AN ANALYSIS OF INTERRATER AGREEMENT BETWEEN THE MOTIVATION ASSESSMENT SCALE (MAS), QUESTIONS ABOUT BEHAVIORAL FUNCTION (QABF) AND ANALOG ASSESSMENT OUTCOMES

Carla Michelle Smith, B.S.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

May 2010

APPROVED:

Richard G. Smith, Major Professor and Chair,
Department of Behavior Analysis
Shahla Ala'i-Rosales, Committee Member
Manish Vaidya, Committee Member
Thomas L. Evenson, Dean of the College of Public
Affairs and Community Service
Michael Monticino, Dean of the Robert B.
Toulouse School of Graduate Studies

Smith, Carla Michelle. An analysis of interrater agreement between the Motivation Assessment Scale (MAS), Questions About Behavioral Function (QABF), and analog assessment outcomes. Master of Science (Behavior Analysis), May 2010, 45 pp., 4 tables, 8 illustrations, references, 40 titles.

An analysis of interrater agreement across multiple respondents on anecdotal assessments was compared with experimental functional analysis outcomes for correspondence. Experiment 1 evaluated the agreement of multiple respondents on the function of problem behavior for 22 individuals across 42 target behaviors using the Motivation Assessment Scale (MAS) and the Questions About Behavioral Function (QABF). Results showed agreement on the primary maintaining consequence for 4 or 5 of the 5 respondents in 52% (22/42) of the individual's target behaviors with the MAS and 57% (24/42) with the QABF. Experiment 2 examined whether correspondence occurred between the anecdotal assessment results and experimental functional analysis (EFA) results for 7 individuals selected from Experiment 1. Correspondence between the QABF assessment and the EFA was found for 6 of 7 participants, and 4 of the 7 showed correspondence between the EFA and the MAS. This study showed that the QABF had higher correspondence with analog assessments than the MAS thus, supporting the previous findings.

ACKNOWLEDGEMENTS

I would like to thank my mentor and friend Dr. Richard Smith for helping to shape my behavior analytic repertoire over the last 8 years—you inspire me. I would also like to thank my committee members, Dr. Shahla Ala`i-Rosales and Dr. Manish Vaidya for their influence in both this manuscript and my education. I must recognize the students and staff at the Behavior Analysis Resource Center (BARC) who contributed to this research, specifically Amy Peterson-Pace, Joseph Dracobly, Claire Anderson, Danielle Russell, Katy Atcheson, Jessica Hobbs, Martha Staff, and Audrey Shivers. I would also like to thank Anney Fahrenholz for the opportunity to collect data on her thesis and for her support as I replicated her work.

I must thank my family and friends for their sustained support of my education and career: Carleton and Marilyn Steele, Lauren Smith, Scott Smith, Leah Smith, Eunice Smith, LaJuana Hale, Holly Clarkson, Sharon Conrad, and Sue Kaeufer. Finally, and most importantly, I thank my husband and best friend, Steve Smith, for his life-altering support to help me become a behavior analyst.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF ILLUSTRATIONS

Page

CHAPTER 1

INTRODUCTION

Functional assessment is a widely used approach to individualize behavior plans for treating maladaptive behavior. The relationship between aberrant behavior and the environment is analyzed to determine the operant function of the behavior, instead of attributing the cause of problem behavior to an underlying pathology. Carr (1977) proposed that several contingencies of reinforcement might function to maintain self-injurious behavior (SIB) and since that time the field of applied behavior analysis has generated a large body of research in functional assessment methodology. Functional assessment identifies environmental variables that potentially influence an individual's challenging behavior, including antecedents that function to evoke the behavior and consequences that reinforce, or maintain the behavior (Neef & Peterson, 2007). These data-driven methods of assessing problem behavior have been applied in many settings including schools, hospitals, clinics, residential facilities, and homes (Vollmer, Marcus, Ringdahl & Roane, 1995; Iwata, Pace, Kalsher, Cowdery, & Cataldo, 1990; Iwata, Pace, et al., 1994) and with many different populations of individuals including those with intellectual disabilities (Asmus, Vollmer, & Borrero, 2002). Three general assessment methods have emerged in the literature: descriptive assessment, where information is obtained by direct observation in the natural environment (e.g., Bijou, Peterson, & Ault, 1968); anecdotal assessments, such as structured interviews or checklists (e.g., Durand & Crimmins, 1988); and experimental functional analysis, where environmental events are systematically manipulated to test the effect on behavior (e.g., Iwata, Dorsey, Slifer, Bauman & Richman, 1982/1994).

Experimental functional analysis (EFA) involves exposing an individual to a series of structured antecedent and consequent conditions to determine if the challenging behavior is

maintained by positive reinforcement, negative reinforcement, or automatic reinforcement (Zarcone, Rodgers, Iwata, Rourke, & Dorsey, 1991). Experimental functional analysis is the most widely studied assessment technique and is considered the "gold standard" of assessment (Applegate, Matson & Cherry, 1999; Hanley, Iwata, & McCord, 2003). Procedures developed by Iwata and colleagues (Iwata, Dorsey, et al., 1982/1994) paved the way for hundreds of direct and systematic replications. Experimental functional analysis methodology has been replicated across settings, populations, and behavioral topographies (Hanley et al., 2003). The validity of experimental functional analysis has been demonstrated repeatedly through research showing that it results in identification of the maintaining reinforcer and serves as a basis for the development of effective function-based treatments (Iwata, Vollmer, & Zarcone, 1990; Asmus et al., 2002). An advantage of experimental functional analysis is the ability to demonstrate cause-effect relations between the environment and behavioral events. Descriptive and anecdotal assessments infer those relationships through observed correlations between environmental events and problem behavior (descriptive assessment) or caregiver reports (anecdotal assessment) but do not show a cause and effect relationship (Asmus et al., 2002). There are several disadvantages to experimental analysis, however, including the level of expertise required to administer it, the possibility that high intensity behavior episodes will occur during assessment, and the possibility that treatments developed in a clinical setting may not be effective in the natural environment (Sturmey, 1995).

Descriptive assessment refers to a set of objective methods for collecting, quantifying, and interpreting data from observations in the natural environment (Bijou et al., 1968). There are several formats for conducting descriptive assessments; two common formats include the scatterplot and the A-B-C assessment. The scatterplot procedure allows one to record the

occurrence of a target behavior using a grid that locates the behavior in time. The interventionist can then identify temporal patterns (to the extent that they exist) and correlations with environmental events that correspond to specific times of day or days of the week, (Iwata, Kahng, Wallace, & Lindberg, 2000; Touchette, MacDonald, & Langer, 1985). The A-B-C assessment involves continuous recording of antecedent events (A), behavior (B), and consequences (C) by observers in the natural environment (Bijou et al., 1968; Asmus et al., 2002). Descriptive assessment provides a record of environmental and behavioral interactions as they occur in natural environments and thus may reveal idiosyncratic forms of reinforcement and schedules of consequences operating in those contexts (Mace, 1994). However, descriptive assessments cannot demonstrate a cause and effect relationship because no relevant variables are directly manipulated (Neef & Peterson, 2007). Furthermore, the "noise," (chaotic context) and intermittent schedules of reinforcement that characterize many natural environments can mask functional relationships or result in correlations between variables that are not actually part of the functional maintaining contingency (e.g., Lerman & Iwata, 1993; Thompson & Iwata, 2007). It has been suggested that combining descriptive assessment in the natural environment (to generate hypotheses and identify potential idiosyncratic contingencies) and experimental functional analysis (to confirm or disconfirm descriptive assessment outcomes) may be an efficient sequence for identifying contingencies operating in the natural environment (Mace & Lalli, 1991; Asmus et al., 2002).

Finally, indirect or anecdotal assessment employs interviews, rating scales, checklists, or questionnaires to determine possible sources of reinforcement maintaining problem behavior. Anecdotal assessments are conducted by obtaining this information from a caregiver who is presumably familiar with the circumstances surrounding the individual's problem behavior.

Respondents may include teachers, parents, direct care staff, or, in some cases, the individual whose behavior is being assessed (Neef & Peterson, 2007). Anecdotal assessments have the advantage of being efficient, inexpensive, and easy to administer (Paclawskyj, Matson, Rush, Smalls, & Vollmer, 2001). Although widely used, anecdotal assessments have well-known limitations. They rely on respondents' memories and opinions instead of direct observation of the behavior in question (Neef & Peterson, 2007). Anecdotal assessments generate a hypothesis about the function of aberrant behavior, but do not directly test this hypothesis. Also, the reliability and validity of indirect assessment has frequently been questioned (Bihm, Kienlen, Ness, & Poindexter, 1991). It has been recommended that these assessments be used only to gather preliminary information prior to experimental functional analysis (Iwata et al., 1982/1994).

One of the most widely used and studied anecdotal assessments, the Motivation Assessment Scale (MAS) (Durand & Crimmins, 1988), provides an illustrative example of some of the issues surrounding indirect assessment. Durand and Crimmins, the developers of the instrument, used the MAS to assess the self-injurious behavior (SIB) of 50 developmentally disabled children. Teachers who worked with the children for the academic year served as primary respondents to the MAS and their outcomes were compared with those from teacher aides from the same classrooms. Pearson correlation coefficients were completed for rater pairs and showed significant correlations between raters for the individual questions (correlations were significant at the .001 level and ranged from .66 to .92). The validity of MAS outcomes was evaluated for 8 subjects by comparing analog assessments (using procedures developed by Carr & Durand, 1985) with the MAS results. Results indicate that the outcomes of analog and MAS assessments matched in all cases. Durand and Crimmins (1988) concluded that the MAS was "a

reliable scale that can predict how individuals will behave in analogue assessment settings" (p. 113).

Subsequent studies have produced mixed findings. Some studies have produced positive outcomes with the MAS (Bihm et al., 1991; Singh et al., 1993; Spreat & Connelly, 1996); however, other results have been less impressive. For example, Zarcone et al. (1991) were unable to replicate the outcomes of the original MAS study. Results of this study showed agreement in only 16 of 55 rater pairs in determining the source of reinforcement. The Pearson correlations for each individual ranged from -.30 to .81 (M = .27). Several additional studies have reported similar outcomes (e.g., Newton & Sturmey, 1991; Paclawskyj et al., 2001; Sigafoos, Kerr, & Roberts, 1994; Thompson & Emerson, 1995). Thus, although Durand and Crimmins (1988) reported encouraging outcomes for the MAS, several attempts to replicate their findings have met with failure.

The Questions About Behavioral Function (QABF) (Matson & Vollmer, 1995) is an anecdotal assessment used by schools, institutions, and service providers to help determine sources of reinforcement for problem behavior. The developers of this assessment have conducted a series of studies investigating its reliability and validity. For example, Paclawskyj, Matson, Rush, Small, & Vollmer (2000) examined test-retest reliability, interrater reliability, and stability of the QABF, producing encouraging results. A factor analysis by this group showed statistically significant differences among subscales (Paclawskyj et al., 2000). Another study showed the QABF was able to determine a clear behavioral function in 84% of 398 individuals. When results were assigned to SIB, aggression or stereotypy groups a clear function was shown in 83% of the SIB group, 74% of the aggression group, and 93% of the stereotypy group (Matson, Bamburg, Cherry, & Paclawskyj, 1999). Furthermore, Matson et al. (1999) evaluated

the treatment utility of the QABF in the natural environment by comparing target behavior rates for individuals whose treatments were based on QABF results and individuals whose treatment was based on a standard treatment (interrupt, block, and redirect) instead of functional assessment. Paclawskyj et al. (2001) performed an evaluation of convergent validity between the QABF or MAS and experimental functional analyses and found that the QABF had moderate agreement (56.3%) agreement with the experimental functional analyses and the MAS had 43.8% agreement with experimental functional analyses. The authors noted when undifferentiated experimental functional analyses for 3 participants were excluded, percentage agreement increased to 69.2% for the QABF and EFA and 53.8% for the MAS and EFA (Paclawskyj et al., 2001). By contrast with generally positive outcomes of these studies, a study of key psychometric properties of the MAS and the QABF by a different research team (Shogren & Rajah, 2003) showed less than satisfactory interrater reliability, with both scales falling into the fair to good range. Both assessments were found to be comparable in measuring similar constructs and in terms of reliability.

Based on concerns about reliability and validity, it may be tempting to dismiss anecdotal approaches to functional assessment. However, Fahrenholz (2004) investigated a variation of typical administration and interpretation procedures to improve the utility of anecdotal assessment in applied settings. This researcher assessed the extent of agreement among groups of raters and attempted to determine if high levels of group agreement were associated with correspondence between anecdotal assessments and experimental functional analysis results. Fahrenholz solicited responses to 2 anecdotal assessments (MAS and Functional Analysis Screening Tool [FAST] [Iwata & DeLeon, 1996]) from 5 respondents for each resident and evaluated agreement within and across assessments for 28 residents. Fifty-four percent of the

assessments showed agreement among 4 of 5 respondents with the MAS and 79% of assessments showed agreement among 4 of 5 respondents with the FAST. Based on these results a sample of 6 participants for whom at least 4 respondents showed agreement underwent experimental functional analyses. Four of 6 participants had experimental functional analysis that produced interpretable data (i.e. problem behavior was observed during the functional analysis) and those results corresponded with results from the anecdotal assessments. A limitation of this study was that the sample of residents who participated in experimental functional analysis only represented 2 of the 4 subscale categories (tangible and sensory reinforcement).

The current study attempted to extend the literature by investigating the potential utility of two anecdotal assessments – QABF and MAS – with multiple respondents. The extent of agreement among 5 respondents was evaluated for each assessment and correspondence with experimental functional analysis outcomes was evaluated for a sample of participants from each of 4 subscale categories (attention, tangible, escape, and sensory).

CHAPTER 2
EXPERIMENT 1: EVALUATION OF INTERRATER AGREEMENT WITHIN AND ACROSS ASSESSMENTS

Method

Participants and Setting

Experiment 1 was conducted at a large, state-sponsored residential facility for individuals with intellectual disabilities. Assessments were administered in secluded areas of the residential apartments or quiet areas in the vocational training setting.

Residents. Twenty-seven individuals who resided at the facility participated in Experiment 1. The residents' ages ranged from 27 to 66 years and all were diagnosed with mental retardation. Refer to Table 1 for each individual's age and functioning level.

Target behaviors. All participants had a history of problem behavior of sufficient severity to necessitate the development of behavior support plans by their unit psychologist and individual support team. The behavioral definitions and topographical descriptions used in Experiment 1 were developed by the individuals' unit psychologists and were part of each individual's behavior support plan. Twelve participants presented with a single target behavior and 15 presented with 2 target behaviors. Motivation Assessment Scale and Questions About Behavioral Function were completed for each target behavior; thus, for residents who presented with two target behaviors, 20 assessments were completed (5 MAS & 5 QABF for target behavior 1 and 5 MAS & 5 QABF for target behavior 2). Due to an error in administration of the assessments, data from only 4 respondents are reported for Barbara. Eleven residents presented with self-injurious behavior (SIB); 10 residents presented with physical aggression toward others (PAO); 7 residents presented with verbal disruptive behavior (VDB); 4 residents presented with physically disruptive behavior (PDB); 3 residents presented with pica; 3 residents presented with stealing (STE); and 4 residents presented with other target behaviors, including

stereotypic behavior (STEO); rumination (RUM); mouthing objects (MO); and aggression toward property (AGP). Refer to Table 1 for a description of residents' target behaviors and definitions.

Respondents. Respondents for Experiment 1 were 113 staff members who had worked regularly with residents as direct care, vocational, or unit management staff for a minimum of 6 months. Educational background for staff members was unavailable. Prior to the start of the study the facility discontinued a longstanding hiring policy requiring a high-school education or equivalent for employment. All respondents were employees of the facility at the time of the interviews.

Materials

Materials used in Experiment 1 included writing utensils and 2 sets of each questionnaire (MAS and QABF). The general information sections of each assessment (name, residence, date, rater, target behavior, and setting descriptions) were completed by the interviewer prior to the start of the assessment.

MAS. The Motivation Assessment Scale (MAS; Durand & Crimmins, 1988) was a 16-question assessment with 4 questions corresponding to each of four subscale categories: escape, sensory, attention, or tangible. Respondents answered questions using a 7-point Likert-type scale with scores indicating the extent to which the rater observed the behavior, from 0 (*never)* to 6 (*always)*. Intermediate values allowed respondents to score 1 (*almost never*), 2 (*seldom*), 3 (*half of the time*), 4 (*usually*), or 5 (*almost always*). The 4 questions corresponding to each category were summed and rank ordered by point value. The category with the highest value was assumed to represent the maintaining contingency.

QABF. The Questions about Behavioral Function (QABF; Matson & Vollmer, 1995) was designed to assess the functions of aberrant behavior. The QABF was a 25-question assessment with 5 questions assigned to each subscale. The scale allowed respondents to use a 4-point Likert style scale (0-3) to score how often the client demonstrated the target behavior. Respondents chose from 0 (*never*), 1 (*rarely*), 2 (*some*), and 3 (*often).* Completed assessments were scored for the number of items endorsed (i.e. if a question was answered) and for the point total assigned to each subscale category (attention, escape, non-social, physical, and tangible). Subscale categories were then rank-ordered according to the score and the category with the highest point value was assumed to represent the maintaining contingency.

Administration Procedures

Twenty-three graduate and undergraduate students were trained to administer the QABF and MAS. Training included reading and discussion of the MAS and QABF manuals, role-playing administration procedures with a senior trainer providing feedback, observing a senior trainer administer an assessment, and receiving feedback following initial assessment delivery. Assessments were administered in secluded areas of client apartments or vocational work settings. To assist staff who may have had difficulty reading the assessment, each respondent was given a copy of the questionnaire to "read along" as the interviewer read each question aloud and scored answers. Interviewers provided only the information on the assessment; no additional information or clarification was given. If staff asked questions, they were told "answer the best you can." The MAS was administered first, followed by the QABF. After completing both instruments the interviewer thanked the respondent and left the area. If the individual presented with two target behaviors, interviewers returned at another time to administer assessments for the second target behavior.

Respondent Agreement Evaluation

Assessments were scored and summarized in an office away from respondents. Inter-scorer agreement was calculated by having two trained graduate students score each assessment. Resulting scores were compared on a question-by-question basis. Agreement for both the MAS and QABF was 100%.

Agreement across assessments (within respondents). Agreement across assessments and within respondents was scored if the respondent identified the same maintaining contingency with both instruments. The MAS was organized according to 4 categories of maintaining variables including: sensory, escape, attention, and tangible. The QABF was organized according to 5 categories of maintaining variables including: non-social, escape, attention, tangible, and physical. Agreement was scored if a respondent ranked sensory as the maintaining variable on the MAS and either non-social or physical as the maintaining variable on the QABF. If a respondent scored two categories as the highest ranking (i.e. there was a tie) on one assessment, both categories were compared with the highest ranking category from the other assessment. For example, agreement was scored if a respondent scored both attention and escape as the highest ranking categories on the QABF (i.e., attention and escape received the same score, which was higher than scores for other categories) and scored attention as the highest ranking category on the MAS.

Agreement across assessments (across respondents). Agreement across assessments (across respondents) was scored if 4 or 5 of the 5 respondents for each resident identified the same maintaining variable for both the MAS and QABF.

Agreement within assessments (across respondents). Agreement within assessments (across respondents) was scored if 4 or 5 of the 5 respondents for each resident identified the same maintaining variable within either the MAS and QABF.

Results

Five respondents completed both the MAS and QABF for each target behavior (except for Barbara, as noted previously). Table 2 shows respondent agreement within and across assessments. For the MAS, 4 or 5 of the 5 respondents agreed on a primary maintaining consequence for 52% (22/42) of individual's target behaviors. For the QABF, 4 or 5 of the 5 respondents agreed on a primary maintaining consequence for 57% (24/42) of individual's target behaviors. Respondents were in agreement across both MAS and QABF for 26% (11/42) of individual's target behaviors. Perfect agreement (5 of 5) occurred with 12 (29%) of MAS respondents, 7 (17%) of QABF respondents and 3 (7%) across both the MAS and QABF.

Table 3 shows the number of respondents who identified specific categories of maintaining consequences across individuals and target behaviors. Results are shown for the QABF (left) and MAS (right). Within-assessment ties between primary maintaining consequences account for instances where the number of identified maintaining consequences was greater than 5 for a given target behavior. Shaded rows indicate target behaviors for which 4 or 5 of the 5 respondents agreed on the primary category of maintaining consequence. Shaded rows that extend across both assessments indicate across-respondent agreement for both the MAS and the QABF. These data offer a detailed view of the summary data presented in Table 2 and Figure 1.

Figure 1 shows percentages of primary categories of maintaining variables across respondents for the MAS and QABF (there were 210 responses for each assessment tool). MAS

respondents scored sensory as the primary maintaining variable 82 times (39%); a tangible contingency was identified 81 times (38.6%); an attention contingency was identified 40 times (19%); and an escape contingency was identified 31 times (14.8%). Sixteen respondents' MAS scores showed a two-way tie between categories and 4 showed a 3-way tie, resulting in a total number of 234 contingencies identified by the 210 respondents (111%).

Of the 210 responses to the QABF, 63 (30%) identified escape as the primary maintaining contingency; a non-social reinforcement contingency was identified 62 times (29.5%); a tangible contingency was identified 60 times (28.6%); an attention contingency was identified 38 times (18.1%); and a physical reinforcement contingency was identified 14 times (6.7%). QABF respondents scored 26 two-way ties and 1 three-way tie between primary maintaining variables resulting in a total number of 237 contingencies identified by the 210 respondents (113%).

Discussion

Experiment 1 outcomes showed agreement across 4 or 5 raters in 55% of all assessments administered. Agreement occurred for 22 of 42 target behaviors (52%) in MAS assessments and in 24 or 42 target behavior (57%) in QABF assessments. The outcomes for the MAS are consistent with the findings reported by Fahrenholz (2004), showing that 15 of 28 (54%) of MAS assessments showed agreement across at least 4 of 5 respondents.

The results of Experiment 1 did not reveal substantial differences in agreement between the QABF and the MAS. The QABF produced slightly greater overall agreement (57% versus 52%), but was also more likely to show ties between identified categories than the MAS. The QABF produced 27 ties, whereas the MAS produced 20 ties. These results suggest that the overall agreement percentages for the QABF may have been artificially inflated based on an

increased tendency for individual raters to identify more than one primary contingency. That is, the mathematical probability of agreement across raters is increased when individual raters identify 2 or more categories; therefore, the less a given instrument produces differentiation among responses (*not* a desirable quality for an assessment instrument), the more likely that correspondence across respondents will be observed. In effect, if a common response to the question "which of these contingencies is responsible for this person's problem behavior?" is "one of two (or three) contingencies" instead of "this *one particular contingency*", then more opportunities for agreement with other raters will be available. So, ironically, the instrument that produces less certainty will, logically, produce more agreement. It should be noted that, in some cases, within-rater ties may also occur as a function of multiple maintaining contingencies (i.e., problem behavior may, in fact, be maintained by multiple contingencies of reinforcement). Therefore, it is not possible to determine if higher levels of agreement within the QABF are due to a better ability to identify multiple controlling contingencies or a decreased ability to distinguish among possible maintaining contingencies.

There were some limitations to Experiment 1. First, target behaviors were identified and defined by residents' unit psychologists and support teams prior to the start of the study. Some target behavior definitions included multiple topographies, which may have increased the possibility of maintenance by different contingencies. For example, Jolinda's target behavior was SIB, which was defined as biting fingers/hands, head banging, and slapping self. It is possible that these very different topographies of behavior were maintained by different consequences.

Another limitation was turnover among staff (respondents), reported by the facility to be 64.57% in 2009 (S. Musgrave, personal communication, February 18, 2010). Thus, although

administration guide for the MAS suggests that respondents are acquainted with individuals being assessed for at least one year (the QABF administration manual suggests that informants should know individuals for at least 6 months), it was necessary to obtain responses from caregivers who had known individuals for as few as 6 months in the current study. All staff worked with the residents for a minimum 6 months, but some staff worked with individuals for greater than 10 years. It is plausible that differences in scoring and, thus, the ability to identify maintaining variable, were at least in part a function of the length of time staff worked with an individual. The rate at which a target behavior occurred may also have affected staff's ability to accurately identify variables with anecdotal assessments. For example, it may be easier to identify the environmental variable associated with behavior that occurs more frequently in natural settings. Future research should investigate how differences in settings, staff tenure, and rate of problem behavior affect anecdotal assessment outcomes.

Based on the results of Experiment 1, 8 residents were selected to participate in Experiment 2. Participants selected all showed agreement within the QABF assessment and 7 of the 8 also showed MAS agreement. Two individuals with agreement in maintaining contingencies were selected for each QABF subscale category (i.e., non-social, attention, tangible, and escape) with the exception of the physical subscale category. In Experiment 2 experimental functional analyses of each participant's problem behavior were conducted.

CHAPTER 3

EXPERIMENT 2: EVALUATION OF CORRESPONDENCE BETWEEN ANECDOTAL AND EXPERIMENTAL ANALYSIS OUTCOMES

Method

Participants, Setting, and Materials

All sessions were conducted in a behavior analysis clinic for the assessment and treatment of behavior disorders, which was located on the campus of the residential facility at which Experiment 1 was conducted. Experimental functional analysis (EFA) sessions were conducted in one of the clinic's observation rooms. Rooms were 3.7 m by 3.7 m and contained a table, two chairs, and materials appropriate for the experimental session. A one-way mirror was installed in one wall of each room for unobtrusive observation and recording of session data.

For pica assessments, materials included simulated non-food items that could be safely consumed by participants. Simulated pica items for Greg were all natural soap (made from edible oils and wax) placed in a soap dish; and mixtures of water, white vinegar, apple cider vinegar, Simply Thick Gel, and food coloring placed in bath gel bottles, hand sanitizer pump, and spray bottle to simulate bath products, cleaning chemicals, and hand sanitizer. Simulated pica items for Vern were rice paper to simulate paper, onion skins to simulate paper and leaves, dried seaweed to simulate leaves, and brown fettuccini to simulate leaves and twigs. These items were continuously available throughout all EFA sessions for both participants.

Participants. A sample of 8 individuals were selected from among Experiment 1 participants for whom 4 or 5 of the 5 respondents showed agreement within the Questions About Behavioral Function (QABF) (see Table 1 for demographic information). Participants were selected in order to obtain representation from all subscale categories showing agreement among at least one group of respondents (non-social, attention, tangible, and escape) and based on

availability and continuing need for behavioral intervention. Greg was a 46-year-old man diagnosed with profound mental retardation and Vern was a 41-year-old man diagnosed with profound mental retardation. Both men presented with pica and their QABF assessments indicated that pica was maintained by non-social reinforcement. Jolinda was a 55-year-old woman diagnosed with profound mental retardation who presented with self-injurious behavior (SIB). Annie was a 33-year-old woman diagnosed with profound mental retardation who presented with physical aggression to others (PAO). QABF results indicated that target behaviors for Jolinda and Annie were maintained by positive reinforcement in the form of tangible items. Karen was a 47-year-old woman diagnosed with profound mental retardation who presented with verbal disruptive behavior (VDB). Asa was a 51-year-old man diagnosed with profound mental retardation who presented with SIB. QABF assessment results for Karen and Asa indicated that their problem behaviors were maintained by negative reinforcement in the form of escape from task demands. Chad was a 27-year-old man diagnosed with profound mental retardation and J.J. was a 29-year-old man diagnosed with mild mental retardation. Both men presented with SIB that, according to QABF results, was maintained by attention.

Target behaviors. The operational definitions used for the EFA were based on residents' target behavior definitions found in their formal behavior support plans. Definitions were refined for the EFA when necessary based on direct observations of problem behavior and a review of the client's records, including the daily reports of problem behavior across all settings. Greg and Vern's target behavior was pica, defined as the insertion of simulated pica items into the mouth. Jolinda presented with SIB, defined as biting her fingers/hands, head banging, or slapping herself. Annie's target behavior was PAO, defined as scratching, hitting, pinching, biting, pushing, grabbing, or 5-finger contact to another person. Karen presented with VDB,

defined as yelling, screaming, or crying. Asa's target behavior was SIB, defined as biting his hand/finger. Chad's target behavior was SIB, defined as his head striking any object (including people). J.J. presented with SIB, defined as hitting himself with his hand on any part of his body (such as head, face, chin, forehead, leg, etc.) or hitting his head or hand against an object.

Experimental Functional Analysis (EFA)

Observation procedures. Target behaviors were recorded by trained observers using handheld computers. Observers used frequency measures for: pica (Greg, Vern); head banging (Chad, J.J.); hitting body parts against objects (J.J.); hitting others (Annie); biting (Annie); pinching (Annie); hitting self (J.J.); and scratching, pushing, or grabbing others (Annie). Duration measures were used for: yelling, screaming, and crying (Karen), head banging (Jolinda), slapping self (Jolinda), and biting hands/finger (Asa, Jolinda).

Interobserver agreement (IOA). A second trained observer collected data during a percentage of each participant's EFA sessions. IOA was calculated by dividing each session into 1-s intervals, summing the number of intervals in which the primary and secondary observers agreed on the occurrence or non- occurrence of the target behavior, dividing the result by the total number of intervals in the session and multiplying the outcome by 100. A second observer independently and simultaneously scored 85% of Greg's sessions, 65% of Vern's sessions, 80% of Jolinda's sessions, 60% of Annie's sessions, 64% of Karen's sessions, 63% of Asa's sessions, and 38% of J.J.'s sessions. IOA was calculated slightly different for J.J. due to difficulty in determining the exact second of the onset of head up. IOA for J.J.'s sessions were calculated as above but with a moving 2 s window (e.g., if the primary observer recorded an event at time *x*, agreement was scored if the secondary observer recorded the same event at time *x*-1 s, time *x*, or time *x*+1 s), dividing the result by the total number of intervals in the session, and multiplying

the outcome by 100. The mean agreement for J.J.'s precursor experimental functional analysis was 99% (range 85%-100%). The mean agreement for Greg's EFA IOA was 99% (range 99%-100%) Vern's EFA IOA agreement was 99% (range 97%-100%); Jolinda's IOA agreement was 99% (range 96%-100%); Annie's mean IOA agreement was 96%; Karen's mean IOA agreement was 98% (range 89%-100%); and Asa's mean IOA agreement was 98% (range 91%-100%).

General procedures. Experimental functional analysis procedures similar to those described in Iwata et al., (1982/1994) were implemented in Study 2. Because of the low frequency and intensity of his SIB, an EFA of precursor behavior was conducted for J.J., as described by Smith and Churchill (2002). Experimental contingencies were in effect for J.J.'s precursor behavior (head up) during the analysis and no consequences were presented contingent on SIB. The operant function of J.J.'s SIB was inferred from outcomes of the precursor assessment. All eight participants were exposed to 3 test conditions (Alone/no interaction, attention, and demand) and a control condition (play), presented within a multielement format. Six of the participants (excluding Greg and Vern) were also exposed to a tangible condition. Each session lasted for 10 min. One to 6 sessions were run per day in the following order: alone/no interaction, attention, play, tangible (if relevant), and demand. Sessions were run at the same time each day for 3 to 5 days a week and the number of sessions conducted each day was arranged so as to start with a different session on successive days (i.e., no day ended with a complete cycle through conditions) so that sequencing patterns were unlikely to develop. Graduate students who were trained in facility protocols for management of aggressive behavior, protection of human subjects, and cardiopulmonary resuscitation served as therapists. The following sections describe each experimental condition.

Alone (Greg, Vern, Jolinda, Chad) / no interaction (Annie, Karen, Asa, J.J.). The participant was placed in the observation room alone. For no interaction sessions a therapist was seated by the door but did not interact with the individual. No materials were present in the room with the exception of simulated pica items for Greg and Vern (see *Participants, Setting, and Materials*). There were no consequences for target behaviors during this condition. The purpose of this condition was to evaluate whether the target behavior was maintained by non-social reinforcement.

Attention. During this condition the individual was in the observation room with the therapist. Recreational and leisure materials were present. Vern, Greg, Annie, Karen, Asa, Chad, and J.J. had magazines for leisure items and Vern had additional recreational toys. Leisure materials for Jolinda were beads, playing cards, and 2 small toys. The therapist delivered attention in the form of statements of concern, touch, eye contact, or reprimands contingent on each occurrence of the target behavior. The purpose of this condition was to evaluate whether the target behavior was maintained by social positive reinforcement in the form of attention.

Play. During this condition the participant was in the observation room with the therapist and recreational and leisure materials were present. The therapist interacted with the participant at least every 30 s. Interactions included prompting the individual to interact with the leisure items, discussing preferred topics, singing, or responding to requests. Demands were not made of the individual and consequences were not presented contingent on target behaviors. The purpose of this condition was to serve as a control against which to compare response measures from other conditions.

Tangible. During this condition the individual was in the observation room with the therapist. Leisure and recreational materials were not present; however, the targeted tangible item was present but unavailable to the individual. The individual was given 5 s access to the tangible item (or a small amount of the edible item) prior to the start of the session. The item was kept in plain view of the individual, but access was contingent on the occurrence of the target behavior, after which 5 s of access or another bite was provided. The purpose of this condition was to evaluate whether the target behavior was maintained by social positive reinforcement in the form of access to tangible or food items. Participant preferences were assessed prior to the EFA using a multiple stimulus without replacement preference assessment (MSWO) (DeLeon & Iwata, 1996). The following items were identified as highly preferred for the 6 individuals participating in this condition: magazines (Jolinda), vanilla wafer cookies (Annie), diet soda (Karen), chocolate chip cream cookie (Asa), and BBQ potato chips (Chad). Pizza flavored chips ranked first in the MSWO assessment for J.J. followed by chewy chocolate chip cookie and pudding ranked second. A 45-trial paired choice assessment (Fisher, Piazza, Bowman, Hagopian, Owens & Slevin, 1992) was conducted with J.J. because he did not always engage with items during his preference assessment but repeatedly made the manual sign for cookie. Chewy chocolate chip cookie was consistently chosen and was used for the tangible sessions of J.J.'s precursor FA.

Demand. During this condition the individual was in the observation room with the therapist. The therapist delivered a demand every 30 s using a 3-prompt sequence. The first prompt was a verbal prompt (e.g., "fold the towel"). If the individual did not comply within 5 s a modeling prompt was delivered (e.g., the therapist folded the towel and said, "fold the towel like this"). If the individual complied after the first (verbal) or second (modeling) prompt, the

therapist delivered verbal praise. If a third prompt was necessary, the therapist provided physical guidance to complete the task (e.g., the therapist said, "fold the towel like this" and physically guided the individual to fold the towel). If at any point in the sequence the individual engaged in the target behavior the therapist turned away from the client and terminated the task demand (e.g., the therapist turned away and said, "Never mind, you don't have to"). The purpose of this condition was to evaluate whether the target behavior was maintained by social negative reinforcement in the form of escape or avoidance of demands.

Results and Discussion

Results of each participant's experimental functional analysis are presented in Figures 2 through 8. Table 4 presents a comparison of anecdotal assessment results with EFA results for Experiment 2 participants.

Greg. Greg's results are shown in Figure 2. Pica occurred exclusively in the alone condition and remained at 0 in the attention, play, and demand conditions, suggesting a non-social function for Greg's pica. Anecdotal assessment results were consistent with the EFA. Four of 5 respondents to the QABF identified a non-social function and 5 of 5 respondents to the MAS identified a sensory function, demonstrating perfect correspondence with the results of the analog assessment.

Vern. Vern's EFA results are displayed in Figure 3. Pica occurred across all four conditions, but a differential pattern can be seen from sessions 12 through 20, with highest levels of pica in the alone condition (M = 8.0 responses per session), followed by attention (M = 5.1 responses per session), play (M = 3.1 responses per session), and demand (M = 1.0 response per session) conditions. Comparison of the EFA and the anecdotal assessments results show correspondence among the outcomes. Both MAS and QABF results show 5 of 5 respondents

agreed on a non-sensory or automatic reinforcement function of Vern's pica; thus, anecdotal assessment results corresponded perfectly with results from the experimental analysis.

Jolinda. Jolinda's EFA results are displayed in Figure 4. Jolinda engaged in more SIB during the alone condition (M = 3.4 responses per session) than during other assessment conditions; however, she also engaged in SIB during other test conditions (tangible M = .72 responses per session; demand M = .36 responses per session; attention M = .18 responses per session; control M = 0 responses per session). No SIB occurred during the final 3 cycles of assessment. Based on these inconsistent and largely undifferentiated outcomes, the EFA does not provide substantial support for any account of Jolinda's SIB. These outcomes show no correspondence with those from the anecdotal assessments, in which 4 of 5 respondents identified social positive reinforcement in the form of tangible items as the likely maintaining contingency for Jolinda's SIB for both assessments.

Annie. Annie's EFA results are shown in Figure 5. Following one cycle during which no responding was observed in any condition, PAO occurred at consistently high frequencies in tangible sessions (M = 32.5 responses per session) and at lower frequencies in the no interaction (M = 5.5 responses per session), attention (M = 4.0 responses per session), play (M = 2.2 responses per session), and demand (M = 2.0 responses per session) conditions. These outcomes strongly indicate that Annie's PAO was maintained by social positive reinforcement in the form of access to tangible items. Outcomes of both the MAS (5 of 5 respondents) and QABF (4 of 5 respondents) correspond to the EFA results. A potential limitation of Annie's assessments is that the anecdotal assessments were administered following the end of Annie's EFA. In all other cases, MAS and QABFs were administered prior to the EFA. However, it is unlikely that staff responses to the MAS and QABF were affected by the results of the EFA, as neither the

respondents nor the assistants who administered the assessments had knowledge of the EFA or its outcomes at the time of administration.

Karen. Karen's EFA results are shown in Figure 6. VDB occurred exclusively in the demand condition (*M* = 115 s per session), strongly indicating that Karen's problem behavior was maintained by social negative reinforcement in the form of escape from task demands. Anecdotal assessment results from both the MAS (4 of 5 respondents) and QABF (4 of 5 respondents) suggested an escape function for Karen's VDB, showing perfect correspondence among the MAS, QABF, and EFA.

Asa. Asa's EFA results are shown in Figure 7. Asa's SIB consistently occurred during a high percentage of intervals in the escape condition (*M* = 18.7% of intervals) relative to other test conditions. Lower levels of problem behavior were observed in play (*M* = 7.2% of intervals), alone (*M* = 5.5% of intervals), attention (*M* = 2.8% of intervals), and tangible (*M* = 1.2% of intervals) conditions. These outcomes suggest that his SIB was maintained by social negative reinforcement in the form of escape from task demands. QABF results showed that 5 of 5 raters identified escape as the primary maintaining contingency; however, the MAS did not produce agreement among 4 of 5 raters for Asa's SIB. Therefore, only the results of Asa's QABF corresponded with EFA results.

Chad. Chad sustained a fall, unrelated to his participation in the study, after participating in only 4 sessions of his EFA. Based on injuries related to the fall and concerns about the severity of his SIB, he was not eligible for further participation in the study at the time of this report. It was not possible to evaluate correspondence between EFA and anecdotal assessment for Chad.

J.J. J.J.'s precursor EFA results are presented in Figure 8. J.J.'s precursor to SIB consistently occurred at higher rates in the attention condition (M = 1.7 responses per min) than in other test condition, suggesting that his SIB was maintained by social positive reinforcement in the form of caregiver attention. Lower levels of problem behavior were observed in play (M = 1.1 responses per min), tangible (M = .69 responses per min), no interaction (M = .13 responses per min), and demand (M = .06 responses per min) conditions. SIB occurred only once during the precursor EFA, during the first presentation of the attention condition. QABF results showed that 4 of 5 raters identified attention as the primary maintaining contingency; however, the MAS did not produce agreement among 4 of 5 raters for J.J.'s SIB. Therefore, only the results of J.J.'s QABF corresponded with EFA results.

CHAPTER 4

GENERAL DISCUSSION

Effective and efficient treatment of behavior disorders depends, in part, on the identification of maintaining variables for problem behavior. Experimental functional analysis (EFA), descriptive assessment, and anecdotal assessment are common methods used to assess the function of problem behavior. Each assessment method has associated strengths and weaknesses, and all require varying degrees of time and resources. Many consumers of behavioral services, such as schools, families, clinics, or institutions, have limited time and resources to assess the function of problem behavior. Use of anecdotal assessments with multiple respondents can help interventionists identify the function of problem behavior quickly and easily when agreement among respondents is observed.

The purpose of the current study was twofold. Experiment 1 evaluated within and across assessment agreement of the Motivation Assessment Scale (MAS) and Questions About Behavioral Function (QABF) among 4 or 5 respondents. Respondents agreed on the function of problem behavior in 57% of QABF assessments and 52% of MAS assessment. This replicates Fahrenholz's (2004) findings regarding levels of agreement among respondents using these procedures. Experiment 2 investigated the extent of EFA correspondence with anecdotal assessments. Correspondence between the QABF and EFA was found for 6 of 7 participants (85.7%). For the sole case in which correspondence was not observed, the outcomes of EFA were not sufficiently differentiated to determine correspondence. Correspondence between the MAS and EFA was found for 4 participants, or 66.6% of cases in which it was possible to determine correspondence. Thus, the QABF showed higher correspondence with analog assessments than the MAS.

The current study supports the previous findings of Paclawskyj et al. (2001) that correspondence can be obtained between the QABF, MAS and EFA. In the present study 1 of 7 (14%) EFAs showed uninterruptable patterns of responding. Correspondence could not be assessed because no function was identified in the EFA; therefore, it is possible that anecdotal assessments for Jolinda more accurately identified the function of her SIB. Staff reports and clinic observations suggested a tangible function, as was indicated by anecdotal assessment outcomes, but in the absence of a clear EFA or function-based treatment outcomes the validity of anecdotal assessment results remains unconfirmed.

Some limitations of the present study are worth noting. First, contingencies for J.J.'s EFA were placed on a precursor to the target behavior instead of the target behavior evaluated in his anecdotal assessments. However, a systematic analysis of treatment based on J.J.'s precursor EFA was conducted, and the outcomes provided further evidence supporting the validity of the assessment (Dracobly, 2009). Second, procedures should be refined for undifferentiated EFAs, such as Jolinda's EFA, to attempt to better approximate conditions that occur in the natural environment. Third, both participants whose anecdotal assessments indicated a sensory reinforcement function presented with pica. It is possible that the operant function of pica is particularly discriminable; therefore, future investigations should make efforts to include participants who exhibit a variety of behaviors maintained by automatic reinforcement. Finally, treatments based on the operant functions identified by anecdotal assessments and experimental functional analysis were not evaluated; showing effective treatment in natural environments would lend additional external validity to the current findings.

The results of the current study, combined with those from previous investigations, suggest multiple-respondent anecdotal assessment represents a promising approach to functional

assessment. Both the MAS and QABF showed agreement among 4 or 5 respondents in a little over half of the cases assessed, with the QABF showing slightly higher rates of agreement among respondents. Furthermore, for all cases in which differentiated EFA results were compared with anecdotal assessments, correspondence was observed for the QABF. This information could be important to clinicians in settings where resources are limited and it is necessary to assess behavior quickly, economically, and with as little risk to participants as possible. Results from this study build on previous research (Fahrenholz, 2004) suggesting that multiple-respondent anecdotal assessment may represent a fast and efficient means of identifying the operant function of problem behavior for many participants (Matson et al., 1999).

Table 1

Resident's Demographic Information, Target Behaviors and Topographical Descriptions

Resident	Age	Functioning Level	Target Behavior	
Annie	31	Profound	PAO	Scratching, hitting, biting, grabbing another person
Asa	51	Profound	AGP	Biting furniture or overturning furniture, slam doors
			SIB	Banging head on hard surfaces & biting hand
Barbara	41	Severe	SIB	Hand mouthing
Carl	39	Mild	VDB	Verbally abusive threatening behavior
			PAO	Hitting, scratching, grabbing, kicking, spitting at others
Chad	27	Profound	SIB	Head banging
			PAO	Making contact with another person which causes injury to the other person
Derek	40	Profound	PDB	Bucking in wheelchair, hitting, or pushing objects
			PAO	Pushing, kicking, slapping or biting others
Donnie	56	Profound	PDB	Displacing training materials, overturning furniture, throwing objects, stripping beds & expelling mucus
			VDB	Brief loud yelling or screaming
Garfield	47	Profound	PAO	Hitting, biting, pinching, shoving or pushing others
			STE	Stealing or attempting to steal food items
Genna	57	Mild	VDB	Yelling, threatening , cursing or whining
Greg	46	Profound	PICA	Ingesting non-food items
Jack	56	Moderate	VDB	Yelling
			PAO	Hitting others
Jerry	66	Profound	PAO	Hitting, kicking, biting, scratching, spitting at, pushing, throwing objects at others
			RUM	Bringing up and re-chewing stomach contents
J.J.	29	Mild	SIB	Hitting, slapping self
Joe	52	Profound	PAO	Hitting, wrapping arms around others and bringing them to the ground.
Jolinda	55	Profound	SIB	Biting fingers/hands, head banging, slapping self
			SIB-PICK	Picking at skin, sores or other scabs
Jon	34	Severe	STEO	twirling shirt, tapping on objects, hording items (socks), seeking out object to the exclusion of all else.
Karen	47	Profound	VDB	Yelling
			SIB	Hand biting
Kate	46	Profound	MO	Mouthing Objects
Mark	50	Profound	SIB-SCR	Scratching skin, itching skin or rubbing skin
			SIB	Striking self, hitting elbow or body part against hard object or surface

(table continues)

Table 1 *(continued)*

Resident	Age	Functioning Level	Target Behavior	
Marion	58	Profound	PICA	Attempting to ingest non-food items
			SIB	Hand Mouthing
Martin	48	Profound	PDB	Slapping tables & walls, throwing materials, stripping, grabbing others clothes while yelling
Mike	50	Profound	PAO	Pushing others
			STE	Taking food items from others
Peg	50	Severe	VDB	Yelling or continuously talking for more 1 min. or more about inappropriate topics (e.g. telling on others or blaming others)
Rob	49	Profound	VDB	Yelling and screaming
			PDB	Slamming doors, pounding windows, dropping to floor, wheelchair bucking and public masturbation
Ted	52	Profound	PAO	Kicking, biting and hitting
Vern	41	Profound	PICA	Ingesting non-food items
Vynita	48	Profound	SIB	Picking/scratching scores & scabs
			STE	Taking food/drink

Table 2

Respondent Agreement Across and Within Assessments

	Total Agreement	4 of 5 Agreement	5 of 5 Agreement
QABF	24 of 42 (57%)	17 of 42 (40%)	7 of 42 (17%)
MAS	22 of 42 (52%)	10 of 42 (24%)	12 of 42 (29%)
MAS & QABF	11 of 42 (26%)	8 of 42 (19%)	3 of 42 (7 %)

Table 3

Individual Results, With Respondent Groups Listed Across Primary Maintaining Consequences

Resident	Behv.	QABF					MAS			
		N/S	ATT	TAN	ESC	PH	SEN	ATT	TAN	ESC
Annie										
	PAO	0	0	4	3	1	0	0	5	0
Asa										
	AGP	2	0	1	3	0	0	2	1	3
	SIB	0	0	1	5	0	1	0	3	3
Barbara										
	SIB	0	4	0	0	0	1	4	0	0
Carl										
	VDB	0	0	1	5	0	1	1	3	0
	PAO	1	1	0	3	0	1	2	3	1
Chad										
	SIB	0	4	2	1	0	2	0	5	1
	PAO	0	2	3	1	0	0	1	5	0
Derek										
	PDB	0	0	0	3	2	0	0	5	1
	PAO	0	0	0	4	1	0	0	4	1
Donnie										
	PDB	4	1	0	0	0	2	2	2	0
	VDB	3	2	0	0	1	3	1	1	0
Garfield										
	PAO	0	0	2	3	0	1	0	2	2
	STE	0	0	5	1	0	3	0	2	0
Greg										
	PICA	4	0	0	1	0	5	0	0	0
Genna										
	VDB	0	0	1	5	0	0	1	3	3
Jack										
	VDB	0	1	1	4	0	0	3	2	0
	PAO	0	1	2	4	0	0	4	3	0
J.J.										
	SIB	1	4	0	0	0	4	1	0	0
Jerry										
	PAO	0	0	4	1	0	1	0	3	1
	RUM	5	0	1	0	0	5	0	0	0

Note. On the QABF, N/S = non-social; ATT = attention; TAN = tangible; ESC = escape; PH = physical. On the MAS, SEN = sensory; ATT = attention; TAN = tangible; ESC = escape. Shaded areas indicate agreement of 4 or 5 respondents.

(table continues)

Table 3 *(continues)*

Resident	Behv.	QABF					MAS			
		N/S	ATT	TAN	ESC	PH	SEN	ATT	TAN	ESC
Joe										
	PAO	0	3	2	1	0	1	3	0	1
Jolinda										
	SIB	0	2	4	0	0	0	2	4	0
	PIC	2	0	3	0	0	2	2	1	0
Jon										
	STEO	1	1	3	1	0	3	1	1	0
Karen										
	VDB	0	0	1	4	0	3	0	2	4
	SIB	0	0	1	3	1	1	1	2	1
Kate										
	MO	3	1	0	1	1	5	0	0	0
Martin										
	PDB	4	1	0	0	1	4	1	0	0
Mark										
	SCR	4	0	0	0	2	5	0	0	1
	SIB	4	0	1	0	0	4	0	1	0
Marion										
	PICA	2	2	1	0	0	1	0	4	0
	SIB	1	2	2	0	0	1	0	3	1
Mike										
	PAO	1	0	3	1	0	0	0	5	0
	STE	1	0	4	0	0	4	1	0	0
Peg										
	VDB	2	2	0	1	1	2	1	1	1
Rob										
	VDB	2	1	1	2	1	0	3	1	2
	PDB	1	1	0	2	1	0	3	1	1
Ted										
	PAO	1	2	2	0	0	1	0	3	2
Vern										
	PICA	5	0	0	0	0	5	0	0	0
Vynita										
	SIB	5	0	0	0	1	5	0	0	1
	STE	3	0	4	0	0	5	0	0	0

Note. On the QABF, N/S = non-social; ATT = attention; TAN = tangible; ESC = escape; PH = physical. On the MAS, SEN = sensory; ATT = attention; TAN = tangible; ESC = escape. Shaded areas indicate agreement of 4 or 5 respondents.

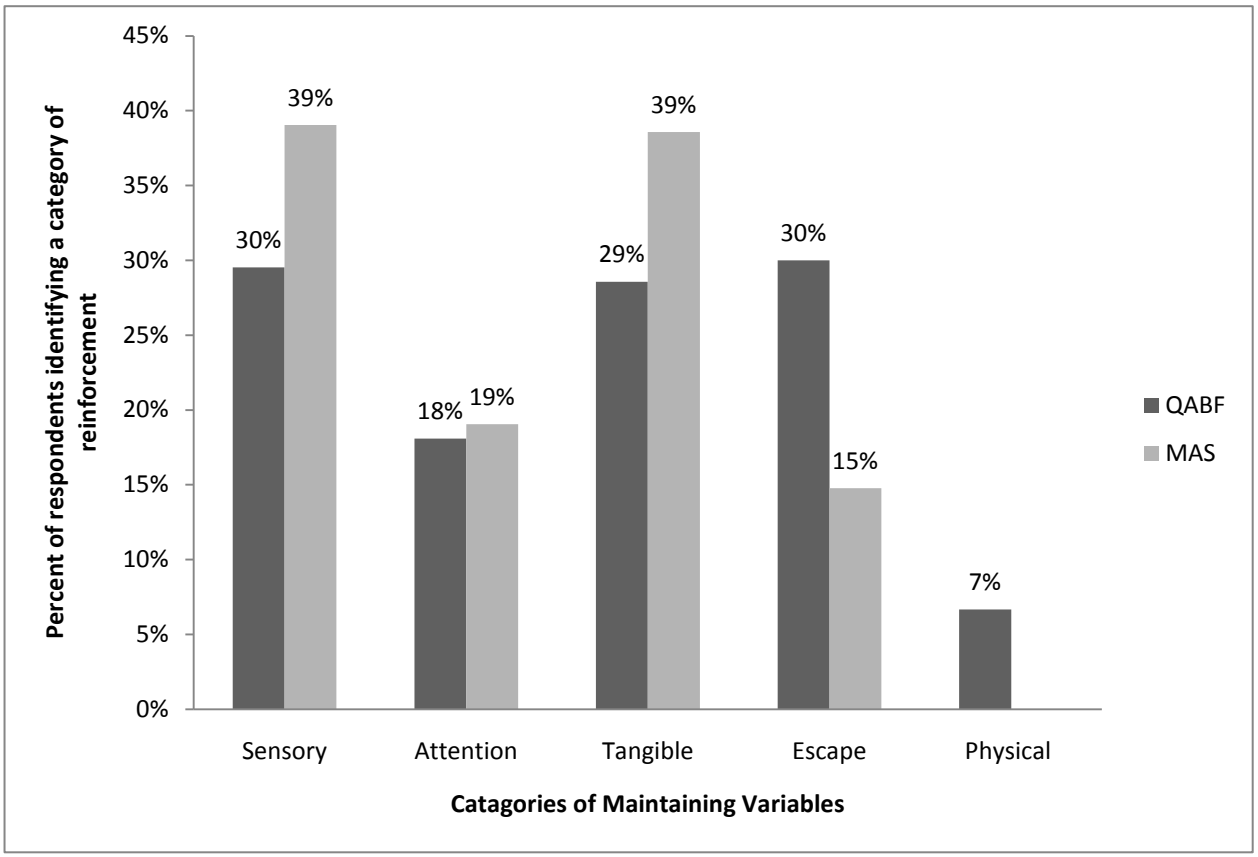

Figure 1. Categories of reinforcement that were identified by respondents of the QABF are displayed by a black bar (left) and MAS results are identified by a gray bar on the right.

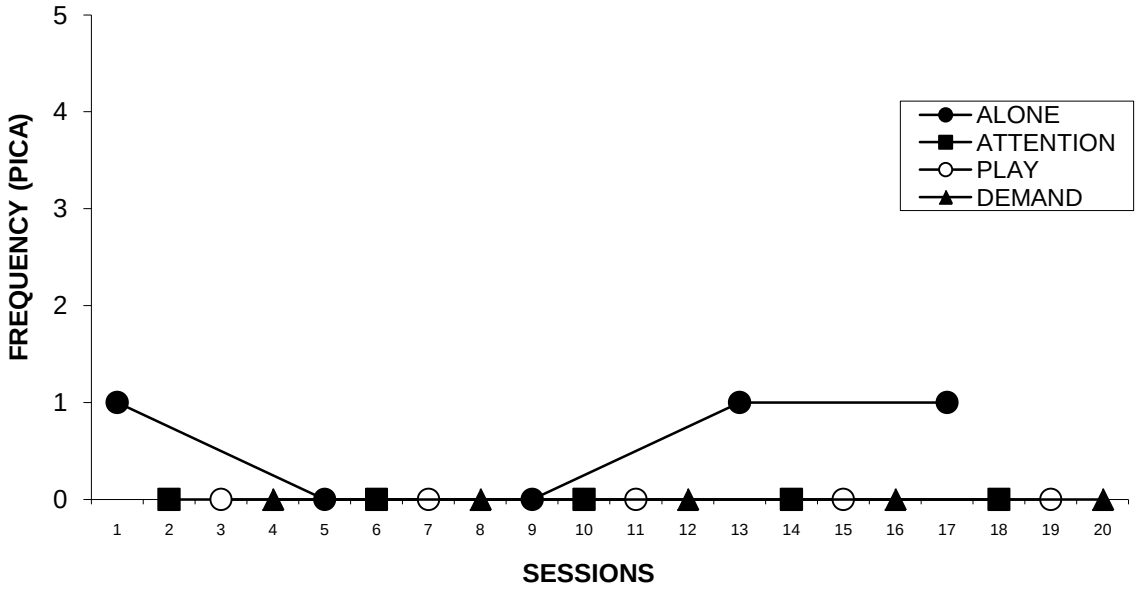

Figure 2. Greg's experimental functional analysis graph.

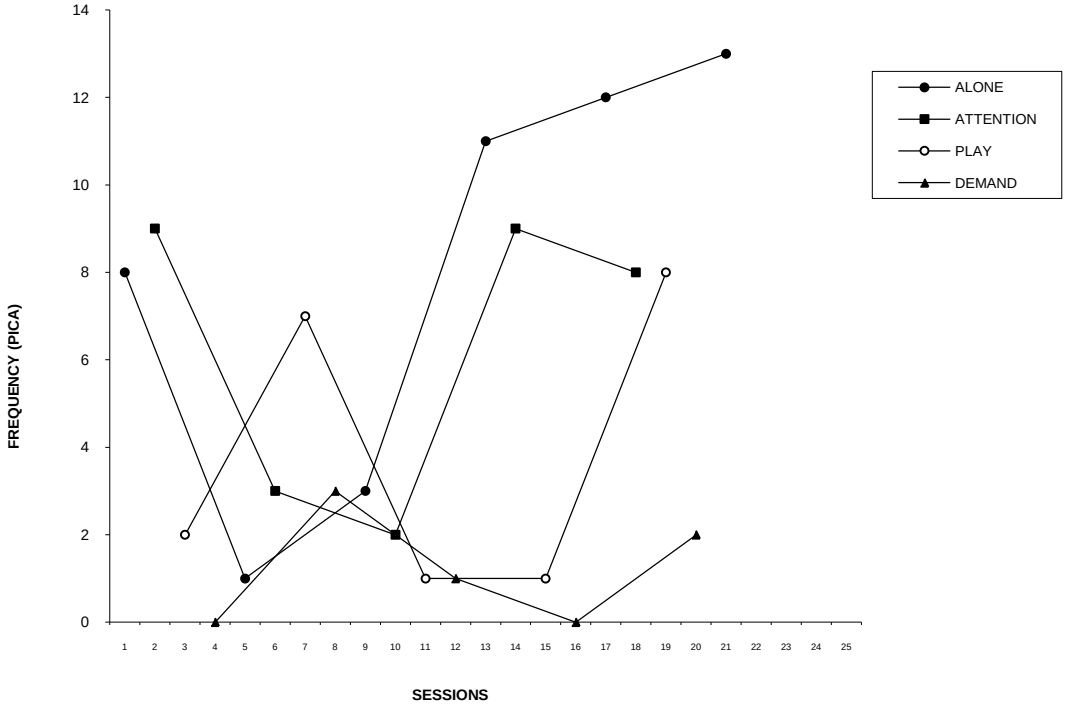

Figure 3. Vern's experimental functional analysis graph

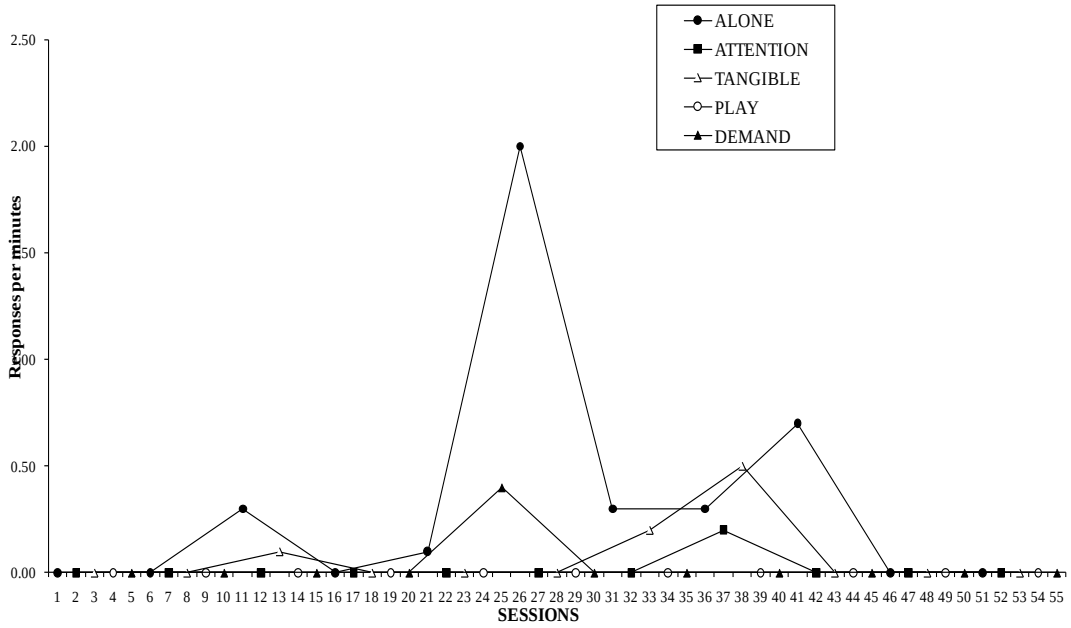

Figure 4. Jolinda's experimental functional analysis graph.

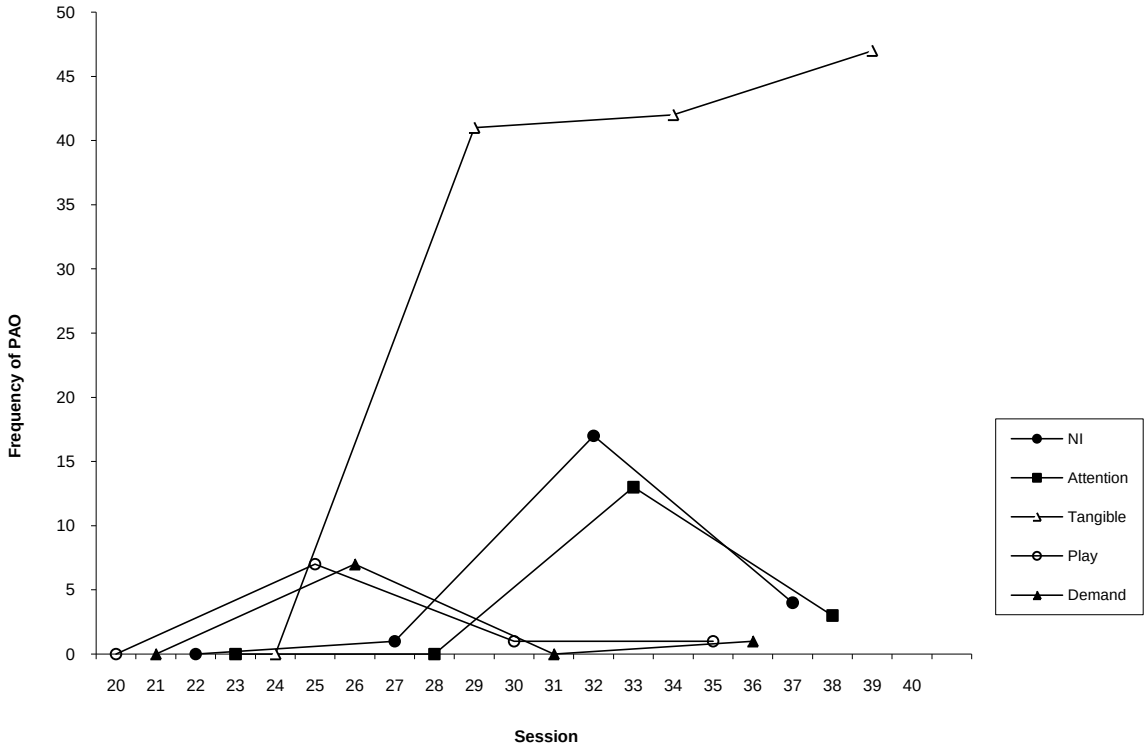

Figure 5. Annie's experimental functional analysis graph.

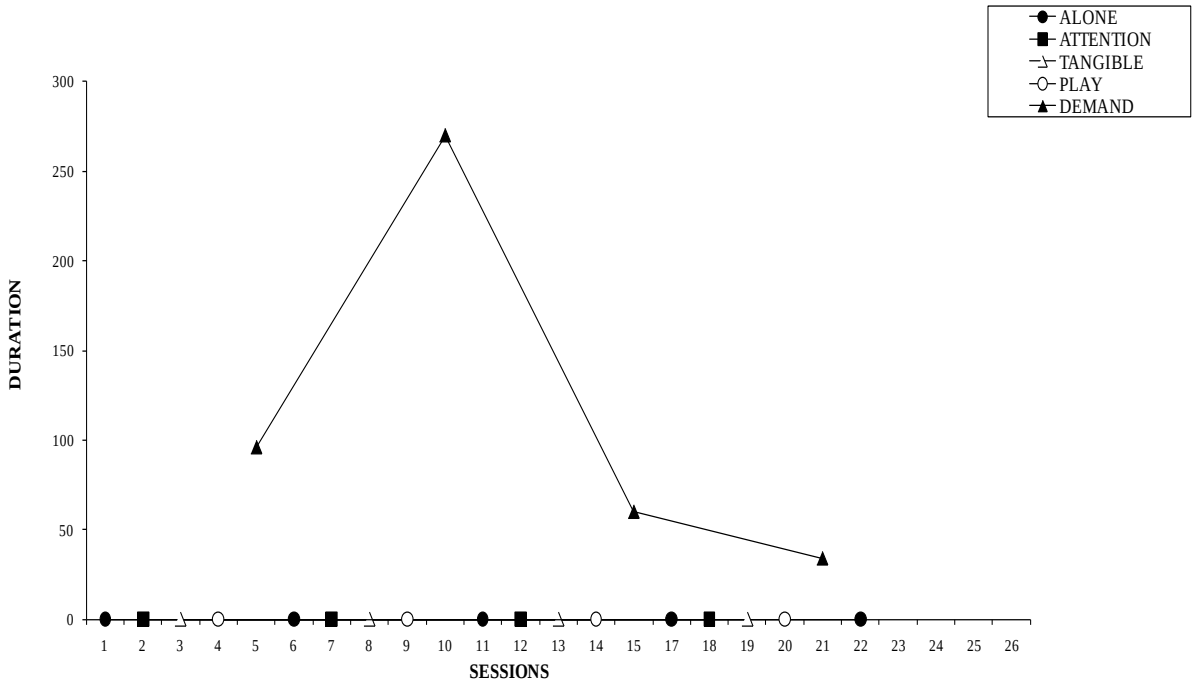

Figure 6. Karen's experimental functional analysis graph.

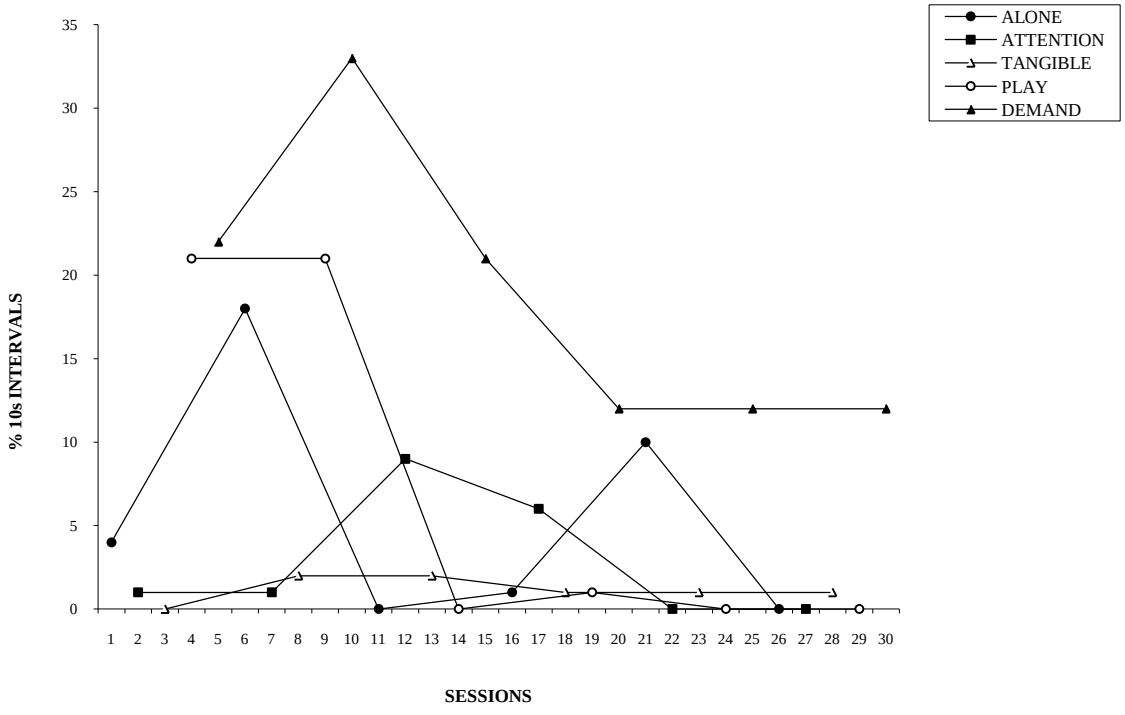

Figure 7. Asa's experimental functional analysis graph.

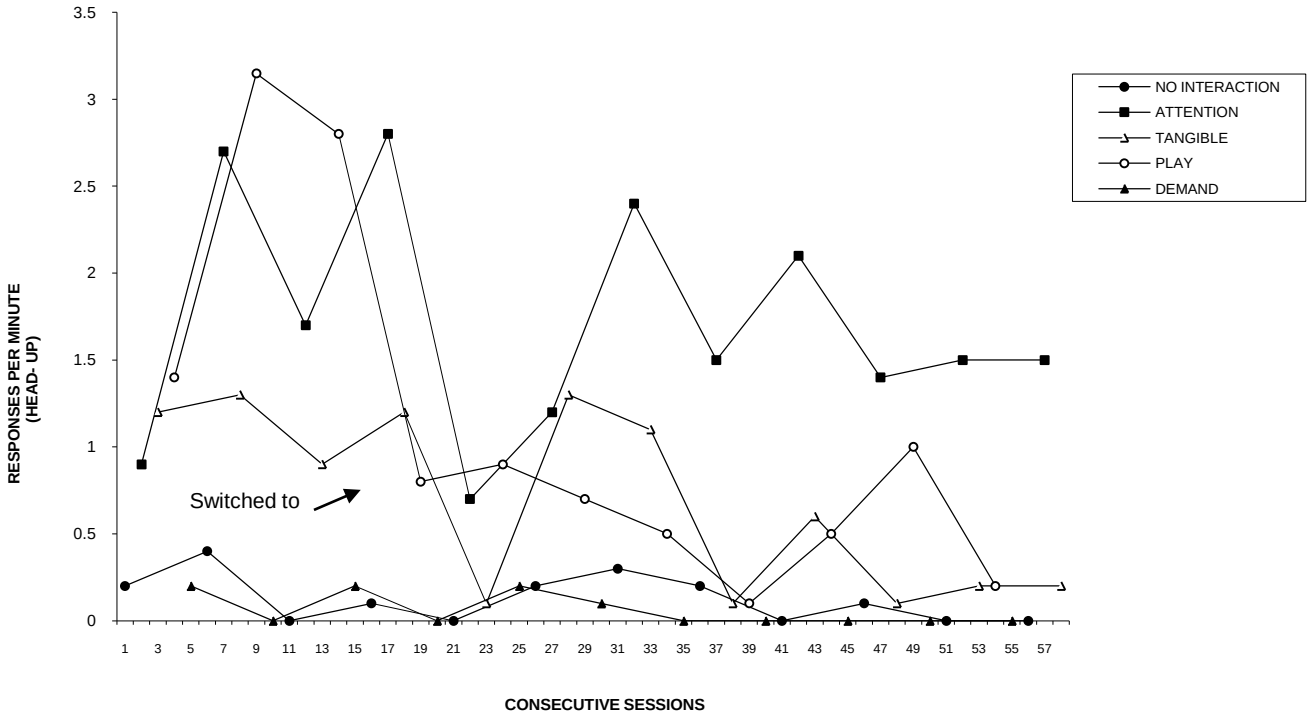

Figure 8. J.J.'s precursor experimental functional analysis graph.

Table 4

Correspondence Between the Anecdotal Assessment and the Experimental Functional Analysis

Resident	Behv.	QABF					MAS				EFA
		N/S	ATT	TAN	ESC	PH	SEN	ATT	TAN	ESC	
Annie	PAO	0	0	4	3	1	0	0	5	0	TAN / Agree
Asa	SIB	0	0	1	5	0	1	0	3	3	ESC / Agree
Greg	PICA	4	0	0	1	0	5	0	0	0	N/S / Agree
J.J.	SIB	1	4	0	0	0	4	1	0	0	ATT / Agree
Jolinda	SIB	0	2	4	0	0	0	2	4	0	Unclear EFA
Karen	VDB	0	0	1	4	0	3	0	2	4	ESC / Agree
Vern	PICA	5	0	0	0	0	5	0	0	0	N/S / Agree

Note. On the QABF, N/S = non-social; ATT = attention; TAN = tangible; ESC = escape; PH = physical. On the MAS, SEN = sensory; ATT = attention; TAN = tangible; ESC = escape. Shaded areas indicate agreement of 4 or 5 respondents.

REFERENCES

Applegate, H., J. E., Matson, J. L., & Cherry, K. E. (1999). An evaluation of functional variables affecting severe problem behaviors in adults with mental retardation by using the Questions about Behavioral Function Scale (QABF). *Research in Developmental Disabilities, 20,* 229-237.

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.

Bihm, E., Kienlen, T., Ness, E., & Poindexter, A. (1991). Factor structure of the motivation assessment scale for persons with mental retardation. *Psychological Reports, 68,* 1235-1238.

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.

Carr, E. G. (1977). The motivation of self-injurious behavior: A review of some hypotheses. *Psychological Bulletin, 84,* 800-816.

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18,* 111-126.

DeLeon, I. G. & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforce preferences. *Journal of Applied Behavior Analysis, 29,* 519-532.

Dracobly, J. D. (2009). *Progressing from identification and functional analysis of precursor behavior to treatment of self-injury.* (master's thesis). University of North Texas, Denton, Texas.

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

Durand, V. M. & Crimmins, D. B. (1988). *The Motivation Assessment Scale (MAS) administration guide.* Topeka, KS: Monaco & Associates.

Fahrenholz, A. R. (2004). *Multiple-respondent anecdotal assessments for behavior disorders: An analysis of interrater agreement and correspondence with functional analysis outcomes* (master's thesis). University of North Texas, Denton, Texas.

Fisher, W. W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25,* 491-498.

Hanley, G. P., Iwata, B. A. & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36,* 147-185.

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,* 1-20.

Iwata, B. A., Pace, G. M., Kalsher, M. J., Cowdery, G. E., & Cataldo, M. F. (1990). Experimental analysis and extinction of self-injurious escape behavior. *Journal of Applied Behavior Analysis, 23,* 11-27.

Iwata, B. A., Vollmer, T. R., & Zarcone, J. R. (1990). The experimental (functional) analysis of behavior disorders: Methodology, applications, and limitations. In A. C. Repp & N. N. Singh (Eds.), *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities* (pp. 301-330). Sycamore, IL: Sycamore.

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

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., Rodgers, T. A., Lerman, D. C., Shore, B. A., Mazaleski, J. L., Goh, H. L., Cowdery, G. E., Kalsher, M. J., McCosh, K. C., Willis, K. D. (1994). The functions of self-injurious behavior: An experimental-epidemiological analysis. *Journal of Applied Behavior Analysis, 27,* 215-240.

Iwata, B. A. & DeLeon, I. G. (1996). *The Functional Analysis Screening Tool (FAST).* Unpublished manuscript, University of Florida, Gainesville, FL 32611.

Iwata, B. A., Kahng, S., Wallace, M. D., & Lindberg, J. S. (2000). The functional analysis model of behavioral assessment. In J. Austin & J.E. Carr (Eds.), *Handbook of applied behavior analysis.* Reno, NV: Context Press.

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

Mace, F. C. (1994). The significance and future of functional analysis methodologies. *Journal of Applied Behavior Analysis, 24,* 385-392.

Mace, F. C. & Lalli, J. S. (1991). Linking descriptive and experimental analyses in the treatment of bizarre speech. *Journal of applied Behavior Analysis, 24,* 553-562.

Matson, J. L., Bamburg, J. W. & Cherry, K. E., Paclawskyj, T. R., (1999). A validity study of the Questions About Behavioral Function (QABF) Scale: Predicting treatment success for self-injury, aggression and stereotypies. *Research in Developmental Disabilities, 20,* 163-176.

Matson, J. L., & Vollmer, T. R. (1995). *The Questions About Behavioral Function (QABF) user's guide.* Baton Rouge, LA: Scientific Publishers Inc.

Neef, N. A. & Peterson, S. M. (2007). Functional behavior assessment. In J.O. Cooper, T.E. Heron & W.L. Heward (Eds.), *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Newton, J. T. & Sturmey, P. (1991). The Motivation Assessment Scale: Inter-rater reliability and internal consistency in a British sample. *Journal of Mental Deficiency Research, 35,* 472-474.

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.

Paclawskyj, T. R., Matson, J. L., Rush, K. S., Smalls, Y. & Vollmer, T. R. (2001). Assessment of the convergent validity of the Questions About Behavioral Function (QABF) Scale with analogue functional analysis and the Motivation Assessment Scale: A behavioral checklist for functional assessment of aberrant behavior. *Journal of Intellectual Disability Research,, 45,* 484-494.

Shogren, K. A. & Rojahn, J. (2003). Convergent reliability and validity of the Questions About Behavioral Functions and the Motivation Assessment Scale: A replication study. *Journal of Developmental and Physical Disabilities, 15,* 367-375.

Sigafoos, J., Kerr, M. & Roberts, D. (1994). Interrater reliability of the3 Motivation Assessment Scale: failure to replicate with aggressive behavior. *Research in Developmental Disabilities, 15,* 333-342.

Singh, N. N, Donatelli, L. S., Best, A., Williams, D. E., Barrera, F. J., Lenz, M. W., Landrum, T. J., Ellis, C. R., & Moe, T. L. (1993). Factor structure of the motivation assessment scale. *Journal of Intellectual Disability Research, 37,* 65-74.

Smith, R. G., & Churchill, R. M. (2002). Identification of environmental determinants of behavior disorders through functional analysis of precursor behaviors. *Journal of Applied Behavior Analysis, 35,* 125-136.

Spreat, S. & Connelly, L. (1996). Reliability analysis of the motivation assessment scale. *American Journal of Mental Retardation, 100,* 528-532.

Sturmey, P. (1995). Analog baselines: A critical review of the methodology. *Research in Developmental Disabilities, 16,* 269-284.

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.

Thompson, S. & Emerson, E. (1995). Inter-informant agreement on the Motivation Assessment Scale: another failure to replicate. *Mental Handicap Research, 8,* 203-208.

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.

Vollmer, T. R., Marcus, B., Ringdahl, J., & Roane, H. (1995). Progressing from brief assessments to extended experimental analyses in the evaluation of aberrant behavior. *Journal of Applied Behavior Analysis, 28,* 561-576.

Zarcone, J. R., Rodgers, T. A., Iwata, B.A., Rourke, D. A., & Dorsey, M. F., (1991). Reliability analysis of the Motivation Assessment Scale: A failure to replicate. *Research in Developmental Disabilities, 12,* 349-360.