MULTIPLE-RESPONDENT ANECDOTAL ASSESSMENTS FOR BEHAVIOR DISORDERS: AN ANALYSIS OF INTERRATER AGREEMENT AND CORRESPONDENCE WITH A FUNCTIONAL ANALYSIS AND TREATMENT OUTCOMES

Heather Moore, B.S.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

December 2007

APPROVED:

Richard G. Smith, Major Professor and Chair of the Department of Behavior Analysis
Shahla Ala’i Rosales, Committee Member
Janet Ellis, Committee Member
Thomas Evenson, Interim Dean of the College of Public Affairs and Community Service
Sandra L. Terrell, Dean of the Robert B. Toulouse School of Graduate Studies

An analysis of interrater agreement across multiple respondents on two anecdotal assessments, the Motivation Assessment Scale (MAS) and the Functional Analysis Screening Tool (FAST), was completed for an individual who displayed aggressive behavior. The results of the assessments indicated high agreement across assessments and respondents that the problem behavior was maintained by social positive reinforcement in the form of contingent delivery of tangible items. By contrast, a subsequent experimental analysis indicated that the behavior was maintained by escape from demands. A treatment was implemented based on the functional analysis outcomes to determine if the functional analysis had correctly identified the maintaining variable of the aberrant behavior. Results of the treatment analysis showed significant reductions in the occurrence of aberrant behavior suggesting that the MAS and FAST may not have accurately identified the maintaining variable of the aberrant behavior.
Copyright 2007
by
Heather Moore
ACKNOWLEDGEMENTS

I would like to begin by extending my gratitude to Dr. Richard Smith. I am thankful for the wisdom, support, encouragement and opportunities that he has provided to me throughout the many years I have been part of the department. In addition, I would like to thank all of the professors in the behavior analysis department for helping to shape my behavioral repertoire. I have learned so much from each of them and feel fortunate to have been involved in a program with such expertise. Furthermore, I am thankful to my many student colleagues for their support in my thesis. A special thanks is warranted to Jessica Hobbs and Bryan Lovelace for giving their time and assistance on a daily basis. Lastly, but certainly not least, thank you to my wonderful family. Thanks to my dad for his insistence on me attending graduate school and for the financial backing to make it possible. Many, many thanks to my loving husband, Brian, who has provided endless support and encouragement throughout my college education, and thanks to my two sweet children who have been so patient while I devoted numerous hours to my studies.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>GENERAL METHOD</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Participants and Setting</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>ANECDOTAL ASSESSMENT</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Setting and Materials</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Administration Procedures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Respondent Agreement Evaluation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>FUNCTIONAL ANALYSIS</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Setting</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Response Definition and Measurement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>General Procedure</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>TREATMENT</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Participant and Setting</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Response Definition and Measurement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experimental Design</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Treatment Analysis Procedures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>DISCUSSION</td>
<td>32</td>
</tr>
<tr>
<td>7.</td>
<td>REFERENCE LIST</td>
<td>47</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1. Table of primary maintaining variable listed across respondents for the MAS and FAST .............................................16
2. Functional Analysis Graph ..................................................................................................................39
3. Functional Analysis: Average Frequency within session .........................................................40
4. Treatment Analysis: Problem behavior and “No” ................................................................41
5. Treatment Analysis: PAO and AGP ...........................................................................................42
6. Conditional Probabilities: PAO within a trial ...........................................................................43
7. Conditional Probabilities: AGP within a trial ...........................................................................44
8. Conditional Probabilities: PAO within 5 s of a trial ending ..................................................45
9. Conditional Probabilities: AGP within 5 s of a trial ending ..................................................46
CHAPTER 1
INTRODUCTION

Developing interventions for challenging behaviors is a primary responsibility for behavior analysts working with persons with developmental disabilities. Problem behaviors often make it difficult to implement instructional programming, prohibit inclusion into less restrictive environments, and, perhaps, most importantly, may cause injury to the individual engaging in the behavior or others in the immediate environment. In order to select effective interventions for challenging behaviors it is necessary to determine why the behavior is occurring. A variety of procedures, collectively known as “functional assessment,” have been developed to identify the environmental events that evoke and maintain challenging behavior. Three general approaches to functional assessment have been identified, including indirect, or anecdotal methods; direct observation, or descriptive methods; and analog, or experimental (also known as functional analysis) methods (Miltenberger, 1997).

Each method of functional assessment shares a common goal of aiding in the development of interventions to decrease challenging behavior and increase appropriate behavior. However, each method has unique procedural characteristics, as well as relative strengths and weaknesses. For example, indirect methods involve gathering information through interviews and questionnaires with individuals (e.g., parents, teacher, or other caregivers) who frequently interact with the person engaging in the challenging behavior. This may be conducted informally with caregivers giving anecdotal reports or in a more structured format utilizing interviews and questionnaires designed specifically for the purpose of functional assessment, including behavioral
interviews (Lalli, Browder, Mace & Brown, 1993), the Motivation Assessment Scale (MAS; Durand & Crimmins, 1988), or the Functional Assessment Screening Tool (FAST; Iwata & DeLeon, 1995). Anecdotal assessment provides initial information regarding the challenging behavior and the conditions in which it is more or less likely to occur. Anecdotal assessments are relatively quick and easy to conduct; however, their reliability and validity have been questioned due to a reliance on caregivers’ memory and possible biases regarding the challenging behavior and the individual (Kearney, 1994; Kwak, Ervin, Anderson, & Austin, 2004; Zarcone, Rodgers, Iwata, Rourke, & Dorsey, 1991).

Direct observation, or descriptive assessment, consists of gathering information by an observer who collects objective information about problem behavior in the environment in which it typically occurs (Miltenberger, 1997). Descriptive procedures may document occurrences of antecedents, challenging behavior, consequences, and/or other relevant conditions and events as they occur in the natural environment. Data may be documented using different formats, including a scatterplot (Touchette, MacDonald, & Langer, 1985), antecedent-behavior-consequence (ABC) recording (Miltenberger, 1997), ABC checklists (Miltenberger, 1997), and continuous or interval-based recording procedures (Hayes, Barlow, Nelson-Gray, 1999). The advantage of descriptive assessment, relative to anecdotal assessment, is that there is not a reliance on the subjective and potentially faulty memory of a third party for information, because the observer records events as they occur and, thus, the information is typically assumed to be more accurate. However, these assessments generally are more time
consuming to conduct and may not be practical for challenging behaviors that occur infrequently.

Experimental (or analog, or functional analysis) methods involve systematically manipulating antecedents and consequences and observing their effects on behavior (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994). These procedures are able to demonstrate a functional relationship between controlling variables and challenging behavior. There has been abundant research on the utility of experimental methods to identify the determinants of a wide range of problem behavior (Neef & Iwata, 1994); however, they are often time consuming to conduct; they require trained therapists and data collectors to be involved; and professional expertise to interpret the data (Miltenberger, 1997).

Research has demonstrated that the “use of functional assessment can be a beneficial and effective approach to understanding and overcoming challenging behavior” (Reid, 2000, p.241) and, as a result, functional assessment is being utilized in many institutions and clinical settings prior to development of treatment (Ellingson, Miltenberger, & Long, 1999). However, the variety of assessments available for use, with their associated strengths and limitations, suggests a need to determine the most efficient way to consistently produce results that accurately identify the maintaining variables of aberrant behavior.

Because anecdotal methods are generally viewed as the easiest and least time consuming of functional assessment procedures, determining which anecdotal assessments produce outcomes that allow for the development of effective treatment would be beneficial. Several studies have evaluated the general utility of the MAS
For example, Durand and Crimmins (1988), developers of the MAS, evaluated its reliability by administering it to primary and secondary raters (teachers and teacher assistants) who provided responses about the self-injurious behavior (SIB) of children in their classrooms. The results indicated that item correlations across raters ranged from .66 to .92 and that agreement on functional condition ranks ranged from .66 to .82. Durand and Crimmins concluded that “raters can agree on the category of maintaining variables presumably influencing a particular child’s self-injurious behavior” (p.104). Subsequently, Durand and Crimmins evaluated the effectiveness of treatments designed to correspond to MAS outcomes for a subgroup of the children who received assessments. Outcomes indicated that treatments based on the MAS outcomes effectively decreased SIB for all participants. Zarcone and colleagues (Zarcone et al., 1991) systematically replicated the original study by Durand and Crimmins, administering the MAS across primary and secondary raters for 55 different individuals who engaged in SIB. Their results, in contrast with those from Durand and Crimmins’ previous study, indicated low interrater agreement with agreement on item correlation ranging from -.51 to .55 and agreement on functional condition ranks ranging from -.80 to 1.00. Zarcone et al. (1991) concluded that a failure to do so should “raise serious questions about the utility of the MAS as a diagnostic or prescriptive tool for clinical as well as research purposes” (p.357). Several additional evaluations of the MAS have yielded inconsistent outcomes; although most have failed to replicate Durand and Crimmins’ initial findings of high measures of reliability and validity (e.g., Conroy, Fox, Bucklin, & Good, 1996; Newton & Sturmey, 1991; Sigafoos,
Kerr, & Roberts, 1994), others have shown more encouraging outcomes (e.g., Bihm, Kienlen, Ness, & Poindexter, 1991; Kearney et al., 1994). Although reasons for these inconsistent outcomes remain unclear, some potential influences that have been suggested in the literature include differences in population characteristics (e.g., Zarcone et al., 1991), differences in respondent characteristics (e.g., Kearney et al, 1994; Zarcone et al., 1991), and the frequency of challenging behavior (e.g., Kearney et al., 1994).

Fahrenholtz and Smith (2004) extended this area of research by assessing agreement among multiple respondents and multiple assessment tools (MAS and FAST), and evaluating correspondence between the outcomes of these assessments and those from experimental analyses. Initially, overall agreement among multiple respondents using the MAS and the FAST was evaluated for 28 individuals who engaged in self-injury or physical aggression. For each participant, five respondents (staff members and administrators at a large residential and training facility) were questioned using the MAS and FAST. Results indicated that, using the MAS, 4 out of 5 respondents agreed on the primary hypothesized variable for 54% of participants. Using the FAST, 4 out of 5 respondents agreed on the primary hypothesized variable for 79% of the participants. Agreement across assessments within respondents was scored when a single respondent was consistent across assessments in identifying a hypothesized maintaining variable; this occurred for 77% of the participants. Subsequently, experimental functional analyses were conducted with 6 of the individuals for whom 4 out of 5 respondents had agreed (across MAS and FAST assessments) on the hypothesized operant function of their problem behavior. The
functional analyses for 4 of the 6 participants produced outcomes that were consistent with those produced by the anecdotal assessments. For the other 2 participants, results of the experimental analyses were inconclusive. The results of this study suggest that agreement among multiple respondents using MAS and FAST assessments is predictive of corresponding functional analysis results. A limitation of this study was that function-based treatments were not implemented to further evaluate the validity of the anecdotal assessment outcomes.

The purpose of the current study was to further evaluate the reliability and validity of the MAS and FAST by replicating and extending the Fahrenholz and Smith (2004) study. The goal of this study was to answer 3 experimental questions. First, to what extent would multiple respondents using multiple assessments agree on the hypothesized maintaining variable of aberrant behavior of a participant? Second, would the results of an experimental functional analysis correspond with the results of the anecdotal assessments? Finally, would a treatment based on the hypothesized function suggested by the outcomes of the anecdotal and experimental assessments yield an effective intervention?
CHAPTER 2

GENERAL METHOD

Participant and Setting

The anecdotal assessments, functional analysis and treatment were conducted at a large, state-operated residential and training facility for adults with developmental disabilities.

*Participant.* The participant, John, was a Caucasian male, 58 years of age, who functioned in the severe range of mental retardation. John was able to follow simple directions; however, he had limited functional speech. At the time of the study he received daily doses of 100 mg of Zoloft® (sertraline HCl), 3000 mg of Depakote® (Divalproex sodium), 2.5 mg of Ativan® (benzodiazipines), and 450 mg of Mellaril® (thioridazine). John resided in a large residential and training facility serving persons with developmental disabilities.

John engaged in physical aggression toward others (PAO) and aggression toward property (AGP), which were targeted for reduction in a formal behavior support plan. His PAO was defined as hitting, kicking, biting, and pinching another person or throwing objects that hit another person. AGP was defined as overturning furniture; throwing objects onto the ground or against the wall; hitting objects with fists, open hand or elbow; and kicking objects with his foot. Exclusions from this definition were knocking on the table or door with his knuckles and stomping. John’s PAO was recorded by attending staff, and documented to occur an average of 146 times/month. AGP was documented to occur 93 times/month.
Respondents. The participant’s physical aggression toward others and aggression toward property were rated by five different respondents. All respondents had a minimum of a high school diploma or equivalent and had worked with the resident as direct-care staff for at least 6 months.
CHAPTER 3
ANECDOTAL ASSESSMENT

Setting and Materials

All anecdotal assessments were conducted in quiet areas of the participant’s home. The materials included 2 copies of the Motivation Assessment Scale (MAS) and Functional Assessment Screening Tool (FAST) questionnaires (1 for the respondent and 1 for the interviewer), a sheet of paper on which the MAS likert scale and what each number represented were printed in 14-point font, and ball point pens.

MAS. The MAS is a 16-item questionnaire designed to identify contingencies of reinforcement that maintain aberrant behavior (Durand & Crimmins, 1988). This assessment is organized so that each question presented corresponds with 1 of 4 categories of possible maintaining contingencies: attention, tangible, escape, or sensory. Four questions are allocated to each category and respondents are asked to rate each question using a 7-point (0-6) Likert-type scale. A score of zero indicates that the behavior never occurs under the circumstances identified in the question, a score of 6 indicates that the behavior always occurs under the circumstances identified in the question, and scores between those values indicate intermediate ratings (e.g., a score of 3 indicates that the behavior occurs “sometimes” under the circumstances indicated by the question). A score is obtained for each category by summing the points from the 4 questions in that category. Categories are then rank-ordered according to the total scores. The category with the highest score is hypothesized to be the maintaining contingency of reinforcement for the aberrant behavior.
**FAST.** The FAST is a 20-item questionnaire designed to identify contingencies of reinforcement that maintain aberrant behavior (Iwata & DeLeon, 1996). This assessment is organized so that each question corresponds with 1 of 4 possible categories of maintaining contingencies: social positive reinforcement in the form of attention and/or preferred items, social negative reinforcement in the form of escape from task or demands, automatic positive reinforcement in the form of sensory stimulation, and automatic negative reinforcement in the form of pain attenuation. Five questions are allocated to each of the categories. Each question is answered in a Yes/No format with a point awarded for every “Yes” response given by the respondent. The points are summed for each category, and the category receiving the most points is considered the hypothesized maintaining contingency.

**Administration Procedures**

Graduate and undergraduate behavior analysis students were trained to administer the anecdotal assessments. The MAS and FAST were administered individually to each respondent to reduce the possibility of influence among respondents. Both the respondent and the interviewer had a copy of the assessments. Prior to beginning the assessments, background information (e.g., name of resident, ID number of respondent, date, etc.) was collected and entered on the forms by the interviewer. Definitions of the problem behaviors, including topographical descriptions, were read aloud to the respondent. Before the MAS administration, a sheet of paper listing the Likert-scale numbers and a description of what each number represented was given to the respondent. Each question was read aloud by the interviewer to the respondent and the respondent was instructed to reply to each question with a number
from the Likert-scale (MAS) or with a Yes/No answer (FAST). If the respondent indicated difficulty answering a question or asked for clarification, the interviewer repeated the question. No additional information or clarifications were provided. The respondents were encouraged to answer each question to the best of her or his ability and were allowed as much time as necessary to answer each question. When the assessments were completed the interviewer thanked the respondent for participating, left the area, and calculated the scores.

Respondent Agreement Evaluation

Reliability of scoring. All data from the MAS and FAST assessments were calculated by 2 graduate or undergraduate students trained in administration and scoring assessments. Agreement was assessed for each question and was 100\% for both assessments.

Agreement across assessments (within respondents). Agreement across assessments was scored if the hypothesized maintaining variable identified by a single respondent was consistent across both questionnaires. The MAS categorizes the possible maintaining variables as attention, tangible, escape and sensory. The FAST categorizes the possible maintaining variables as social positive reinforcement, social negative reinforcement, automatic positive reinforcement, and automatic negative reinforcement. Due to the differences in the categorization of reinforcement contingencies, general guidelines were established in order to assess agreement across the instruments. If social positive was the highest ranked category on the FAST, agreement was scored if either attention or tangible was the highest ranked category on the MAS. If sensory was the highest ranked category on the MAS, agreement was
scored if either automatic positive or automatic negative was the highest ranked category on the FAST. If a respondent scored escape on the MAS, agreement was scored if social negative was the highest ranked category on the FAST. On some occasions, the outcome of given assessment (MAS or FAST) was a tie between two categories of reinforcement as the highest ranked category. In this event, the two categories were compared to the other highest ranked category from the other assessment to assess agreement across instruments. Agreement was scored if either of the two categories matched the highest category from the other assessment. For example, if social positive reinforcement and social negative reinforcement tied as the highest ranked categories on the FAST, agreement was scored if tangible, attention, and/or escape was the highest ranked category on the MAS.

Agreement within assessments (across respondents). Agreement across assessments was scored if 4 or 5 of the 5 respondents agreed on the hypothesized maintaining variable for either the MAS or the FAST. Ties between categories within an assessment were treated in a similar manner to the assessment of agreement across assessments; the tied categories were compared with the highest ranked category from the other respondents’ assessments to determine agreement within assessments (across respondents). For example, if Respondent 1 scored both attention and escape as the highest ranked category on the MAS, and 3 or 4 of the other 4 respondents scored either attention or escape as the highest ranked category, agreement within the MAS assessment (across respondents) was scored.

Agreement across assessments (across respondents). Agreement across assessments was scored if 4 or 5 of the 5 respondents agreed on the highest ranked
category across both the MAS and the FAST. Ties between categories within an
assessment were treated as described above (agreement within assessments [across
respondents]). Overall agreement across assessments (across respondents) was
scored only if 4 or 5 out of 5 respondents agreed on 1 maintaining variable for both the
MAS and the FAST. For example, if 4 or 5 respondents scored tangible as the
maintaining variable on the MAS and social positive reinforcement as the maintaining
variable on the FAST, then agreement was scored. However, if 2 respondents scored
attention as the highest ranked category on the MAS, 3 respondents scored tangible as
the highest ranked category on the MAS, and all respondents scored social positive on
the FAST, agreement was not scored. Although, in this case, agreement was seen
between the MAS and FAST, agreement within the MAS was not achieved.

Results

The highest ranked categories for physical aggression toward others (PAO) and
aggression toward property (AGP), as rated by each respondent, are displayed in Table
1. This table uses a scatterplot format to permit a detailed visual analyses of agreement
across assessments within and across respondents. Yellow shading indicates across
assessment (within-respondent) agreements, bold type indicates within assessment
(across respondents) agreement, and green shading indicates across assessment
(across respondents) agreement.

The results of the anecdotal assessment process show significant agreement
across each type of analysis. For PAO, all respondents scored tangible as the highest
ranked category for the MAS, and 4 of the 5 respondents scored social positive
reinforcement as the highest ranked category for the FAST. Thus, substantial
agreement within and across respondents and assessments indicated that John’s PAO was maintained by positive reinforcement, probably in the form of the delivery of tangible items. For AGP, 4 of the 5 respondents scored tangible as the highest ranked category for the MAS, and 4 of 5 respondents scored social positive reinforcement as the highest ranked category for the FAST. Thus, as with PAO, substantial agreement within and across respondents and assessments indicated that John’s AGP was maintained by positive reinforcement, probably in the form of the delivery of tangible items.

Respondents 1, 2, and 5 did not show across assessment (within respondent) agreement for one behavior each. Respondent 1’s MAS outcomes indicated that AGP was maintained by escape from demands, whereas this respondent’s FAST outcomes indicated that AGP was maintained by social positive reinforcement. Respondent 2’s MAS outcomes indicated that AGP was maintained by access to tangibles, whereas this respondent’s FAST outcomes indicated that AGP was maintained by social negative reinforcement. Respondent 5’s MAS outcomes indicated that PAO was maintained by access to tangibles, whereas this respondent’s FAST outcomes indicated that PAO was maintained by negative automatic reinforcement.

Several within-assessment ties were observed. Respondents 2 and 4 scored both social positive reinforcement and social negative reinforcement as highest ranked categories when rating John’s PAO using the FAST. Similarly, Respondents 3 and 4 scored both social positive reinforcement and social negative reinforcement as highest ranked categories when rating John’s AGP using the FAST.
The MAS results identified access to tangibles as the hypothesized maintaining consequence for John’s PAO and AGP. These outcomes corresponded with those from the FAST, which identified social positive reinforcement as the likely maintaining consequence. However, several records (2 of 5 FAST results for PAO, 3 of 5 FAST results for AGP, and 1 of 5 MAS results for AGP) identified negative reinforcement as a potential maintaining variable. In order to further evaluate the variables related to John’s PAO and AGP an experimental functional analysis was conducted.
Table 1

**Primary Maintaining Variable Listed across Respondents for John’s PAO and AGP**

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Respondent 1</th>
<th>Respondent 2</th>
<th>Respondent 3</th>
<th>Respondent 4</th>
<th>Respondent 5</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAS</td>
<td>FAST</td>
<td>MAS</td>
<td>FAST</td>
<td>MAS</td>
<td>FAST</td>
</tr>
<tr>
<td>PAO</td>
<td>Tan</td>
<td>$+$</td>
<td>Tan</td>
<td>$+$</td>
<td>Tan</td>
<td>$S$&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>AGP</td>
<td>Esc</td>
<td>$+$</td>
<td>Tan</td>
<td>$-$</td>
<td>Tan</td>
<td>$S$/&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Note: On the FAST: $S^{+}$ = Social Positive; $S^{-}$ = Social Negative; $A^{-}$ = Automatic Negative
On the MAS: Esc = Escape; Tan = Tangible

Yellow shading = within-respondent agreements
Bold type = agreement across respondents
Green shading = agreement across instruments
CHAPTER 4
FUNCTIONAL ANALYSIS

Setting

All functional analysis sessions were conducted at the Behavior Analysis Resource Center (BARC), a clinic for assessment and treatment of problem behaviors, located on the residential campus. Sessions were conducted in a 3.7-m by 2.4-m room, containing a table, two chairs, and appropriate materials for each analogue condition. The observation room was equipped with a one-way mirror suitable for unobtrusive observation and data collection.

Response Definitions and Measurement

The operational definitions used during the functional analysis were the same definitions used during the anecdotal assessments. Physical aggression toward others (PAO) was defined as hitting, kicking, biting, and pinching another person or throwing objects that hit another person. Aggression toward property (AGP) was defined as overturning furniture, throwing objects onto the ground or the wall, hitting objects with fists, open hand or elbow, and kicking objects with his foot. Excluded from this definition were knocking on the table or door with his knuckles and stomping. During all sessions a trained observer collected data on PAO and AGP using event recording. Event recording was also used to collect data on the delivery of tangibles, delivery of demands, and compliance. Delivery of attention was recorded as a duration measure. Data were recorded on handheld computers installed with Instant Data 0.8B®, data collection software.
Interobserver agreement (IOA) data were collected simultaneously but independently by a second observer during 43% of sessions. IOA was calculated by dividing each session into 1-s bins, summing the number of bins in which the observers agreed on the occurrence or nonoccurrence of target behaviors, dividing the results by the total number of seconds in the session and multiplying the result by 100. IOA scores averaged 97.2% (range = 93.9% - 99.4%).

General Procedure

The participant was exposed to four conditions similar to those described by Iwata, Dorsey, Slifer, Bauman, and Richman (1982/1994). Attention, tangible, play and demand conditions were presented within 10-min sessions using a multielement design. The conditions were conducted in the above order. One to 4 sessions were completed per day, at the same time each day, 2-3 times a week. The following is a detailed description of the experimental conditions.

Attention. The therapist was present in the observation room with the participant. Leisure materials (e.g., books, toys, cards, magazines, etc) were placed on the table. The therapist began the session by telling the participant hello and that the therapist would be doing some work, should he/she be needed. No attention was provided unless the participant engaged in target behaviors. Contingent on the target behavior, the therapist approached and made eye contact with the client and provided attention in the form of statements of concern or reprimands (e.g. “Do you need something?” or “Don’t do that, you’re going to hurt yourself!”). This condition was designed to evaluate whether the target behavior was maintained by social positive reinforcement in the form of attention.
**Tangible.** Prior to this condition, a multiple-stimulus without replacement preference assessment was completed for the participant (DeLeon, 1996). Leisure and edible preference assessments were conducted separately. The leisure preference assessment identified a western movie as the most preferred leisure item/activity and the edible preference assessment identified pudding as most preferred food. Movies were selected to use during the functional analysis because anecdotal observations indicated that the presentation of soda occasioned disruptive and aggressive behavior (e.g., throwing soda on the therapist). During tangible sessions the therapist was in the observation room with the participant. Prior to the session, the therapist allowed the participant to select a movie from 4 available movies. The therapist started the movie and the participant was allowed to view the movie for 2 min, at which time the therapist turned off the movie and the session began. The therapist stood in the room and did not attend to the participant. At any time if the participant emitted a target behavior the therapist turned the movie on for 2 min. This condition was designed to evaluate if the target behaviors were maintained by social positive reinforcement in the form of access to a leisure item/activity (movies).

**Play.** The therapist was in the observation room with the participant. Leisure items (e.g., books, toys, cards, magazines, etc.) were placed on the table. No demands were given, and the therapist interacted with the subject at least once every 30 s (e.g., talked about leisure items, the participant’s day, prompted engagement with activities, etc). Target behaviors were ignored. This condition was designed to serve as a control condition against which to compare levels of target responding observed in test conditions.
Demand. During this condition, the therapist was in the observation room with the participant. Every 30 s the therapist initiated a task trial (e.g., “John, clap your hands.”) using graduated prompting at 10-s intervals (verbal prompting, modeling, and physical guidance). Compliance prior to the third prompt resulted in verbal praise (e.g., “Good job!”) and withdrawal of the task demand for the remainder of the 30-s interval. Contingent on a target behavior, the trial was terminated for 30 s. This condition was designed to evaluate if the target behaviors were maintained by social negative reinforcement in the form of escape from task trials.

Results

Results of the functional analysis, summarized for both PAO and AGP, are presented in Figure 1. These data show that problem behavior occurred at higher levels in the demand condition than other conditions during the first 3 cycles of the multielement assessment but subsequently decreased to levels approximating those observed in the other conditions. During the final 4 cycles of assessment, problem behavior occurred at mean levels of 38 responses/session in the play condition, 29 responses/session in the demand condition, 23 responses/session in the tangible condition, and 9 responses/session in the tangible condition. Substantial overlap among data paths was observed during the final 4 cycles of the analysis.

Figure 2 shows the average number of problem behaviors per min, displayed across minutes of the session. Data paths correspond to the experimental conditions. These data show that PAO and AGP occurred at relatively stable levels within demand sessions, corresponding roughly to the rate at which demands were placed throughout the sessions. Relatively high levels of problem behavior were recorded during the initial
minutes of attention sessions; however problem behavior decreased steadily to zero or near-zero levels across minutes of attention sessions. Lower levels overall and a general decreasing within-session trend was observed during tangible sessions. Problem behavior persisted throughout play sessions at moderate and stable levels.

The outcomes of the functional analysis were most consistent with a social negative reinforcement account for John’s problem behavior. Although problem behavior showed a general decreasing trend across demand sessions, inspection of within-session data revealed that his problem behavior persisted at levels roughly corresponding to the rate of demand presentation within those sessions. Therefore, it is likely that John’s problem behavior became more efficient over the course of the analysis, permitting him to escape most task requests with just a few instances of problem behavior, rather than the bursts of behavior observed during the first 3 cycles of the analysis. The moderate levels of problem behavior observed in the play condition also are consistent with this account, based on procedural similarities between demand and play conditions (in each condition, the therapist approaches the participant and prompts activity; thus, escape behavior is sometimes observed in play sessions). Finally, relatively low levels of problem behavior during attention and tangible conditions, coupled with within-session decreasing trends in both of these conditions, indicates that John’s problem behavior did not produce strong reinforcement during these sessions.

The results of the functional analysis did not show correspondence with the results of the anecdotal assessments. Whereas the anecdotal assessment outcomes suggested that PAO and AGP were maintained by social positive reinforcement,
probably in the form of the delivery of tangible/activity items, the functional analysis outcomes indicated that social negative reinforcement in the form of escape from demands maintained John’s problem behavior. A function-based treatment, developed based on the results of the functional analysis, was subsequently evaluated.
CHAPTER 5
TREATMENT ANALYSIS

Participant and Setting

Based on the outcomes of the functional analysis, John participated in a treatment analysis. Sessions were 10 min in length and 1 to 3 sessions were conducted at the same time each day, 1 to 4 days per week. Sessions during the treatment analysis were conducted in the same observation room as during the functional analysis.

Response Definitions and Measurement

The operational definitions and data collection procedures for physical aggression toward others (PAO) and aggression toward property (AGP) were identical to those used during the functional analysis. In addition, event recording was used to score occurrences of an alternative response of “No”. “No” was defined as any audible verbalization of the word “No”. Excluded from the definition was headshakes or verbalizations that did not start with an “n” sound and end with an “o” sound.

Interobserver agreement (IOA) data were collected simultaneously but independently by a second observer during 64% of baseline conditions and during 53% of treatment conditions. IOA was calculated by dividing each session into 1-s bins, summing the number of bins in which the observers agreed on occurrence or nonoccurrence of target behaviors, dividing the results by the total number of seconds in the session and multiplying the result by 100. IOA scores averaged 98% (range = 97.2% - 98.8%) during baseline session and averaged 98.1% (range = 96.6% - 98.9%) during treatment conditions.
Experimental Design

The effects of a series of treatments were evaluated using a reversal design.

Treatment Analysis Procedures

Baseline sessions were identical to the demand condition of the functional analysis. During these sessions, John was asked to clap his hand every 30 s, using graduated prompting at 10 s intervals (verbal, modeling, and physical guidance) and, contingent upon a target behavior he was provided a 30-s escape from demands.

A differential reinforcement of alternative behavior (DRA) procedure was utilized to treat John’s problem behaviors. John was asked to clap his hands every 30 s using graduated prompting. During DRA, PAO or AGP did not produce any programmed consequences (i.e., the presentation of demands was unaltered by problem behavior). An alternative response (“No”) produced a 30 s escape from demands. “No” was chosen as the alternative response because John had previously been observed to emit this response during some baseline sessions.

During the initial DRA condition substantial treatment effects were not evident and, as a result, 3 additional components were sequentially introduced into treatment. First, a “safety signal” was integrated into the DRA procedure. Contingent on the participant saying “No” the therapist would say, “Ok, you don't have to” as she walked away. Second, a “rule,” stated prior to the session and approximately every 4 min during the session was introduced. The therapist always stated the rule the same way, saying “John, I’m going to ask you to clap your hands; if you don't want to just say ‘No’, but if you hit or kick me or the wall you will still have to clap”. Lastly, guided compliance
was used when PAO or AGP occurred following a demand, rather than merely continuing the prompting sequence.

Following a brief reversal to baseline, treatment was reintroduced. Problem behavior again showed decreases and “No” increased; however, these effects were not considered clinically adequate, so another change was made to the procedure. The demands were presented continuously, omitting the interval between trials, unless an appropriate response was emitted. In addition, reinforcement for appropriate responses was increased. If the participant said ‘No’ during task trials a 30-s escape from the task was provided; in addition, if the participant independently completed the task or if the 3-prompt sequence was completed without PAO or AGP, he was provided with a 30 s escape.

The third baseline condition was identical to the first two baseline conditions with the exception that continuous demands were presented. The demands were presented continuously to remain consistent with procedures in the second treatment condition. Following the third baseline condition, the previous treatment condition was reinstated.

Results

Figure 3 displays the total number of problem behaviors (PAO and AGP) and “No” responses within each session. The scale for problem behavior is displayed along left Y-axis and the scale for “No” displayed along the right Y-axis. During the first 4 sessions of baseline, “No” responses were not recorded, so there are no data for those sessions. During the initial baseline the mean number of problem responses/session was 66.9 (range = 19-196), with “No” occurring at a mean of 4.5 responses/session (range = 0-13). The introduction of the first treatment condition did not result in
substantial treatment effects. The mean number of problem behaviors per session was 58.3 (range = 16-154) and “No” occurred at a mean of 20.4 responses/session (range = 11-36). Incorporating the safety signal and rule statement into the DRA procedures produced minimal, if any, beneficial effects. The mean number of problem behaviors per session was 59.4 (range = 14-192) and “No” occurred at a mean of 13.8 responses/session (range = 7-20). However, following the implementation of guided compliance a substantial decrease in problem behavior and an increase in occurrences of “No” was observed. The mean number of problem behaviors per session was 20.2 (range = 0-104) and “No” occurred at a mean of 16.9 responses/session (range = 5-35).

Following the treatment condition there was a brief return to baseline, which was identical to the initial baseline and produced similar outcomes. The mean number of problem behaviors per session in the second baseline condition was 54.67 (range = 20-122) and “No” occurred at a mean of 8.3 responses/session (range = 4-12).

The second baseline was followed by a return to the treatment condition. Problem behavior again showed decreases and “No” increased slightly; however, these effects were not considered clinically adequate. The mean number of problem behaviors per session was 20.8 (range = 4-63) and “No” occurred at a mean of 9.54 responses/session (range = 3-19). After making the demands continuous and increasing the interval following either compliance or the absence of problem behavior throughout the 3-prompt sequence, PAO and AGP decreased to near zero levels for the remainder of the condition. The mean number of problem behaviors per session following implementation of continuous demands and longer escape intervals was 1.6 (range = 1-2) and “No” occurred at a mean of 5 responses/session (range = 1-7).
During the third return to baseline a slight increase in problem behavior and a decrease in “No” responses were observed. The mean number of problem behaviors per session in the third baseline condition was 18.2 (range = 13-29) and “No” occurred at a mean of 5.7 responses/session (range = 1-9).

Finally, during the third treatment condition problem behaviors decreased again, and “No” responses continued to occur, although at lower levels compared to previous treatment conditions. The mean number of problem behaviors per session in the third treatment component was 8.3 (range = 1-14) and “No” occurred at a mean of 4.1 responses/session (range = 3-6).

In Figure 4, PAO and AGP are shown in separate data paths. In general, AGP showed more variability within and between conditions. Significantly more AGP occurred during the first 3 sessions of the baseline condition, with a mean of 141.7 responses/session (range = 77-186), compared with a mean of 4 PAO responses/session (range = 0-10). However, following the third baseline session, AGP decreased sharply to a level approximately equal to that observed for PAO.

During the first segment of the first treatment condition significant variability occurred in measures of AGP while PAO occurred at more stable levels. PAO occurred at a mean of 21.5 responses per session (range = 16-32) while AGP occurred at a mean of 37.2 responses per session (range = 0-129). Inclusion of the safety signal and rule statement produced little if any change across these measures. PAO occurred at a mean of 19.5 responses per session (range = 0-62) while AGP occurred at a mean of 44.7 responses per session (range = 0-130). However, after guided compliance was added as a consequence for problem behavior, measures of both AGP and PAO
decreased, although two “spikes” in AGP were observed. PAO occurred at a mean of 7.6 responses/session (range = 0-17) while AGP occurred at a mean of 13.4 responses/session (range = 0-95).

Following a reinstatement of baseline procedures AGP increased to significantly higher levels with a mean of 46 responses/session (range = 9-113); whereas, PAO remained relatively unaffected with mean of 8.6 responses/session (range = 4-11).

In the second treatment condition PAO initially decreased to near zero levels but then showed an increasing trend that continued throughout the condition. PAO occurred at a mean of 6.2 responses/session (range = 1-14) during this condition. AGP showed some variability during the first 14 sessions of this condition but decreased to near zero levels immediately after the interval between trials was eliminated and the interval following either compliance or absence of problem behavior throughout the 3-prompt sequence was increased. AGP occurred at a mean of 11 responses/session (range = 0-52). In the final return to baseline measures of PAO remained relatively low, with a mean of 12 responses session (range = 9-15) while AGP initially increased and then decreased to near zero levels, with a mean of 6.17 responses/session (range = 0-20). In the final treatment condition both AGP and PAO continued to occur at low levels with a mean of 4.63 responses/session (range = 1-12) for PAO and a mean of 3.36 responses/session (range = 0-12) for AGP.

Figures 5-8 display conditional probabilities of target behaviors occurring within trials and following trials for each condition. Conditional probabilities were calculated to determine the proportion of trials that contained problem behavior (during trials) and the proportion of trials that were immediately followed by problem behavior (within the first 5
s following termination of trials). Conditional probabilities of problem behavior during trials (or within 5 s of trial termination) were calculated by dividing the number of trials containing problem behavior (or that were followed by problem behavior within 5 s) by the total number of trials presented in each condition.

Figure 5 displays the conditional probabilities of PAO occurring within a trial across conditions. During the initial baseline the conditional probability of PAO during a trial was almost .5 (i.e., nearly 50% of trials contained PAO). The first significant decrease in the probability of PAO occurring during trials in treatment was not seen until the rule was stated, at which point a decrease to .302 was observed. There was a further decrease when guided compliance was introduced, and the conditional probability of PAO during trials decreased to .22. During the first return to baseline the conditional probability of PAO during trials increased to a level approximating that observed during the first baseline (.515). The conditional probability of PAO during trials decreased again (to .263) when treatment was reintroduced and even further (to .113) when trials were presented continuously and the interval following either compliance or absence of problem behavior throughout the 3-prompt sequence was increased. The final return to baseline resulted in an increase in the probability of PAO during trials, although the probability was not as great as during the two previous baselines (probability of PAO during trials = .387). The final treatment produced a decrease in the probability of PAO during trials to .201.

Figure 6 displays the conditional probabilities of AGP occurring within a trial across conditions. Overall, AGP was less likely to occur within trials than PAO. During the initial baseline the conditional probability of AGP during trials was .179. This
decreased to .09 when treatment was introduced. There was not another decrease until guided compliance was introduced, after which the conditional probability of AGP occurring within a trial decreased to .013. Similar conditional probabilities were observed in the following baseline and treatment conditions.

Figure 7 displays the conditional probabilities of PAO occurring within 5 s of a trial ending across conditions. During the initial baseline, the conditional probability of PAO followed within 5 s of a trial ending was .179. This decreased to .064 after treatment was introduced and decreased further to .02 after the safety signal was introduced. However, an increase occurred following inclusion of the rule statement and after initiating guided compliance. During the return to baseline there was an increase to .172, which was very similar to the original baseline conditional probability (.179). When treatment was reinstated the probability again dropped to .075 and then to 0 when the interval between trials was omitted. The return to baseline showed a slight increase to .067 while returning to treatment showed another slight decrease to .044.

Figure 8 displays the conditional probabilities of AGP occurring within 5 s of a trial ending across conditions. During the initial baseline the conditional probability of AGP within 5 s of a trial ending by was .235. The probability decreased after treatment was introduced to .106 and again after the safety signal was added, to .08. However, after inclusion of the rule statement an increase to .182 occurred. After guided compliance was introduced the probability decreased again to .079. During the return to baseline condition the probability increased to .13 and the subsequently decreased to .058 when treatment was reinstated. When the interval between trials was omitted the
conditional probability of AGP following a trial decreased to zero. There was a slight 
increase to 0.37 during the third baseline component and an even greater increase to 
0.071 in the last treatment component.
CHAPTER 6
DISCUSSION

The use of functional assessment to determine the function of aberrant behavior is an essential component for the development of effective treatment. Anecdotal assessments are an attractive option for many clinicians due to the ease and speed with which they can be administered. However, based on previous research the reliability and validity of these assessments to accurately assess aberrant behavior remains undetermined.

The purpose of this study was to replicate Fahrenholtz (2004) with an individual participant to determine if functional analysis outcomes would correspond with results obtained through the use of a modified procedure for analyzing anecdotal assessment outcomes. Using procedures described by Fahrenholz et al., agreement across respondents and assessments tools was assessed. The results of the Fahrenholz et al. study were extended by conducting a treatment analysis based on the hypothesized operant function of problem behavior, as identified via the anecdotal assessments and experimental analysis.

In the current study the modified analysis of Motivation Assessment Scale (MAS) outcomes indicated that the maintaining consequence for physical aggression toward others (PAO) and aggression toward property (AGP) was access to tangible items and the Functional Assessment Screening Tool (FAST) outcomes indicated that these behaviors were maintained by social positive reinforcement, thus indicating agreement across assessments. That is, at least 4 out of 5 respondents to the MAS and FAST agreed that the hypothesized function for both PAO and AGP was a form of positive
reinforcement. These outcomes replicated those produced by Fahrenholtz (2004), showing that it was possible to obtain agreement across multiple respondents and assessments regarding the function of aberrant behavior. However, the results of this study differed from those reported by Fahrenholtz (2004) in that when a functional analysis was conducted the results of that analysis did not correspond with the results of the anecdotal assessments.

There are several possible explanations as to why the results of the analogue functional analysis did not correspond with the results of the MAS and FAST. One explanation might be that the variables maintaining the challenging behavior in the analog analysis are not the same as those that maintain it in the natural environment (Yarbrough & Carr, 2000). For example, in the current case, the primary maintaining consequence for the participant’s problem behavior in residential or training contexts may have been access to tangibles; however, in the analog conditions tested in a clinical setting, the problem behavior was sensitive to escape from demands as reinforcement. It may be that the conditions in the observation room acted as an abolishing operation and decreased the momentary value of the tangibles that would otherwise serve as reinforcers in the natural environment.

Another possible explanation for the different outcomes across assessments could be the influence of the social context immediately preceding the analogue assessment (O’Reilly, Lancioni, & Emerson, 1999). Staff frequently reported that the participant engaged in many challenging behaviors prior to, during, and immediately following the transition from home to work. The assessments were conducted in the same building as the participant’s workshop, and he was always seen during his
scheduled vocational/work time, typically 5 – 10 min after he arrived at work. Alternatively, it is possible that being at the workshop functioned as an establishing operation, increasing the reinforcing properties of escape within the demand condition of the analogue functional analysis (Michael, 1993).

It is also possible that, although agreement was obtained with the anecdotal assessments, those assessment outcomes did not accurately identify the contingency of reinforcement maintaining the participant’s problem behaviors. Previous research outcomes supporting the validity of experimental analyses (e.g., Iwata, Pace, Dorsey, 1994) combined with outcomes indicating that anecdotal assessments often produce unreliable or invalid results (e.g., Zarcone et al., 1991; Conroy et al., 1996; Sigafoos et al., 1994) suggest that this may have been the case in the current study.

Following the functional assessments, a function-based treatment was developed to evaluate the effects on the targeted behaviors. It was decided that the treatment would be based on the results of the functional analysis because (a) it has been shown to produce the reliable and valid data regarding the function of challenging behaviors (Iwata, 1994) and (b) very little problem behavior was observed during attention and tangible conditions of the experimental analysis. Thus, no appropriate baseline against which to compare treatments corresponding to these operant functions was available.

The functional analysis indicated that the hypothesized function of the aberrant behavior was escape from demands. Based on these results a treatment was developed that utilized escape extinction and differential reinforcement. Initially, the treatment was not successful, so additional modifications were systematically implemented. The use of a safety signal and rule statement did not produce significant
and consistent effects; however after guided compliance was added measures of PAO and AGP substantially decreased. A final treatment component, omitting the interval between demands and increasing escape intervals for the absence of aberrant behavior, resulted in immediate decreases in both PAO and AGP. Furthermore, low measures of these behaviors were observed throughout the remainder of the study.

There are several possible explanations as to why it was not possible to totally eliminate the participant’s problem behaviors. First, AGP typically occurred in episodes that consisted of several occurrences of kicking or hitting the wall. Thus, a single episode in a session resulted in recording many discrete responses. In addition, it is very likely that the aberrant behavior was reinforced in natural environments (e.g., the workshop and at the participant’s residence) thus, intermittent reinforcement may have increased the resistance of the participant’s behavior to extinction. Finally, aberrant behavior sometimes occurred during the escape interval while in the treatment component. Problem behaviors that occurred during this interval may have maintained due to adventitious reinforcement in the form of task avoidance, because task demands often did not immediately follow problem behavior.

Analysis of the conditional probabilities further suggests that problem behavior was, in fact, maintained by escape from task demands. These outcomes indicated that problem behavior was much more likely to occur during, rather than after, the delivery of task demands. Thus, the onset of task demands appeared to set the occasion for problem behavior, while offset of demands set the occasion for cessation of problem behavior, as would be predicted if problem behavior was maintained by escape.
Some limitations of the study should be discussed. First, the study included only one participant. Including more participants would provide additional information regarding utility and validity of anecdotal assessments. Secondly, escape extinction was utilized to address the hypothesized function based on the functional analysis while a treatment based on the hypothesized function of the anecdotal assessments was not assessed. However, as stated previously, it was not possible to produce an adequate baseline condition under which to test the effects of treatments corresponding to the anecdotal assessment outcomes. The failure of the attention and tangible conditions to produce substantial rates of problem behavior provides strong evidence that these contingencies were not involved in the production and maintenance of this participant’s problem behaviors. An attempt to establish attention or tangible baselines in natural environments may have clarified whether or not the participant’s problem behavior might have been maintained by multiple contingencies of reinforcement, and future research might combine controlled procedures in clinical settings with naturalistic observations when anecdotal and experimental analyses produce inconsistent outcomes.

It is possible that the guided compliance component of the escape extinction condition may have had punishing effects and thus would have decreased the participant’s aberrant behavior regardless of the type of reinforcement contingency responsible for its maintenance. However, as noted above, a preponderance of evidence, including elevated measures of problem behavior in conditions during which task demands were presented, low measures of problem behavior during tests for positive reinforcement functions, and conditional probabilities showing a higher
likelihood of problem behavior during, rather than immediately following, task demands, provides support for the conclusion that the participant’s problem behaviors were maintained via negative reinforcement.

Some extraneous events not under control of the experimenter may have affected the outcomes. First, there were two substantial breaks in the experiment. One was due to an injury to the participant causing a 13-day delay between Sessions 68 and 69 (fractured jaw), and a second was due to administrative issues (a delay in re-obtaining consent following initial expiration) causing a 27-day delay between Sessions 46 and 47. Inspection of data trends before and after these events did not indicate any substantial threats to the current interpretations. In addition, several medication changes were made during the course of the study, including 2 decreases in Ativan® (benzodiazepines) dosages (one occurred between Sessions 27 and 28 and another between Sessions 78 and 79), a decrease in Mellaril® (thioridazine) dosage (occurred between Sessions 43 and 44), and tapering of Zoloft® (sertraline HCl) (beginning following Session 68 with complete discontinuation following Session 76). All the medication changes were reductions and/or discontinuations, which might have been expected to cause increases in problem behavior following the change. The majority of the changes were made during treatment components and there did appear to be a slight increase in problem behavior following the change. However, the behavior eventually decreased to levels that would be predicted as a result of treatment. Medication changes and changes in conditions did not occur simultaneously; thus, there is still strong evidence that the behaviors were controlled by the experimental
manipulations. However, these medication changes may have hampered the ability of the treatment to completely eliminate the aberrant behaviors.

Overall, the findings of this study would suggest that anecdotal assessment tools such as the MAS and FAST should continue to be interpreted with caution and used in combination with more rigorous analyses (Miltenberger, 1997; Zarcone et al., 1991), as the validity of these assessments remains undetermined.

Future research should continue to assess the reliability and validity of anecdotal assessment tools. This is an important area of research due to the limited resources available to many practitioners in applied settings. It is essential for many of these practitioners to find the most efficient and accurate methods for identifying the function of challenging behavior. Currently, however, it is unclear that treatments based solely on anecdotal evidence will ultimately produce satisfactory outcomes.
Figure 1. Total number of occurrences of problem behavior (Physical Aggression to Others and Aggression to Property).
Figure 2. Average number of PAO and AGP that occurred minute-by-minute across the conditions.
Figure 3. Results of problem behavior and “No” from the treatment analysis sessions.

Data are presented as total number of responses/session.
Figure 4. PAO and AGP displayed separately. Data are presented as total number of responses/session.
Figure 5. Conditional probabilities of PAO occurring within a trial across conditions.
Figure 6. Conditional probabilities of AGP occurring within a trial across conditions.
Figure 7. Conditional probabilities of PAO occurring within 5 s of a trial ending.
Figure 8. Conditional probabilities of AGP occurring within 5 s of a trial ending.
REFERENCES


functional analysis outcomes. Unpublished manuscript, University of North Texas.


Unpublished manuscript, University of Florida.


