB (3.5). A regression analysis was conducted to investigate whether accuracy of recording predicted self-reported confidence regarding the selection of behavioral function. The results were statistically significant, $F(1, 46) = 12.902$, $p = .001$ with an adjusted R^2 value of .202. According to Cohen's (1988) guidelines, that is a medium effect.

Discussion

The purpose of this study was to replicate and extend current research by evaluating the accuracy of a structured form of A-B-C recording when used in a continuous manner (the Serial Recording form). There is limited research evaluating the accuracy of direct assessment methods using the definition provided by Hayes et al. (1986) wherein data are compared to an

indisputable standard (Lerman et al., 2009; Mayer & Reed, 2013). The lack of such research is problematic as these assessment methods are frequently used by school-based practitioners (Kern et al., 2004; Love et al., 2009). The research that does exist has examined the accuracy of A-B-C narrative recording and A-B-C structured recording (Lerman et al., 2009; Mayer & Reed, 2013)—two methods of direct assessment that are touted practitioner-friendly due to their perceived ease of use and feasibility in schools relative to more complex forms of behavioral observations (Lerman & Iwata, 1993; Thompson & Borrero, 2011). The existing research has evaluated accuracy using A-B-C recording in a discontinuous manner, in which environmental events are only recorded if they occur immediately before or after the target behavior (TB). Even when using discontinuous recording, which reduces the number and range of behaviors and environmental events to be recorded, participants did not achieve an acceptable level of accuracy in their recordings, using the 80% standard for acceptability provided in behavior analytic research (Cooper et al., 2007).

Accuracy

Accuracy in scoring. As with prior research (Lerman et al., 2009; Mayer & Reed, 2013), unacceptable levels of accuracy between participant recordings and the gold standard Serial Recording form were found in the current study, with participants' total accuracy scores for student behaviors and environmental events recorded being below the criterion of 80%. Participants were more accurate when scoring the occurrence of behavior and environmental events in isolation (range, 32 to 70%), compared to the accuracy with which they recorded the correct sequence of those events and behaviors (Unadjusted range, 8 to 50%; Adjusted range, 30 to 85%); however, both occurrence and sequential agreement yielded unacceptable accuracy scores.

Student behaviors. Accuracy scores for the occurrence of the TB exceeded those of other inappropriate behavior (OIB) engaged in by the student and appropriate behavior (AB) engaged in by the student (See Table 2). This finding might be expected given existing practices, in which participants are primed to look for instances of the pre-determined TB and ignore other behaviors displayed by the target student (Thompson & Borrero, 2011). For instance, a school psychologist recruited to address a child's call-outs would come into the classroom to observe that specific behavior, along with its antecedents and consequences, and ignore other behavior exhibited by the target student. Given that the school psychology participants in this study were asked to collect data on more than simply the TB, efforts were made during training to prepare them for the recording of other behaviors. For example, they were provided with extensive operational definitions, as well as examples and non-examples of environmental events and student behaviors, video examples of each student behavior, and the opportunity to observe and score a 2-min training video. Furthermore, care was taken to select a TB that was overt and presumably easier to identify within the context of continuous observation (leaving the instructional area) recording procedures. Despite efforts to increase accuracy through the provision of explicit definitions and the selection of a TB that was purposefully selected to be easy to observe, overall, participants did not achieve acceptable levels of accuracy in coding the occurrence of TB (mean, 67%; range 10 to 100%).

There was considerable variability in participants' levels of accuracy in coding the occurrence of OIB (mean, 43%; range, 0-80%) and AB (mean, 48%; range, 10-80%), both of which were significantly less accurate than their coding of the TB. This suggests that participants were more accurate when scoring discrete, clearly-defined instances of behavior, rather than a code which encompasses multiple student behaviors. This finding is not surprising, as recording

multiple behaviors is more complex than recording a single behavior (Cooper et al., 2007). Furthermore, outside of research-related activities it is rare that a school-based practitioner would be expected to collect data on a number of OIB or on AB. Collection of data on behaviors other than the TB is likely neglected by practitioners due to a combination of the added strain such a task would place on school resources and a lack of education on the rationale behind A-B-C continuous recording procedures.

Environmental events. Regarding participants' recording of environmental events, participants were most accurate in recording attention directed towards the target student. Teacher attention is an overt environmental event and is likely easier to capture than a covert event, such as escape from work (ESC), which participants were the least accurate in recording. The operational definition for ESC, which was provided to participants during training, encompassed a wide range of student behaviors: leaving the area, putting head down, and engaging in OIB. It is possible that participants were more likely to attend to the physical act of the student leaving the area as opposed to the subtler ways the target student escaped instruction. As there was a range in participants' experience with FBAs it is possible that some participants were unfamiliar with the functional definition of ESC as outlined in the behavior analytic literature, wherein individuals engage in any number of behaviors to get out of, or delay, an aversive stimulus (Cooper et al., 2007). The author attempted to control for this possibility by providing a definition of ESC together with descriptions of examples and non-examples of ESC, as well as examples of ESC in the training video they watched prior to the test video. Despite these efforts, participants were unable to accurately record this environmental event.

Accuracy across time. Although significant differences in accuracy across the minutes of the video were observed, participants did not become consistently more or less accurate over

time. Results indicated that participants were most accurate in their recording of the first minute, with accuracy of the first minute exceeding each of the remaining nine minutes. Their accuracy was lowest for the final minute, with their accuracy for all of the minutes being greater than the last minute. It is possible participants were most accurate scoring the first minute because they had just watched the instructional video and least accurate scoring the last minute because of fatigue. Another possible explanation for the differences in accuracy across minutes is the frequency and/or the order in which certain student behaviors and environmental events appeared in the video. For example, there were no instances of OIB in the first minute and only one instance of ESC—two events which were associated with lower accuracy scores. Conversely, there was one instance of OIB and three instances of ESC in min 10. Unfortunately, this video was not developed to systematically examine the factors impacting accuracy for each minute. If future research were to examine this research question, the specific types of student behaviors and environmental events would have to be systematically distributed across the minutes of the video. Although the total instances of problem behavior (TB + OIB) within each minute were controlled for in this study, the specific types of student behaviors were not controlled for, nor were the number and types of environmental events within each minute.

Accuracy of function. Despite the unacceptable accuracy in recording the occurrence and sequence of environmental events and student behaviors, the majority of participants selected the correct function of behavior (ESC). This finding suggests participants were not establishing their decisions solely on the data. It is also possible that the TB, leaving the instructional area, primed participants toward identifying ESC as the function due to a perception that leaving the instructional environment was synonymous with ESC, despite the attention provided as a distractor consequence on several occasions. It is interesting to note that several of

the minutes that participants were least accurate in recording contained more OIB-ESC sequences than TB-ESC sequences, which could possibly be due to participants being more likely to record ESC if it followed the overt act of the student leaving the instructional area.

School-based Implementation

Bias as a function of condition. It is common practice in schools for practitioners to use indirect assessment methods before conducting direct observations (Floyd et al., 2005; Hanley, 2012). This practice introduces the possibility of confirmation bias, with observers searching for evidence to support the function identified in the interview or rating scale and ignoring conflicting information. The accuracy scores for the Informed condition and the Neutral condition did not differ significantly using both the adjusted and unadjusted accuracy scores. Participants who were primed with the incorrect function of attention were not more likely to select attention as the primary function over the true function of ESC. However, those participants primed with the incorrect function of attention were significantly more likely to select attention as a secondary function, as opposed to participants who were not provided with the hypothesis that attention was maintaining the TB. Taken together, these findings provide minimal evidence that participants were biased by the hypothesis they received. Additional research is needed to determine the extent to which school psychologists are influenced by information collected earlier in the assessment process and how that might impact treatment decisions.

Standard practice components. Along with examining the accuracy of this A-B-C recording method and the accuracy with which participants identified the function of a TB; this study sought to gain a broader understanding of how school psychologists are conducting FBAs. As there is confusion as to what is required for an FBA in both federal and state guidelines, there

is a need to gain an understanding of school-based practitioners' application of the mandate. Results from this study suggests that the participants were not incorporating many standard practice components in their FBAs. The implications of this finding are unclear, as there is limited direction provided by the federal and state governments as to what should be included in an FBA (Asmus et al., 2002; Stage et al., 2006; Zirkel, 2011). However, the number of due process hearings schools have been involved in over absent or inadequate FBAs (i.e., FBAs with limited data sources) since the IDEA amendments suggests that a failure to conduct a thorough FBA can have serious legal implications (Drasgow & Yell, 2001; Poucher, 2016).

Alarmingly, components which are indisputably essential for the development of an FBA, such as hypothesis development and direct observation of the TB, were not included in every FBA conducted by this sample of school psychologists and school psychology graduate students (75% and 66%, respectively). Perhaps more troubling is the finding that participants were not regularly conducting FBAs. One-third of participants reported having completed fewer than five FBAs in their career. Anecdotal information provided by the participants in this study suggested that in their school districts' classroom teachers were typically responsible for conducting the FBAs.

Experience. Prior experience with conducting FBAs did not significantly predict participants' accuracy of recording; however, this is not surprising as participants only reported incorporating direct observation in 66% of their most recent FBAs. Interestingly, there was a negative correlation between years of experience practicing as a school psychologist and accuracy of recording. The longer a participant's tenure as a practicing school psychologist, the less accurate they were with the procedure. Assuming that most school psychologists learn to conduct behavioral observations while in graduate school, it is feasible that the further removed

participants are from instruction in these methods the less likely they are to retain and use the strategies in their practice. Regardless of experience, participants are unlikely to implement experimental functional analyses, the gold standard for identifying the function of behavior, in their FBAs. This is a finding consistent with prior research (Blood & Neel, 2007; Doggett et al., 2001).

Limitations

Findings from this study should be interpreted with caution due to a number of associated limitations. First, participants were recruited from only three states in the Southeast (Georgia, Tennessee, and South Carolina), with the majority of participants working in school districts in Georgia (67%); however, it is likely that these participants matriculated from a variety of training programs. Although participants were recruited from a total of 13 different school districts, a second limitation is that two districts provided half of the participants (50%) and school districts dictate practice as they control the resources and the training provided to their practitioners. Research suggests that many school districts follow a "cookbook approach" to conducting FBAs, wherein school personnel are provided with a set number of components to complete in a particular sequence (Weber et al., 2005). If the districts represented in this study followed a similar approach to FBA implementation, it would impact the data collected on the standard practice components each participant includes in their FBAs. Similarly, if one school district in our sample failed to provide participants with a thorough understanding of FBAs, it is likely all participants within that district were similarly disadvantaged.

A third limitation is that participants were taught to use Serial Recording form by watching a 12-min training video. Other researchers incorporated more extensive training before and throughout their studies; such as having participants watch a 60-min lecture prior to

recording (Lerman et al., 2009) or receiving written task clarification and verbal performance feedback (Mayer & Reed, 2013). A fourth limitation is that the findings from this study are based on participants' observations of one 10-min sample of behavior. With this limited sample, it cannot be said that accuracy in recording is irrelevant to accurate selection of function, despite participants' accurate selection of function in this study. This recording method needs to be examined with other TBs that are both overt and covert in nature, to determine the extent to which characteristics of the TB contribute to participants' ability or inability to accurately identify function. Additionally, although there is no standard as to what constitutes an adequate sample of behavior, there is research to suggest that for behavior that occurs with lower frequency and higher variability, as is characteristic of the behavior typically observed in schools, multiple observations of longer duration are often required to achieve an adequate representation of behavior (Hintze & Matthews, 2004; Tiger et al., 2013). Research is needed to see if this measure is able to be used accurately with multiple behavior samples that are longer than 10-min. A final limitation of this study is that school psychologists were selected as the population from which to sample for this study, with the assumption that school psychologists were regularly implementing FBAs in their school districts. Surprisingly this was not the case. Data evaluating the frequency with which participants were conducting FBAs indicated that 40% of participants had not conducted an FBA within the past three months, and 17% of participants reported that they had never conducted an FBA. On average, participants conducted two FBAs within the past three months.

School psychologists are among the most appropriate candidates for conducting FBAs in schools as they are in the unique position of being an impartial, third-party in the classroom and they likely received training in behavioral assessment in their graduate training (Scott & Kamps,

2007; Van Acker et al.,2005). Despite this fact, anecdotal data from this study suggested that special education and/or regular education teachers were primarily the individuals conducting FBAs, and school psychologists were responsible for synthesizing the information once it was collected. This is problematic for two reasons. First, there is research suggesting that teachers lack training in FBA-related procedures (Couvillon et al., 2009) and second, if teachers are collecting A-B-C recording they are collecting data on their own behavior. When individuals record or monitor their own behavior it changes their behavior (Lipinski & Nelson, 1974), which limits one's ability to collect an accurate and unbiased sample of the environment. It is possible that teachers can be trained to use A-B-C recording to collect data on student behavior; however, teachers should not collect data on their own behavior. Direct observation methods, like A-B-C recording, are considered standard practice for FBAs in schools; however, if the individuals charged with conducting those observations are not accurate in using the available tools, those observations are not likely to generate the data needed to design function-based treatment.

When examining the accuracy of recording from this study alongside the findings from prior research (Lerman et al., 2009; Mayer & Reed, 2013), it is clear that practitioners have difficulty recording student behavior and environmental events accurately—regardless of their position (teacher vs. school psychologist) or the method of recording (discontinuous vs. continuous). There needs to be sufficient training before school psychologists, or any school-based practitioner, can be expected to the Serial Recording form, or any form of continuous recording, as a clinical tool in schools. There is research to suggest that continuous recording procedures provide certain advantages over discontinuous recording (Skinner et al., 2000; Thompson & Borrero, 2011), especially in the area of data analysis and monitoring, but collecting this information accurately is a difficult skill to master (Thompson & Borrero, 2011).

In general, participants did not believe the Serial Recording form was easy to use, and therefore reported that they were unlikely to use it in the future. Despite reporting not wanting to use the Serial Recording form, participants believed it was helpful for determining the function of behavior. It is possible that participants saw the value in a tool that encouraged them to take into account the full scope and sequence of environmental events and behavior; however, this alone was unlikely to motivate participants to use the tool in practice unless they could be trained to be more accurate in their recording.

Directions for Future Research

There is some research to suggest that task clarification plus feedback improves the accuracy of A-B-C structured recording (Mayer & Reed, 2013). Research is needed to determine whether additional training can have the same effect on the Serial Recording form. Research is also needed to determine the full extent to which this form of recording could be used in practice. It is possible the Serial Recording form could be used as a training tool to teach practitioners how to observe and record the flow of classroom events, rather than restrict their focus to isolated events. It could be a method for collecting data on the ecological context in which behavior occurs by providing a ratio of praise to reprimands or time spent in whole group versus small group instruction. Finally, research is needed to see whether or not the Serial Recording form lends itself to more complex analyses, such as the calculation of conditional probabilities.

Beyond the implications for the Serial Recording form, the real implication from this study is that practitioners need more training in function-based assessment and treatment. Even if practitioners received training on behavioral assessment in graduate school, it was seemingly insufficient to sustain them in their practice for an entire career. There needs to be frequent, recurring professional development in order to instill a true understanding of the importance of

identifying behavioral function over simply identifying behaviors and then instruction on how to effectively use that information to guide treatment. A mandate is useless if those charged with its execution do not see its value or have the support to use it effectively.

CHAPTER 4

CONCLUSIONS

The ambiguity surrounding IDEA's definition of an FBA, as well as its parameters for use, has left individual states responsible for setting guidelines for school districts to follow. However, one examination of the state statutes and regulations specific to FBAs found that only 17 state laws define FBAs and those definitions lack specificity, leaving room for misinterpretation (Zirkel, 2011). Without clear guidelines for implementation there is great potential for FBAs to be implemented inadequately. Unfortunately, this concern has proved true, with findings from the studies examining the use of FBAs in schools suggesting that there are multiple flaws with their current implementation (Blood & Neel, 2007; Van Acker et al., 2005). In an attempt to address this problem, several researchers have reviewed the special education literature and published standard practice components which, if included and implemented correctly, would result in an appropriate and "legally defensible" FBA (Weber et al., 2005; Zirkel, 2011).

Even if the existing lists of standard practice components were widely distributed and uniformly adopted, there is still the problem that some of the components considered "standard practice" do not have sufficient research examining their quality as *behavioral* assessment methods—specifically, research examining their accuracy, sensitivity, and treatment utility (Hayes et al., 1986). This is problematic. When conducting FBAs, school-based practitioners are most likely to use rating scales and some form of direct observation such as A-B-C recording (Anderson et al., 2006; Eckert et al., 2005; Floyd et al., 2005; Gable et al., 2014) likely due to the

lack of training and resources required. The research evaluating rating scales has evaluated their quality from a psychometric perspective (Koristas & Iacono, 2013; Paclawskyj et al., 2000; Shogren & Rojahn, 2003), which is not appropriate for use with behavioral assessment data as any variability in the data may be a function of a true change in behavior rather than a measurement flaw. The research examining the quality of direct FBA methods has mostly focused on A-B-C recording (Ellingson et al., 2000; Lerman et al., 2009; Mayer & Reed, 2013; VanDerHeyden et al., 2001), a recording method popular with school-based practitioners (Eckert et al., 2005; VanDerHeyden et al., 2011). In the two studies evaluating the accuracy of commonly used forms of A-B-C recording (Lerman et al., 2009; Mayer & Reed, 2013), participants did not achieve acceptable levels of accuracy in their scoring. If the quality of some of these assessment methods is questionable, then distribution of a list of standard practice components is premature. There is a clear need for more research evaluating the quality of behavioral assessment methods in a way that takes into account the principle tenant of behavioral assessment--that consistency in behavior across time or settings cannot be assumed. As direct FBA methods, such as A-B-C recording, allow for direct observation of the behavior in the natural environment, this type of assessment method was selected for further research.

Chapter 3 described a study that evaluated the accuracy the Serial Recording form, a hybrid form of A-B-C recording that combined the practitioner-friendly format of structured recording with the superior analytic properties of continuous recording. Unlike previous studies, this study used school psychologists as participants rather than teachers as school psychologists can serve as an impartial third-party in the classroom and they likely received training in behavioral assessment in their graduate training (Scott & Kamps, 2007; Van Acker et al., 2005). The findings from this study as well as the other studies examining accuracy suggested that

practitioners have difficulty recording student behavior and environmental events accurately, regardless of recording method or observer.

Although it is reasonable to assume that practitioners can be trained to collect data accurately once high-quality assessment methods have been established in the research, a top-down approach of training practitioners on a prescribed set of assessment methods seems a small fix to a systemic problem. There appears to be a fundamental lack of understanding and training on function-based assessment and treatment. Function-based treatments are more likely to prevent the occurrence of future problem behavior (Cooper et al., 2007) however, if practitioners do not understand the reasons behind identifying the function of behavior it seems unlikely they would they have the motivation to seek out and use quality function-based assessment methods in practice. Even if practitioners received training on behavioral assessment in graduate school, as many of our school psychology participants did, it was seemingly insufficient to sustain them in their practice for an entire career. There needs to be frequent, ongoing professional development in order to instill a true understanding of the importance of identifying behavioral function over identifying behaviors. Practitioners must then receive guidance on how to effectively use that information when designing a treatment plan.

It should be further specified that these trainings should be targeted towards those individuals for whom it would be appropriate to conduct an FBA. Namely, school psychologists or other school-based professionals. These individuals would not be involved in the behavioral contingency they are responsible for recording. Anecdotal data from this study suggested that special education and/or regular education teachers were primarily responsible for conducting FBAs in their school districts. One-third of school psychology participants reported having completed fewer than five FBAs in their career. This is a problem that is best addressed through

a combination of professional development and increased clarity in state and federal laws regarding FBAs.

Chapter 3 sought to address one problem identified in Chapter 2—the lack of research evaluating the quality of one type of assessment method considered standard practice for use in an FBA (Weber et al., 2005). More research must be conducted and those results must be communicated not only to lawmakers, but directly to school districts. It has long been said that if you want to change child behavior, you change adult behavior. The findings gleaned from this dissertation suggest that the research community, district-level administrators, and lawmakers need to get better at changing adult behavior.

REFERENCES

Alberto, P. A., & Troutman, A. C. (2006). *Applied behavior analysis for teachers (7th ed.)*. Upper Saddle River, NJ: Pearson.

Allday, R. A., Nelson, J. R., & Russel, C. S. (2011). Classroom-based functional behavioral assessment: Does the literature support high fidelity implementation? *Journal of Disability Policy Studies, 22,* 140-149.

Alter, P. J., Conroy, M. A., Mancil, G. R., & Haydon, T. (2008). A comparison of functional behavior assessment methodologies with young children: Descriptive methods and functional analysis. *Journal of Behavioral Education, 17,* 200-219.

Anderson, C. M., English, C. L., & Hedrick, T. M. (2006). Use of the structured descriptive assessment with typically developing children. *Behavior Modification, 30,* 352-378.

Anderson, C. M., & Long, E. S. (2002). Use of a structured descriptive assessment methodology to identify variables affecting problem behavior. *Journal of Applied Behavior Analysis, 35,* 137-154.

Anderson, C. M., Rodriguez, B., & Campbell, A. (2015). Functional behavior assessment in schools: Current and future directions. *Journal of Behavioral Education, 24,* 338-371.

Arndorfer, R. E., Miltenberger, R. G., Woster, S. H., Rortvedt, A. K., & Gaffaney, T. (1994). Home-based descriptive and experimental analysis of problem behaviors in children. *Topics in Early Childhood Special Education, 14,* 64-87.

Asmus, J. M., Vollmer, T. R., & Borrero, J. C. (2002). Functional behavioral assessment: A school based model. *Education and Treatment of Children, 25,* 67-90.

Barton-Arwood, S. M., Wehby, J. H., Gunter, P. L., & Lane, K. L. (2003). Functional behavior assessment rating scales: Intrarater reliability with students with emotional or behavioral disorders. *Behavioral Disorders, 28,* 386-400.

Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis, 1,* 175-191.

Blood, E., & Neel, R. S. (2007). From FBA to implementation: A look at what is actually being delivered. *Education and Treatment of Children, 30,* 67-80.

Bloom, S. E., Iwata, B. A., Fritz, J. N., Roscoe, E. M., & Carreau, A. B. (2011). Classroom application of a trial-based functional analysis. *Journal of Applied Behavior Analysis, 44,* 19-31.

Calloway, C. J., & Simpson, R. L. (1998). Decisions regarding functions of behavior: Scientific versus informal analyses. *Focus on Autism and Other Developmental Disabilities, 13,* 167-175.

Camp, E. M., Iwata, B. A., Hammond, J. L., & Bloom, S. E. (2009). Antecedent versus consequent events as predictors of problem behavior. *Journal of Applied Behavior Analysis, 42,* 469-483.

Campbell, J. M. (2003). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: A quantitative synthesis of single-subject research. *Research in Developmental Disabilities, 24,* 120-138.

Chandler, L. K., & Dahlquist, C. M. (2010). *Functional assessment: Strategies to prevent and remediate challenging behavior in school settings (3rd Ed.)*. Upper Saddle River, NJ: Pearson.

Cipani, E., & Schock, K. M. (2007). *Functional behavioral assessment, diagnosis, and treatment*. New York: Springer.

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

Couvillon, M. A., Bullock, L. M., & Gable, R. A. (2009). Tracking behavior assessment methodology and support strategies: A national survey of how schools utilize functional behavioral assessments and behavior intervention plans. *Emotional and Behavioural Difficulties, 14,* 215-228.

Doggett, R. A., Edwards, R. P., Moore, J. W., Tingstrom, D. H., & Wilczynski, S. M. (2001). An approach to functional assessment in general education classroom settings. *School Psychology Review, 30,* 313-328.

Drasgow, E., & Yell, M. L. (2001). Functional behavioral assessments: Legal requirements and challenges. *School Psychology Review, 30,* 239-251.

Durand, V. M., & Crimmins, D. B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18,* 99-117.

Eckert, T. L., Martens, B. K., & DiGennaro, F. D. (2005). Describing antecedent-behavior-consequence relations using conditional probabilities and the general operant contingency space: A preliminary investigation. *School Psychology Review, 34,* 520-528.

Ellingson, S. A., Miltenberger, R. G., Galensky, T. L., & Garlinghouse, M. (2000). Functional assessment and intervention for challenging behaviors in the classroom by general classroom teachers. *Journal of Positive Behavior Interventions, 2,* 85-97.

Ervin, R. A., Ehrhardt, K. E., & Poling, A. (2001). Functional assessment: Old wine in new bottles. *School Psychology Review, 30,* 173-179.

Fisher, W. W., Adelinis, J. D., Thompson, R. H., Worsdell, A. S., & Zarcone, J. R. (1998). Functional analysis and treatment of destructive behavior maintained by termination of "don't" (and symmetrical "do") requests. *Journal of Applied Behavior Analysis, 31,* 339-356.

Floyd, R. G., Phaneuf, R. L., & Wilczynski, S. M. (2005). Measurement properties of indirect assessment methods for functional behavioral assessment: A review of research. *School Psychology Review, 34,* 58-73.

Gable, R. A., Park, K. L., & Scott, T. M. (2014). Functional behavioral assessment and students at risk for or with emotional disabilities: Current issues and considerations. *Education and Treatment of Children, 37,* 111-135.

Georgia Department of Education. (2016). *Special Education Rules Implementation Manual*. Retrieved from (http://www.gadoe.org/Curriculum-Instruction-and-Assessment/Special-Education-Services/Documents/Sp%20Ed%20Implementation%20Manual%20Part%201_8-12.pdf)

Gresham, F. M., Watson, T. S., & Skinner, C. H. (2001). Functional behavioral assessment: Principles, procedures, and future directions. *School Psychology Review, 30,* 156-172.

Hanley, G. P. (2012). Functional assessment of problem behavior: Dispelling myths, overcoming implementation obstacles, and developing new lore. *Behavior Analysis in Practice, 5,* 54-72.

Hanley, G. P., Jin, C. S., Vanselow, N. R., & Hanratty, L. A. (2014). Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments. *Journal of Applied Behavior Analysis, 47,* 16-36.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1986). Evaluating the quality of behavioral assessment. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 463-503). New York: Guilford Press.

Herzinger, C. V., & Campbell, J. M. (2007). Comparing functional assessment methodologies: A quantitative synthesis. *Journal of Autism and Developmental Disorders, 37,* 1430-1445.

Hintze, J. M., & Matthews, W. J. (2004). The generalizability of systematic direct observations across time and setting: A preliminary investigation of the psychometrics of behavioral observation. *School Psychology Review, 33,* 258-270.

Individuals with Disabilities Education Act Amendments of 1997, 20 U.S.C. § 1400 (1997).

Individuals with Disabilities Education Improvement Act, 20 U.S.C. § 1400. (2004).

Iwata, B. A., DeLeon, I. G., & Roscoe, E. M. (2013). Reliability and validity of the functional analysis screening tool. *Journal of Applied Behavior Analysis, 46,* 271-284.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982). Toward a functional analysis of self-injury. *Analysis and Intervention in Developmental Disabilities, 2,* 3-20.

Kahng, S., Iwata, B. A., Fischer, S. M., Williams, D. E., & Smith, R. G. (1998). Temporal distributions of problem behavior based on scatter plot analysis. *Journal of Applied Behavior Analysis, 31,* 593-604.

Kern, L., Hilt, A. M., & Gresham, F. (2004). An evaluation of the functional behavioral assessment process with students with or at risk for emotional behavioral disorders. *Education and Treatment of Children, 27,* 440-452.

Koritsas, S., & Iacono, T. (2013). Psychometric comparison of the motivation assessment scale (MAS) and the questions about behavioral function (QABF). *Journal of Intellectual Disability Research, 57,* 747-757.

Kozlowski, A. M., & Matson, J. L. (2012). Interview and observation methods in functional assessment. In J. L. Matson (Ed.). *Functional assessment for challenging behaviors* (pp. 105-124). New York: Springer.

Lanovaz, M. J., Argumedes, M., Roy, D., Duquette, J. R., & Watkins, N. (2013). Using ABC narrative recording to identify the function of problem behavior: A pilot study. *Research in Developmental Disabilities, 34,* 2734-2742.

Lerman, D. C., & Iwata, B. A. (1993). Descriptive and experimental analyses of variables maintaining self-injurious behavior. *Journal of Applied Behavior Analysis, 26,* 293-319.

Lerman, D. C., Hovanetz, A., Strobel, M., & Tetreault, A. (2009). Accuracy of teacher-collected descriptive analysis data: A comparison of narrative and structured recording formats. *Journal of Behavioral Education, 18,* 157-172.

Lipinski, D., & Nelson, R. (1974). The reactivity and unreliability of self-recording. *Journal of Consulting and Clinical Psychology, 42,* 118-123.

Love, J. R., Carr, J. E., Almason, S. M., & Petursdottir, A. I. (2009). Early and intensive behavioral intervention for autism: A survey of clinical practices. *Research in Autism Spectrum Disorders, 3,* 421-428.

Mace, F. C., Lalli, J. S., & Lalli, E. P. (1991). Functional analysis and treatment of aberrant behavior. *Research in Developmental Disabilities, 12,* 155-180.

Martens, B. K., DiGennaro, F. D., Reed, D. D., Szczech, F. M., & Rosenthal, B. D. (2008). Contingency space analysis: An alternative method for identifying contingent relations from observational data. *Journal of Applied Behavior Analysis, 41,* 69-81.

Matson, J. L., & Minshawi, N. F. (2007). Functional assessment of challenging behavior: Toward a strategy for applied settings. *Research in Developmental Disabilities, 28,* 353-361.

Matson, J. L., & Vollmer, T. R. (1995). *The questions about behavioral function (QABF) user's guide*. Baton Rouge, LA: Scientific Publishers.

Mayer, K. L., & Reed, F. D. G. (2013). Effects of a training package to improve the accuracy of descriptive analysis data recording. *Journal of Organizational Behavior Management, 33,* 226-243.

McDougal, J. L., Chafouleas, S., & Waterman, B. B. (2006). *Functional behavioral assessment and intervention in schools: A practitioner's guide.* Champaign, IL: Research Press.

McIntosh, K., Borgmeier, C., Anderson, C. M., Rodriguez, B., & Tobin, T. J. (2008). Technical adequacy of the functional assessment checklist: Teachers and staff (FACTS) FBA interview measure. *Journal of Positive Behavior Interventions, 10,* 33-45.

McKerchar, P. M., & Thompson, R. H. (2004). A descriptive analysis of potential reinforcement contingencies in the preschool classroom. *Journal of Applied Behavior Analysis, 37,* 431-444.

National Association of State Directors of Special Education. (1998). *Guidance on functional behavioral assessments for students with disabilities*. Retrieved from http://www.p12.nysed.gov/specialed/publications/policy/functionbehav.htm

Nelson, R. O., & Hayes, S. C. (1979). Some current dimensions of behavioral assessment. *Behavioral Assessment, 1,* 1-16.

Newcomer, L. L., & Lewis, T. J. (2004). Functional behavioral assessment: An investigation of assessment reliability and effectiveness of function-based interventions. *Journal of Emotional and Behavioral Disorders, 12,* 168-181.

Paclawskyj, T. R., Matson, J. L., Rush, K. S., Smalls, Y., & Vollmer, T. R. (2000). Questions about behavioral function (QABF): A behavioral checklist for functional assessment of aberrant behavior. *Research in Developmental Disabilities, 21,* 223-229.

Pence, S. T., Roscoe, E. M., Bourret, J. C., & Ahearn, W. H. (2009). Relative contributions of three descriptive methods: Implications for behavioral assessment. *Journal of Applied Behavior Analysis, 42,* 425-446.

Poucher, S. M. (2016). The road to prison is paved with bad evaluations: The case for functional behavioral assessments and behavior intervention plans. *American University Law Review, 65,* 471-523.

Rooker, G. W., DeLeon, I. G., Borrero, C. S. W., Frank-Crawford, M. A., & Roscoe, E. M. (2015). Reducing ambiguity in the functional assessment of problem behavior. *Behavioral Interventions, 30,* 1-35.

Rooker, G. W., Iwata, B. A., Harper, J. M., Fahmie, T. A., & Camp, E. M. (2011). False-positive tangible outcomes of functional analyses. *Journal of Applied Behavior Analysis, 44,* 737-745.

Sasso, G. M., Reimers, T. M., Cooper, L. J., Wacker, D., Berg, W., Steege, M., Kelly, L., & Allaire, A. (1992). Use of descriptive and experimental analyses to identify the functional properties of aberrant behavior in school settings. *Journal of Applied Behavior Analysis, 25,* 809-821.

Scott, T. M., Bucalos, A., Liaupsin, C., Nelson, C. M., Jolivette, K., & DeShea, L. (2004). Using functional behavior assessment in general education settings: Making a case for effectiveness and efficiency. *Behavioral Disorders, 29,* 189-201.

Scott, T. M., Liaupsin, C. J., Nelson, C. M., & Jolivette, K. (2003). Ensuring student success through team-based functional behavioral assessment. *TEACHING Exceptional Children, 35,* 16-21.

Scott, T. M. & Kamps, D. M. (2007). The future of functional behavioral assessment in school settings. *Behavioral Disorders, 32,* 146-157.

Shapiro, E. S. (2003). Behavioral observation of students in schools (BOSS). *Computer Software.* San Antonio, TX: Psychological Corporation.

Shapiro, E. S., & Kratochwill, T. R. (Eds.). (2000). *Behavioral assessment in schools: Theory, research, and clinical foundations* (2nd ed.). New York: Guilford.

Shippen, M. E., Simpson, R. G., & Crites, S. A. (2003). A practical guide to functional behavioral assessment. *TEACHING Exceptional Children, 35,* 36-44.

Shogren, K. A., & Rojahn, J. (2003). Convergent reliability and validity of the questions about behavioral function and the motivation assessment scale: A replication study. *Journal of Developmental and Physical Disabilities, 15,* 367-375.

Shriver, M. D., Anderson, C. M., & Proctor, B. (2001). Evaluating the validity of functional behavior assessment. *School Psychology Review, 30,* 180-192.

Skinner, C. H., Dittmer, K. I., & Howell, L. A. (2000). Direct observation in school settings: Theoretical issues. In E. S. Shapiro & T. R. Kratochwill (Eds.). *Behavioral assessment in schools: Theory, research, and clinical foundations* (2nd ed.; pp. 19-46). New York: Guilford.

Sloman, K. N. (2010). Research trends in descriptive analysis. *The Behavior Analyst Today, 11,* 20-35.

Solnick, M. D. & Ardoin, S. P. (2010). A quantitative review of functional analysis procedures in public school settings. *Education and Treatment of Children, 33,* 153-175.

Stage, S. A., Jackson, H. G., Moscovitz, K., Erickson, M. J., Thurman, S. O., Jessee, W., & Olson, E. M. (2006). Using multi-method-multisource functional behavioral assessment for students with behavioral disabilities. *School Psychology Review, 35,* 451-471.

Stichter, J. P., & Conroy, M. A. (2005). Using structural analysis in natural settings: A responsive functional assessment strategy. *Journal of Behavioral Education, 14,* 19-34.

Tarbox, J., Wilke, A. E., Najdowski, A. C., Findel-Pyles, R. S., Balasanyan, S., Caveney, A. C., Chilingaryan, V., King, D. M., Niehoff, S. M., Slease, K., & Tia, B. (2009). Comparing indirect, descriptive, and experimental functional assessments of challenging behavior in children with autism. *Journal of Developmental and Physical Disabilities, 21,* 493-514.

Thompson, R. H., & Borrero, J. C. (2011). Direct observation. In W. W. Fisher, C. C. Piazza, & H. S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 191-205). New York: Guilford.

Thompson, R. H., & Iwata, B. A. (2007). A comparison of outcomes from descriptive and functional analyses of problem behavior. *Journal of Applied Behavior Analysis, 40,* 333-338.

Tiger, J. H., Hanley, G. P., & Bessette, K. K. (2006). Incorporating descriptive assessment results into the design of a functional analysis: A case example involving a preschooler's hand mouthing. *Education and Treatment of Children, 29,* 107-124.

Tiger, J. H., Miller, S. J., Mevers, J. L., Mintz, J. C., Scheithauer, M. C., & Alvarez, J. (2013). On the representativeness of behavior observation samples in classrooms. *Journal of Applied Behavior Analysis, 46,* 424-435.

Touchette, P. E., MacDonald, R. F., & Langer, S. N. (1985). A scatter plot for identifying stimulus control of problem behavior. *Journal of Applied Behavior Analysis, 18,* 343-351.

U.S. Department of Education, Office of Special Education and Rehabilitative Services. (1999). *Applying positive behavioral support and functional behavior assessment in schools.* (OSEP Center on Positive Behavioral Interventions and Supports Publication Technical Assistance Guide #1). Retrieved from http://www.cesa7.org/sped/discoveridea/topdocs/pbis.pdf.

Van Acker, R., Boreson, L., Gable, R. A., & Potterton, T. (2005). Are we on the right course? Lessons learned about current FBA/BIP practices in schools. *Journal of Behavioral Education, 14,* 35-56.

VanDerHeyden, A. M., Witt, J. C., & Gatti, S. (2001). Descriptive assessment method to reduce overall disruptive behavior in a preschool classroom. *School Psychology Review, 30,* 548-567.

Vollmer, T. R., Borrero, J. C., Wright, C. S., Van Camp, C., & Lalli, J. S. (2001). Identifying possible contingencies during descriptive analyses of severe behavior disorders. *Journal of Applied Behavior Analysis, 34,* 269-287.

Von Ravensburg, H. & Tobin, T. J. (2008). IDEA 2004: Final regulations and the reauthorized functional behavioral assessment. Available at SSRN: https://ssrn.com/abstract=1151394.

Watson, T., & Steege, M. W. (2003). *Conducting school-based functional behavioral assessments: A practitioner's guide.* New York: Guilford.

Weber, K. P., Killu, K., Derby, K. M., & Barretto, A. (2005). The status of functional behavioral assessment (FBA): Adherence to standard practice in FBA methodology. *Psychology in the Schools, 42,* 737-744.

Zirkel, P. A. (2011). Case law for functional behavior assessments and behavior intervention plans: An empirical analysis. *Seattle University Law Review, 35,* 175-212.

APPENDIX A

BACKGROUND QUESTIONNAIRE

1. Which of these best represents your gender?

   Male Female

2. Which of the following best represents your race?

   African-American Asian Caucasian Hispanic Other

3. How many years have you been a practicing school psychologist? _____

4. Please list the school district in which you are currently employed._____

5. What is your terminal degree? (e.g., Eds., PhD., MA) _____

6. Do you have your NCSP?

   Yes No

7. Have you completed any graduate-level coursework in applied behavior analysis and/or behavioral assessment?

   Yes No

8. How many hours would you estimate you have had, either in coursework or professional development, in behavioral assessment or FBA-related training? _____

9. Approximately, how many Functional Behavior Assessments (FBAs) did you conduct in the past three months? _____

10. Out of the past five FBAs you've conducted, please make your best guess as to how many included each of the following components. If you have conducted less than five FBAs, how many have you conducted? _____? Please refer to this number for the following question.

    a. Definition or identification of the target behavior: _____
    b. Review of Records: _____
    c. Checklist data: _____

d. Student interview: _____
e. Other interview types: _____
f. Team meetings: _____
g. Direct observation with no manipulation: _____
h. Scatterplot: _____
i. Antecedent-Behavior-Consequence (ABC) analysis: _____
j. Functional Analysis Observation form: _____
k. Reinforcer identification: _____
l. Ecological context: _____
m. Development of hypothesis: _____
n. Experimental manipulation (EFA): _____

APPENDIX B

BACKGROUND QUESTIONNAIRE FOR GRADUATE STUDENTS

11. Which of these best represents your gender?

    Male Female

12. Which of the following best represents your race?

    African-American Asian Caucasian Hispanic Other

13. How many years of graduate work have you completed? _____

14. Please list the university in which you are currently enrolled._____

15. What degree in school psychology are you currently pursuing? (e.g., Eds., PhD.)

    _____

16. Do you have your Masters degree?

    Yes No

17. What is your anticipated graduation date? _____

18. Have you completed any graduate-level coursework in applied behavior analysis and/or behavioral assessment?

    Yes No

19. How many hours would you estimate you have had, either in coursework or professional development, in behavioral assessment or FBA-related training? _____

20. Approximately, how many Functional Behavior Assessments (FBAs) did you conduct in the past three months? _____

21. Out of the past five FBAs you've conducted, please make your best guess as to how many included each of the following components. If you have conducted less than five

FBAs, how many have you conducted? _____? Please refer to this number for the following question.

a. Definition or identification of the target behavior: _____
b. Review of Records: _____
c. Checklist data: _____
d. Student interview: _____
e. Other interview types: _____
f. Team meetings: _____
g. Direct observation with no manipulation: _____
h. Scatterplot: _____
i. Antecedent-Behavior-Consequence (ABC) analysis: _____
j. Functional Analysis Observation form: _____
k. Reinforcer identification: _____
l. Ecological context: _____
m. Development of hypothesis: _____
n. Experimental manipulation (EFA): _____

APPENDIX C

OPERATIONAL DEFINITIONS OF SETTING EVENTS AND ENVIRONMENTAL EVENTS

Whole group instruction: Whole group instruction should be coded when the teacher is engaged in instruction directed to the entire class. Examples include: Lecturing, moving through a PowerPoint, and writing examples on the board at the front of the class. Non-examples include: Sitting at her desk while the students work independently, working with a student one-on-one, instructing a small group of students (3 or fewer), or whole group directives (e.g., "take out your books).

Small group instruction: Small group instruction should be coded when the teacher is engaged in instruction directed to a small group of students (3 or fewer) in which the target student is included, or when students are working together in a small group (3 or fewer students). Examples include: Guiding a small group of students through a lesson, students working in groups of 3 or less to complete an instructional task. Non-examples include: Whole group instruction, working independently without teacher or peer assistance.

Independent work: Independent work should be coded when students are working on an assigned task independently, without teacher or peer assistance. Examples: Working on an assigned task independently, in the assigned instructional area. Non-examples include: A student engaging in problem behavior while sitting by himself.

Transition: A transition should be coded whenever a group of students move from one activity to another within the observation period. Note that a transition does not have to involve physical movement. Examples include: moving from whole class instruction to small groups or moving from independent work to whole class instruction. Non-examples include: a student moving from one area of the room to another to perform a brief task (e.g., sharpen a pencil, throw a piece of paper away), or a student leaving his seat without being instructed to do so.

Teacher directive to group (TD-G): A teacher directive should be coded whenever a teacher issues a command, question, or directive to the entire class, or to a group that includes the target student, that is related to instruction. To be coded as a separate event, there must be at least 3-s between directives. Examples include: "Let's do this problem together," "Take out your book and turn to page 28," "Who can tell me an answer to this problem?" Non-examples include: commands, reprimands, or praise statements directed specifically to the target student, asking a question to a specific student that is not the target student, speaking to a group of students in which the target student is not included, giving information to the class in which a specific action is not requested (e.g., "Today we're going to learn about fractions.")

Individual teacher attention-Negative (ITA-N): ITA-N should be coded when a teacher issues any negative verbal or non-verbal attention (for example, a stare, vocalization, or gesture)

directly to the target student that is meant to discourage or correct the current behavior of the student. To be coded as a separate event, there must be at least 3-s between events. Examples include: Looking directly at the student while he is engaging in problem behavior, giving the target student a command to stop a certain behavior or correct inappropriate behavior, sternly saying the student's name while he is engaging in problem behavior. Non-examples include: reprimands issued to a student who is not the target student, directives issued to the entire class.

Individual teacher attention-Other: ITA-O should be coded when a teacher issues any verbal or non-verbal attention directly to the target student that is neither punitive nor meant to discourage the target student's behavior. To be coded as a separate event, there must be at least 3-s between events. Examples include: Answering a question from the target student, asking the target student a question related to instruction, or praising the target student's on-task behavior or compliance. Non-examples include: Praise issued to a student who is not the target student, directives issued to the entire class.

Peer attention (PA): Peer attention should be coded whenever peers in the classroom are providing attention directly to the target student. Examples include: laughing at the target student's problem behavior, a conversation between the target student and a peer, peers looking at the student while the target student engages in problem behavior, or a peer telling the target student to stop doing a specific behavior. Non-examples include: peer attention directed towards another student in the classroom (not the target student), whole class choral responding.

Escape from work (ESC): Escape from work should be coded whenever work or instruction is removed by the teacher or delayed by the student as a consequence of student behavior. Examples include: the teacher taking away the target student's work or materials following an instance of problem behavior, or an action that inhibits the target student's ability to see or hear the instructional material (e.g., the target student blocking his eyes, talking to another peer while he should be working, or leaving the instructional area). Non-examples include: transitions from one activity to another.

APPENDIX D

OPERATIONAL DEFINITIONS OF BEHAVIOR

Target behavior (TB): Any instance when the target student leaves his assigned area during instruction and/or at a time when the directions are to remain in his area.

- Examples: leaving his seat to perform an activity unrelated to instruction when instruction is taking place (e.g., sharpening his pencil, throw away trash), leaving his seat to walk around the classroom when he was instructed to remain seated, leaving his seat to approach the teacher or peer during instruction.
- Non-examples: leaving his seat during a transition between activities, leaving his seat when instructed to do so by the teacher.

Other Inappropriate behavior (OIB): Any active off-task behaviors (motor activity or audible verbalizations) that are not directly associated with an assigned task, are not permitted, and prevent the child from receiving instruction. The target behavior (leaving the assigned area) should not be coded in this category. <u>If the target student is out of seat when he engages in other inappropriate behavior, only code the student as engaging in the target behavior.</u>

- Examples: Interrupting the teacher by calling out, doodling, manipulating objects not related to the academic task (e.g., twirling a pencil, playing with paper, eating), turning around in one's seat away from classroom instruction, putting one's head on the desk, and making any audible sound, such as humming/whistling/or tapping an item against his desk.
- Non-examples: Out of seat behavior (TB), sitting in his seat quietly watching the teacher (AB), or passively listening to other students talk about the assigned work in a cooperative learning group (AB).

Appropriate behavior (AB): Student is actively or passively attending to the assigned work and is not engaging in the target behavior or other inappropriate behavior.

- Examples include: Writing, raising a hand to answer a question, talking to the teacher about material, talking to a peer about the assigned material, sitting in his seat listening to a lecture, looking at an academic worksheet, reading assigned material silently, listening to a peer respond to a question, complying with a teacher directive.
- Non-examples include: Talking to others about nonacademic materials (OIB); walking to the pencil sharpener (TB), Calling out (OIB), doodling (OIB).

APPENDIX E

INSTRUCTIONS

Hello and thank you for agreeing to participate in this study. The purpose of this research is to determine whether the behavioral recording method I'm about to describe can be used to accurately identify the function of a child's behavior.

A-B-C recording is a data recording system that allows for the recording of student behavior and the sequence of surrounding events. These events include the antecedents to the behavior, or the things that precede it, and the consequences, or the events that immediately follow the behavior. There are different A-B-C recording methods, but for the purpose of this study I am using a structured form of A-B-C recording. This means I am going to provide you with a list of antecedents and consequences that typically occur in classrooms, as well as definitions of appropriate and inappropriate student behavior, and ask that you record the behavior and events as occurring when you observe them in the video. Our antecedents and consequences are going to be grouped together on the recording sheet under the heading "environmental events", so you do not have to worry about classifying them as one or the other. Just record them in the order you observe them.

Although some A-B-C recording procedures exclusively involve the recording of problem behavior, for the purposes of this study you will need to record both appropriate and inappropriate behavior along with the environmental events that precede or follow the student behavior. This data will provide a more thorough picture of the classroom environment and allow for the development of interventions that promote increases in appropriate behavior. For example, if we see in A-B-C recording that teacher attention usually follows problem behavior, we are likely to think that attention is maintaining the behavior. But if we see that teacher attention also follows appropriate behavior, we may need to look at other variables that might be maintaining the behavior.

We will now go over the different components on the recording sheet. The first column lists potential setting events. You will circle these setting events whenever they occur, and it is possible that multiple setting events may occur within the same minute. You will circle 'W' whenever the teacher is conducting whole group instruction. Circle 'S' if the teacher is instructing a small group of students (3 or fewer), or when students are working together in a small group (3 or fewer students per group). Circle 'I' if students are working independently. Finally, circle 'T' if a transition takes place within the interval. Operational definitions of the setting events can be found on your sheet entitled, "Operational Definitions of Setting Events and Environmental Events."

Now let's quickly go over the environmental events that you will record. [close up on recording sheet].

A teacher directive to the group should be coded whenever the teacher issues a command, question, or directive to the entire class, or to a group that includes the target student, that is related to instruction. Individual teacher attention-Negative should be coded when a teacher issues any negative verbal or non-verbal attention (for example, a stare, vocalization, or gesture) directly to the target student that is meant to discourage or correct the current behavior of the student. For example, telling the target student to sit down or yelling out his name when he is engaging in problem behavior. Individual teacher attention-other is pretty much any other verbal or non-verbal attention directed to the target student. For example, praising his on-task behavior or calling on him to answer a question.

With all of these teacher behaviors it is important to note that in order to be coded as a separate event, there must be at least 3-s between events. For example, the directive: "Ok class- you need to get out your books and turn to page 200," would be coded as one teacher directive to the group. Conversely, if the teacher said, "Johnny! [4-s] You need to sit down right now," that would be coded as two instances of individual teacher attention-Negative.

Escape from work should be coded whenever work or instruction is removed by the teacher or delayed by the student as a consequence of student behavior. This may be something very obvious, like the teacher taking away the target student's materials, or more discrete, like the target student putting his head down or talking when he is supposed to be working. To emphasize this point I would like you to watch the following video clips--one is an example of the target behavior and one an example of other inappropriate behavior. Although the behaviors are different, both result in the same consequence—Escape from work. We will discuss after you view the clips.

Play clips [10-s]

As you saw, the behaviors looked different; however, the consequence was the same. In the target behavior example, the student was able to delay participating in instruction by physically removing herself from the instructional area. In the other inappropriate behavior example, the student was similarly able to delay instruction—this time by occupying herself with food that is not allowed in class. Keep in mind throughout recording that it is possible for multiple environmental events to occur before and after student behavior.

Now let's review the definitions of student behavior you may observe and record. Please follow along on the sheet entitled, "Operational Definitions of Behavior" at the top.

The student's target behavior (TB) for this observation is "leaving the assigned area." It is defined here as "any instance when the target student leaves his assigned area during instruction or at a time when the directions are to remain in his area." Please refer to your worksheet for a comprehensive list of examples and non-examples of the target behavior.

Other inappropriate behavior (OIB) is defined as any active off-task behaviors (motor activity or audible verbalizations) that are not directly associated with an assigned task, are not permitted, and prevent the child from receiving instruction. The target behavior (leaving the assigned area)

should not be coded in this category. If the target student is out of his assigned area when he engages in other inappropriate behavior, only code the student as engaging in the target behavior.

If the student engages in several different types of OIB in quick succession—say he calls out and then several seconds later starts playing with materials unrelated to the instructional task--simply code OIB once and do not code another student behavior until the target student engages in the target behavior or appropriate behavior. For examples and non-examples of OIB behavior, please refer to your worksheet.

Finally, appropriate behavior (AB) is defined as any time when the student is actively or passively attending to the assigned work and is not engaging in the target behavior or other inappropriate behavior. For examples and non-examples of AB behavior, please refer to your worksheet.

Now that you are familiar with the behaviors and environmental events you will observe and record, let's go over the recording procedure itself. [Close up of data sheet]. Here is the data recording sheet. Let's first look across the columns. You can see the environmental events listed across the top of the sheet and the three types of student behaviors listed after those events. Now let's look down the rows. You can see that there is a row for each minute of the video—ten altogether. For each minute, you will circle the setting events that occur within that minute. You will begin each minute by recording the student behavior occurring at that moment. You will then continue to record the sequence of events in the row corresponding to the minute in which it occurs. The video will clearly state the start time for each minute so you know which row to record the sequences in.

As previously mentioned, for this A-B-C recording you are being asked to record both inappropriate and appropriate student behavior. If the target student is following directions when the first minute begins, you do not have to wait for problem behavior before you start recording. It is very important to record every event or behavior in the order in which it occurs. To accomplish this, you will need to use numbers to indicate the order in which the events and behavior occur. Let's watch a brief example and then score it together.

[VIDEO example plays]

In this video, whole group instruction is taking place so I'm going to quickly circle a 'W'. First, we look at the target student. We know from the description that Brooklyn is supposed to be sitting in her seat paying attention to the lesson. She is sitting at her desk eating chips, and is not paying attention to the teacher. So, we look on out data sheet and find 'OIB' and place a '1' right under it. As she is failing to attend to the teacher or the lesson she is escaping work, so we place a '2' right here under Escape from work. Another student noticed her eating the chips and asked her for some. This would be coded as peer attention, so we place a '3' under that column. Then, the teacher issued a reprimand, reminding them that there is no eating in class. So we put a '4' under individual teacher attention-Negative. Finally, the target student puts the chips away and turns to face the teacher, so we put a '5' under Appropriate Behavior. And there's our first behavioral sequence.

It's important to note that after one minute has elapsed, the behavioral sequence for the first minute ends, and a new sequence starts and we start numbering with 1 again. Each time we move down a row, and the student behavior occurring at that moment will be coded with a '1'.

Now you are going to have the opportunity to practice this recording method by watching a 2-min training video. Please refer to the green recording sheet with "Training Sheet" written at the top. While you watch the training video, please fill out the training sheet. When you are done, please check your answers against the Answer Key, which is on the back of your training sheet. Feel free to practice with the training video until you feel comfortable with the recording method.

When you are ready, please take out the blue recording sheet with "Test Sheet" written at the top. Please use this sheet to record environmental events and behavior while watching the 10-min test video. You may only view the test video one time.

When you are done with the test video, please turn the recording sheet over and answer the two multiple choice questions on the back of the recording sheet. They will ask you to identify the function of the target student's target behavior.

There will be a screenshot of the target students at the beginning of the training and test videos, so you will be able to quickly identify the appropriate students for data collection. Although there is a different target student for the training and test videos, for ease of recording both students will have the same target behavior (leaving the assigned area). There will also be a brief statement of the behavioral expectations for the students when the first minute begins, so you are able to accurately record the student's behavior when the video starts.

Finally, please remember the purpose of this project is to evaluate an A-B-C recording method. In practice, decisions should not be based upon one 10-min observation. Please do your best to identify the function based upon the information that you have been provided.

APPENDIX F

TRAINING SHEET

Name: Date:

		Environmental Events					Student Behaviors		
Setting Event	Time	**TD-G**	**ITA-N**	**ITA-O**	**PA**	**ESC**	**TB**	**OIB**	**AB**
W S I T	0:00-1:00								
W S I T	1:01-2:00								

Setting Events:

W = Whole group instruction; S = Small group instruction (≤ 3 students); I = Independent work; T = Transition

Environmental Events:

TD-G = Teacher directive to group; ITA-N = Individual teacher attention-Negative; ITA-O = Individual teacher attention- Other; PA = Peer attention; ESC = Escape from work

Target Student Behaviors:

TB = Target behavior; OIB = Other inappropriate behavior; AB = Appropriate behavior

TRAINING SHEET (BACK)

Name: Date:

		Environmental Events					Student Behaviors		
Setting Event	Time	**TD-G**	**ITA-N**	**ITA-O**	**PA**	**ESC**	**TB**	**OIB**	**AB**
(W) S I T	0:00-1:00	2 5 6 10	7 14		8 13	4 12	3	11	1 9 15
(W) (S) I (T)	1:01-2:00	2 3 4	8 9		7	6	5		1 10

Setting Events:

W = Whole group instruction; S = Small group instruction (≤ 3 students); I = Independent work; T = Transition

Environmental Events:

TD-G = Teacher directive to group; ITA-N = Individual teacher attention-Negative; ITA-O = Individual teacher attention- Other; PA = Peer attention; ESC = Escape from work

Target Student Behaviors:

TB = Target behavior; OIB = Other inappropriate behavior; AB = Appropriate behavior

APPENDIX G

TEACHER REFERRAL (INFORMED)

Timothy is a 10-year-old boy in 5th grade at Appletree Elementary school. He is in his fourth week of school, and has recently started displaying some disruptive behavior in Ms. Davis' math class. Ms. Davis wrote an email to the principal and the school's behavior specialist to request assistance with Timothy's behavior. The following is an excerpt from the email:

Timothy is all over the place and cannot seem to focus on the lesson. I am constantly redirecting him. If I had to guess, I would say Timothy is acting out to get attention from a couple of the other children in the class and to see if he can get a rise out of me.

APPENDIX H

TEACHER REFERRAL (NEUTRAL)

Timothy is a 10-year-old boy in 5th grade at Appletree Elementary school. He is in his fourth week of school, and has recently started displaying some disruptive behavior in Ms. Davis' math class. Ms. Davis wrote an email to the principal and the school's behavior specialist to request assistance with Timothy's behavior. The following is an excerpt from the email:

APPENDIX I

TEST SHEET

Name: Date:

		Environmental Events					Student Behaviors		
Setting Event	Time	**TD-G**	**ITA-N**	**ITA-O**	**PA**	**ESC**	**TB**	**OIB**	**AB**
W S I T	0:00-1:00								
W S I T	1:01-2:00								
W S I T	2:01-3:00								
W S I T	3:01-4:00								
W S I T	4:01-5:00								
W S I T	5:01-6:00								
W S I T	6:01-7:00								
W S I T	7:01-8:00								
W S I T	8:01-9:00								
W S I T	9:01-10:00								

Setting Events:

W = Whole group instruction; S = Small group instruction (≤ 3 students); I = Independent work; T = Transition

Environmental Events:

TD-G = Teacher directive to group; ITA-N = Individual teacher attention-Negative; ITA-O = Individual teacher attention- Other; PA = Peer attention; ESC = Escape from work

Target Student Behaviors:

TB = Target behavior; OIB = Other inappropriate behavior; AB = Appropriate behavior

TEST SHEET (BACK)

1. Based on your observation and recording, please select the function of the target child's problem behavior.
    a. Attention
    b. Escape

2. If you are not certain about your choice, please select the next most likely function of the target child's problem behavior.
    a. Attention
    b. Escape
    c. No second function

APPENDIX J

FOLLOW-UP QUESTIONNAIRE

Please answer the following questions based on how true the following statements are about your thoughts and/or behavior, where ‘1’ does not describe your thoughts or behavior, and ‘5’ completely describes your thoughts or behavior.

1. The A-B-C recording method was easy to use.

   1 2 3 4 5

2. I believe this A-B-C recording method is a helpful method for determining the function of problem behavior.

   1 2 3 4 5

3. I am likely to use this method of A-B-C recording in the future.

   1 2 3 4 5

4. I am confident in my selection of the behavioral function.

   1 2 3 4 5

APPENDIX K

TEST VIDEO SCRIPT

TIME	SPEAKER	ACTION
Min 1		
		Students are working independently Timothy standing at the door, looking outside Timothy is not working All other students are working quietly at their desks
	Teacher	**Alright class, it's time to start wrapping up writing**
		Target student remains standing at the door, looking outside
	Teacher	**Ok, writing is over so we're going to get started on math. I'd like everyone to pull out their math journals**
		Students begin to transition from writing to math
		Students put their writing away and start pulling their math journals out of their desks. Target student remains standing at the door, looking out the window
		Whole group math instruction begins
	Teacher	**Today we're going to be learning about coordinate planes. This is one of our standards that we need to cover by the end of the semester so I need everyone's focus.** Other students are watching the teacher as she starts the lesson.
	Teacher	**Timothy. You need to come sit down.**
		Timothy turns and starts walking back to his chair slowly. The other children continue to keep their eyes forward on the teacher.
		Timothy sits down
	Teacher	**Thank you, Timothy.** **Alright. I'm going to finish drawing our coordinate plane.**

		All students are in their seats watching the teacher draw on the board.
Min 2		Whole group instruction is taking place
		Target student is in his seat looking at the teacher
		All other students are in their seats looking at the teacher
	Teacher	**Allrighty boys and girls, I've drawn a new figure on the board. Can anyone look at it and tell me something familiar?**
		Heather and Brooklyn raise their hands.
	Teacher	**Heather?**
	Heather	**I see numbers going up the side and it's called a number thingy.**
	Teacher	**Oh I'm glad you saw the numbers going up. It's called a number line.**
		When Heather starts answering the question, Timothy leans over to Bob and starts saying her name repeatedly.
	Timothy	**Bob, Bob, Bob, Bob…**
		Timothy is not attending to instruction or the teacher
		All other students are taking notes and paying attention to the teacher and Heather.
	Bob	Bob whispers **"Stop!"**
	Teacher	**Bob and Timothy!**
		Bob and Timothy stop making noise and turn and face the teacher.
	Teacher	**So we have two number lines. One going horizontal and one going vertical. Together they make what's called a coordinate plane. I'm going to write it out for you.**
		Students are watching the teacher and taking notes. Timothy is watching the teacher.
	Teacher	**Alright. I want everyone to look at the board and write coordinate plane, along with its definition, into your journal.**
		Timothy gets out his math journal.

		All students start copying the information down into their journals except for Timothy. He is sitting in his seat looking at the teacher.
		Timothy gets out of his chair holiding his pencil, and walks to the back of the classroom. He starts sharpening his pencil in the pencil sharpener. The other students continue working. Timothy is not attending to the teacher or the instructional materials
Min 3		
		Whole group instruction is taking place Timothy is sharpening his pencil in the back of the classroom. Timothy is not attending to the teacher or the instructional materials The other students are quietly working at their seats.
	Teacher	**Timothy!**
		Timothy turns to look at teacher and holds up his pencil.
	Timothy	**I broke my pencil.**
	Teacher	**Shh. Just take this one.** Teacher grabs another pencil and puts it on Timothy's desk.
		Timothy walks back to his seat. Timothy sits down. All other students are writing in their notebooks.
	Teacher	**All right class, the 2 number lines have names. They're actually called axes. This axis is called the X axis. This is how you spell it. This axis here is called the y-axis. You spell it like this.** Timothy is sitting quietly at his desk. Students are writing in their notebooks and looking at the teacher.
	Teacher	**All right. Please write these names down in your journals.** Timothy puts his head down on his desk. Timothy is not attending to the teacher or the instructional materials Other students are writing and looking at the front of the classroom.
	Teacher	**Now, does anyone notice anything about where the two lines meet?** Brooklyn, Larry, and Abby raise their hands. Timothy keeps his head down.

Min 4		
		Whole group instruction is taking place
		Timothy still has his head down on his desk.
		Timothy is not attending to the teacher or the instructional materials
		Brooklyn, Larry, and Abby are raising their hands.
		Timothy starts playing with a rubber band that is on his wrist.
	Teacher	**Timothy?**
		Abby and Heather turn to look at Timothy.
		Timothy stops playing with the rubber band and sits up
	Timothy	**I don't know**
		Other students turn back to teacher and resume raising their hands.
	Teacher	**Brooklyn.**
	Brooklyn	**They both meet at the bottom corner.**
	Teacher	**Yes, that's right, but there's more to it.**
	Brooklyn	**They both meet at zero.**
	Teacher	**Yes. That's exactly right. That place has a name. It's called the origin. Here's how you spell it.**
		Students start writing in their notebooks.
		Timothy continues to stare quietly at the teacher, but does not write in his notebook.
	Teacher	**Origin. Here we have the numbers 2 and 3. This is called an ordered pair and it's used to find points on the coordinate plane. Notice that there are 2 numbers, separated by a comma. The numbers have parentheses around them. The first number has a name. It's called the X coordinate. It tells us how far down to travel on the x axis. One, two.**
		Timothy starts tapping his pencil on Bob's desk.
		Timothy is not attending to the teacher or the instructional materials
		The teacher is teaching and the rest of the students are paying attention to the teacher
Min 5		
		Whole group instruction is taking place
		Timothy is still tapping his pencil on Bob's desk.
		Timothy is not attending to the teacher or the instructional materials

	The other students are writing in their notebooks or looking at the teacher.
Teacher	**The second number is called the y coordinate. This number right here. The y coordinate tells us how far to travel up the y axis. One, two, three. Right there. All right we're going to try some new numbers now. Four and five. I've written a new ordered pair. Abby, what numbers do you see in the ordered pair?**
Abby	**I see a four and a five.**
	When Abby starts to talk, Timothy reaches over and starts drawing on Bob's paper
Bob	**Stop!**
Teacher	**Timothy. You need to pick up your journal and go to the back of the class.** Timothy rolls his eyes, picks up his journal, walks to the back table Timothy sits down at a desk. All other students continue looking at the teacher or writing in their journals.
Teacher	**Sorry, Abby. What numbers do you see?**
Abby	**I see a four and a five. The four is the x coordinate.**
Teacher	**Very good. You even answered my next question. What does the X coordinate do?**
Abby	**It tells you how far to go on the x axis.**
Teacher	**That's right. It tells you how far to go on this axis. One, two, three, four.**
	Timothy starts sliding his chair back to the computer. Then he turns and starts playing on the computer. Timothy is not attending to the teacher or the instructional materials The teacher continues teaching. The other students continue looking at the teacher.
Min 6	
	Whole group instruction is taking place Timothy is at the back of the classroom on the computer. Timothy is not attending to the teacher or the instructional materials The other students are all looking at the teacher quietly.
Teacher	**And what is the next number in the ordered pair?**

	Larry?
Larry	**It's a five.**
Teacher	**You're right. The five is the y coordinate. What does the y coordinate tell us to do?**
Larry	**It tells us to travel up.**
Teacher	**That's right. It tells us to travel up. Class, what does the y coordinate tell us to do?**
Class	**Travel up** (The whole class responds except Timothy, who is still playing on the computer, and Bob, who turns to look at Timothy).
	Bob starts whispering to Timothy
Teacher	**Nice work. I like how everybody is responding**
	Timothy turns around to answer Bob.
Teacher	**Hey! Timothy. Bob. You guys need to be quiet.** **Timothy, I need you to go back to your table please.** Bob turns back to face the front. Timothy gets up and brings his chair back to his desk. Timothy sits down
Teacher	**So as Larry said, the y coordinate tells us to travel up the y axis. Now I'm going to write some new numbers on the board.** **Brooklyn, what is the x coordinate?**
Brooklyn	**Four.**
	While Brooklyn is responding, Timothy reaches into the desk and pulls out a sheet of paper. Timothy is not attending to the teacher or the instructional materials The other students are paying attention to the instruction
Teacher	**Good job. The x coordinate is four. Heather, what's the y coordinate?**
Heather	**Six**
	While Heather is responding, Timothy starts making a paper airplane at his desk.
Teacher	**Great job girls. Four and six. Next we're going to get some worksheets ready for your group activity.**

		Timothy is making a paper airplane. The other students are quietly looking at the teacher.
Min 7		
		Whole group instruction is taking place Timothy is still making a paper airplane at his desk. Timothy is not attending to the teacher or the instructional materials The other students are sitting quietly at their desks either looking at the teacher or writing in their journals.
	Teacher	The teacher ends whole group instruction and begins to introduce new activity **All right. For group work today we'll be working with coordinate planes. On the coordinate plane you will label the x axis, y axis, and the origin. You will have five ordered pairs to plot on the coordinate plane.**
	Teacher	**Bob, you'll be with Timothy. That's for you guys.** Teacher places a worksheet down in front of Bob. Timothy is still sitting at the desk at the back table, holding his paper airplane.
	Teacher	**Heather and Abby, you guys will also be a group** As the teacher is passing out the next worksheet, Timothy throws his paper airplane at Abby. She turns and looks down at the airplane.
	Teacher	**Timothy! You need to get control of yourself man.** Abby and Heather look at Timothy. Timothy looks down at the ground. The teacher picks up the paper airplane and throws it in the trash. Timothy sitting at his desk
	Teacher	**Alright, back to your worksheets. You will use the ordered pairs traveling along the axes to make points on your coordinate plane. If you have any questions you can ask them to the other people in your group, or raise your hand and I'll come answer them for you.** While the teacher is talking, Timothy puts his head down on his desk. Timothy is not attending to the teacher or the instructional materials
		Small group instruction begins

		The other students start to work on the worksheets with their partners. They are talking quietly with each other. Bob is waiting for Timothy so they can get started on the worksheet.
		The teacher walks over to Timothy who still has his head down
	Teacher	**Timothy, I want you to come over and work with Bob, please.**
		Bob turns to look at Timothy. Timothy gets up, walks to his desk. Timothy sits down at his seat Bob is looking at the worksheet. After several seconds, Timothy stands up and starts walking away. Timothy is not attending to the teacher or the instructional materials Other students are working on their worksheets
Min 8		
		Small group instruction is taking place Timothy is standing at the back door looking out the window. Timothy is not attending to the teacher or the instructional materials The other students are working on their worksheets. The teacher walks over to Abby and Heather.
	Teacher	**How's your work going?**
		She continues to help Abby and Heather with a problem, then notices Timothy standing at the back of the classroom.
	Teacher	**Timothy! It's time to sit down, man.**
		Timothy starts walking back to his seat.
		Teacher moves around the room looking at the other students' work. Timothy sits down at his seat.
	Teacher	**I'll be at my desk, but if you guys have any questions you can raise your hand.**
		Students are working on their worksheets except for Timothy, who is sitting at his desk. Timothy gets out of his seat, walks over to get paper airplane, and brings it back to his desk. Timothy is not attending to his assigned work Other students continue working while he does this.

		Timothy is playing with his paper airplane at his seat. Timothy is not attending to his assigned work The other students are working.
Min 9		
		Small group instruction is taking place Timothy is still at his seat quietly playing with his paper airplane. Timothy is not attending to his assigned work Other students are working Timothy starts pretending to fly his paper airplane around, making swooshing noises. Abby turns to Timothy
	Abby	**Oooh, you're going to get in trouble.**
		Timothy throws his paper airplane. Timothy gets up and starts walking around the room. Timothy is not attending to his assigned work Other students continue working, not paying him any attention. Timothy continues walking around the room, looking at the teacher's things. Timothy hears the teacher coming, quickly picks up his paper airplane, runs back over to his seat. Timothy sits down, holding the paper airplane. Timothy is not attending to his assigned work
	Teacher	**Alright class. I need everyone to start wrapping up.**
Min 10		
		Small group instruction is taking place Timothy is doodling on his paper airplane which is hidden under his work. Timothy is not attending to his assigned work Students are finishing up their worksheets.
		Teacher starts walking around, stops to talk to Larry and Brooklyn and picks up their worksheet. Teacher sees Timothy doodling and not working on his worksheet.
	Teacher	**Timothy, your worksheet is not done. You need to finish your worksheet and you're going to have to miss recess if you're not done with it.**

	Peers look at Timothy (Larry and Bob)
	Teacher takes away his paper airplane and walks over to Abby and Heather.
Teacher	**Did you girls finish your worksheet?**
	Timothy grunts. While the teacher is talking to Abby and Heather he complains out loud.
Timothy	**It's not fair! Why are you always picking on me.**
	Teacher ignores him.
	Timothy gets up out of his chair and sits down on the floor
	Timothy is not attending to his assigned work
	Peers watch him
Teacher	**Alright Timothy, that's it, you're going to the office.**
	Timothy stands up and walks out of the room.

APPENDIX L

DEBRIEFING FORM

Sometimes in research it is necessary not to tell the participants the hypothesis. We cannot always tell people about the purpose of the experiment because it might affect the results–If we tell people the purpose of the experiment or how we predict people will act in the experiment, they may deliberately do whatever it is they think we want them to do, just to help us out and give us the results that they think we want.

As mentioned in the informed consent form, the primary purpose of this study is to determine if school psychologists and school psychologists-in-training are able to accurately collect data on environmental events and student behavior. There is another purpose of the study that was concealed from you prior to your collecting data.

I am also interested in whether having a hypothesis about the function of the student's problem behavior before conducting a direct observation biases both (a) the recording of behavior and (b) the selection of the function of the student's behavior.

For participants assigned to Condition A: You were provided with information from the target student's hypothetical teacher that suggested the student's behavior was maintained by attention.

For **participants assigned to** Condition B: You were not provided with a hypothesis on the function of the student's behavior before collecting data.

The information provided to those participants in Condition A was inaccurate. The script, performed by child and graduate student actors, was written to indicate escape as the function of the student's problem behavior. Often when a behavior referral is made in schools an opinion on behavioral function is provided; however, there is rarely data to support that hypothesis. With this research we are aiming to uncover whether or not participants are still able to (a) collect data accurately and (b) identify the correct function of behavior after receiving an inaccurate hypothesis of the function of behavior. This is important as school psychologists are likely to receive inaccurate information in practice.

Obviously, if we tell people outright what we are studying, it might affect their behavior. Thus, we had to conceal this particular research question until now.

There are many other school psychologists and school psychologists-in-training participating in this research both during this semester and across next semester. The success of this study requires that the people who participate are not aware of this secondary research question:

specifically, that we are interested in determining whether or not a hypothesis of the function of behavior biases participants' accuracy of recording and selection of behavioral function.

If you talk to others about this purpose of the study, it would jeopardize the validity of our findings. Participant responses wouldn't be spontaneous and natural. So if you discuss this study with others, we wouldn't have enough valid data to draw any conclusions about how people naturally behave in this situation. In short, the study would be wasted; your time would be wasted and our time would be wasted. After you leave, please do not discuss the details of this study. If anybody asks you about the study, you could tell them it was a study examining the accuracy of an A-B-C recording procedure.

I hope you see why it is important not to tell anyone the secondary purpose of the experiment. If you have any additional questions or comments, please feel free to contact the co-investigator at any time.

I have distributed this debriefing form to the participant following data collection and allowed her/him to ask questions.

_____ _____

Researcher Date

I have received and read the debriefing form, and I had the opportunity to ask questions about the true purpose and experimental manipulations that took place in this study.

_____ _____

Participant Date