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

Roxanne Wolf, B.A.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

May 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
Wolf, Roxanne. *Multiple-respondent anecdotal assessments for behavior disorders; An analysis of interrater agreement and correspondence with treatment outcomes*. Master of Science (Behavior Analysis), May 2007, 40 pp., 1 table, 6 illustrations, references, 36 titles.

The current study was designed to further evaluate the usefulness of anecdotal assessments. The goal of this study was to evaluate the overall agreement between multiple respondents on the primary function of aberrant behavior using the Motivation Assessment Scale (MAS) and the Functional Analysis Screening Tool (FAST) and, if agreement was obtained, to assess the effectiveness of treatment based on the outcome of the assessments. Results showed that anecdotal assessments were able to identify the general type of contingency maintaining two participants’ problem behavior. However, for one participant the assessments did not correctly identify the specific form of reinforcement (attention or tangible items) that maintained the aberrant behavior.
Copyright 2007

by

Roxanne Wolf
ACKNOWLEDGEMENTS

First and foremost, I would like to sincerely thank Dr. Richard Smith for your guidance and supervision. Over the years you have been my mentor, my friend, and my extended family. You have taught me the skills to make a difference in the lives of others, and for that I am truly grateful. I would also like to thank Dr. Janet Ellis and Dr. Shahla Rosales for their feedback on this project and overall support. I would like to send a “shout out” to Katy Atcheson and Victoria White for being there through the good, the bad, and the ugly! No matter what the circumstance you always stood beside me and I don’t think I could have made it with out either of you (thanks ladies)! In addition, I would like to thank Carla Smith (the “BARC mom”) for always being the first person to volunteer your time, your friendship, and a helping hand (which includes some pretty amazing baked goods). I would also like to thank Jessica Hobbs, Heather Moore, and all the BARC members for their dedication to this project. Finally, I would like to express my appreciation to the faculty in the Department of Behavior Analysis at the University of North Texas for their ongoing support and dedication to our science.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. GENERAL METHOD</td>
<td>7</td>
</tr>
<tr>
<td>3. ANECDOTAL ASSESSMENT</td>
<td>9</td>
</tr>
<tr>
<td>4. TREATMENT ANALYSIS</td>
<td>15</td>
</tr>
<tr>
<td>5. GENERAL DISCUSSION</td>
<td>26</td>
</tr>
</tbody>
</table>

| ACKNOWLEDGEMENTS | iii |
| LIST OF TABLES AND FIGURES | v |
| REFERENCE LIST | 38 |
LIST OF TABLES AND FIGURES

Page

Tables

1. Resident-by-resident results, with primary categories listed across respondents
.................................................................................................................................................. 32

Figures

1. Results of SIB in treatment analysis sessions for Beth................................. 32
2. Results of PAO and AGP in treatment analysis sessions for Beth ............ 33
3. Results of magazine engagement for Beth.................................................. 34
4. Results of target behavior during treatment analysis sessions for .......... 35
5. Results of PAO from treatment analysis sessions for Rachel ................... 36
6. Results of SIB from treatment analysis sessions for Rachel ..................... 37
CHAPTER 1
INTRODUCTION

The functional-analytic approach to the assessment and treatment of aberrant behavior has been evolving for over two decades. In those two decades, substantial progress has been made, with an emphasis on treatments derived from knowledge of the environmental functions of aberrant behavior. Interventions based on the operant function of aberrant behavior have shown to be effective to treat a variety of aberrant behaviors (Hagopian et al., 1997; Iwata, 1994; Neef & Iwata, 1994). Functional assessment has been defined as the process of gathering information about the antecedent and consequences that are functionally related to the occurrence of a problem behavior (Miltenberger, 2001). There are three general methods of functional assessment. Experimental analysis involves systematic manipulation of antecedents and consequences to observe their effects on the target behavior (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982). Descriptive assessment involves having an observer record antecedent, behavior and consequent events as they occur in the settings in which the behavior is typically observed (e.g., at school, home, or work; Anderson & Long, 2002). Anecdotal assessment involves gathering information about aberrant behavior gathered through interviews, rating scales, and questionnaires (Durand & Crimmins, 1988). Each of these functional assessment approaches is associated with individual strengths and limitations.

Experimental analysis has received the most empirical validation and extension of functional assessment procedures. By systematically manipulating the antecedents and consequences that evoke and maintain problem behaviors and eliminating
extraneous variables that may occur in the individual’s natural setting, experimental analysis offers a method to directly observe the effects of specific environmental variables on problem behavior (Smith, Vollmer, & St. Peter, in press). In their landmark study, Iwata, Dorsey, Slifer, Bauman, and Richman (1982) arranged a series of experimental conditions to determine whether attention from adults, escape from demands, or sensory simulation maintained the self-injurious behavior (SIB) exhibited by their participants. Results showed differentiated patterns of SIB across conditions, demonstrating a cause-and-effect relationship between the controlling variables and the problem behavior. Thus, functional relationships among suspected environmental and behavioral events can be confirmed or disconfirmed by systematically manipulating environmental events and observing their effects on behavior.

There are drawbacks to experimental analysis procedures, including the need for trained therapists and observers, the amount of time and resources needed to complete the assessments, and potential risks to participants and others (Miltenberger, 2001; Smith & Churchill, 2000). Standard functional analysis procedures may also not be appropriate for high intensity or low frequency behaviors (Linscheid et al. 1996; O’Neill at al., 1997; Pyles et al. 1997; Sturmey 1995; Vollmer & Smith 1996). Researchers and clinicians have evaluated variations of the original experimental analysis procedures in response to these putative limitations and made the procedures more practical in outpatient and clinical settings, by conducting brief, time-limited analyses (Derby et al., 1992; Northup et al., 1991; Vollmer, Marcus, Ringdahl, & Roane, 1995). However, none has been widely accepted as a viable alternative to supplant more traditional experimental analysis procedures.
A second method of functional assessment is descriptive assessment. Descriptive assessments involve recording occurrences of naturally occurring antecedent and subsequent events and target behaviors in real time. An advantage of the descriptive assessment is that it suggests how reinforcement operates in the natural environment (i.e., the external validity of outcomes is enhanced by conducting the assessment in the setting in which the problem behavior typically occurs). In addition, the observer records potentially relevant environmental and behavioral events as they occur rather than relying on memory, as in anecdotal assessments (Anderson & Long, 2002; Miltenberger, 2001). However, clinicians and experimenters who conduct descriptive assessments have no control over environmental variables; information gathered during these assessments provides correlational, not causal, information regarding environment-behavior relations. Previous research has also indicated that an additional drawback to descriptive assessment methodology is that, although it may be able to suggest if problem behavior is maintained by social or nonsocial contingencies, it seldom provides useful information about the specific source of reinforcement for problem behavior (Lerman, 1993; Mace & Lalli, 1991; Smith, Vollmer, & St. Peter, in press).

A third method of functional assessment is anecdotal assessment. Anecdotal assessments involve gathering information about environmental and behavioral events indirectly, usually from caregivers or teachers. Some common anecdotal assessments include checklists, interviews, and rating scales. Anecdotal assessments are useful as they provide immediate and general information about characteristics of the individual, the behavior, and the environment (Smith, Vollmer & St. Peter, in press). Anecdotal
assessments are also easy to conduct, do not require a substantial investment of time, and a number of interview formats and anecdotal assessments are available to the public (Miltenberger, 2001). A major limitation of anecdotal assessments is that they rely on opinions and memory and, as a result, may produce unreliable or invalid outcomes (Newton & Sturmey, 1991; Zarcone et al, 1991). Anecdotal assessments also generate hypotheses about the functional properties of target behaviors but do not directly test those hypotheses.

Although the literature has yielded mixed results regarding their utility, anecdotal assessments are widely used by practitioners to identify the operant function of problem behavior (Smith, Vollmer, & St. Peter, in press). For example, the Motivation Assessment Scale (MAS; Durand & Crimmins, 1988) is frequently used in clinical, school, and institutional settings. The MAS was developed to identify the potential influences of contingent attention, escape, and self-stimulation on SIB. Interrater reliability was assessed by administering the MAS to pairs of teachers and teacher aides who worked with 50 children with SIB. Correlational analyses compared the teachers’ and teacher aides’ responses on individual items (correlation range = .66 - .92), on mean scores for the subscales (correlation range = .80 - .95), and on ranks within each subscale (correlation range = .66 - .81). Results showed that all three correlational analyses were significant at the .001 level, indicating a high level of agreement within rater pairs. After concluding that the MAS produced reliable outcomes, correspondence with experimental analyses was assessed by comparing the teachers’ ratings on 8 children to results from analog (experimental) analyses. Results of the correspondence study indicated that the MAS results corresponded with the
results of the analog analyses. Thus, Durand and Crimmins’ (1988) outcomes suggested that the MAS was both reliable (in the sense that a high level of agreement across rater pairs who evaluated the same participant was observed) and valid (in the sense that the MAS outcomes were corroborated by the experimental analysis outcomes).

Subsequent attempts to replicate Durand and Crimmins’ (1988) procedures have yielded mixed outcomes. For example, Zarcone, Rodgers, Iwata, Rourke, and Dorsey (1991) conducted MAS assessments with adults in an institutional setting and children in a school setting. The reliability of the MAS was assessed using the procedures described by Durand and Crimmins as well as using two item-by-item percentage agreement procedures. Results indicated that no analysis showed high levels of agreement within rater pairs, and only 29% of the paired respondents actually agreed on the variables maintaining the SIB. Several additional evaluations of the MAS have been conducted, and although a handful supported its reliability and/or validity (Bihm, Kinlen, Ness, & Poindexter, 1991; Durand, Crimmins, Caulfield, & Taylor, 1989; Kearney 1994;), most have failed to replicate Durand and Crimmins’ original results (Conroy, Fox, Bucklin, & Good, 1996; Goza & Ricketts, 1993; Newton & Sturmey, 1991; Sigafoos, Kerr, & Roberts, 1994; Thompson & Emerson, 1995; Zarcone, Rodgers, Iwata, Rourke, & Dorsey, 1991).

Fahrenholz and Smith (2004) evaluated the potential utility of administering anecdotal assessments to multiple respondents for each person assessed. First, they administered the MAS and the Functional Analysis Screening Tool (FAST; Iwata & DeLeon, 1996) independently to 5 respondents for each of 28 individuals whose
problem behavior was assessed. Results indicated that 4 out of 5 respondents agreed on the primary function of aberrant behavior for 10 of the 28 residents. Of the 10 individuals for whom high agreement had been observed (and who remained available for further analysis) experimental analyses were conducted. Results indicated that, when the experimental analyses produced differentiated outcomes, those results corresponded with those of the anecdotal assessments. Thus, although a relatively small percentage (36%) of assessments showed high levels of agreement on the maintaining contingency for problem behavior, those cases showed a high level of correspondence with the maintaining contingencies identified via functional analysis. Based on these outcomes, the researchers suggested that, when multiple anecdotal assessments yield high agreement across raters, the outcomes may be useful for the development of effective, function-based interventions.

The current study was designed to further evaluate the usefulness of anecdotal assessments. The goal of this study was to evaluate the overall agreement between multiple respondents on the primary function of aberrant behavior using the MAS and the FAST and, if agreement was observed across 4 of 5 respondents, to assess the effectiveness of treatment corresponding to the operant function identified in those outcomes. Function-based treatment should be effective only if the assessment has correctly identified the maintaining contingencies and should not be effective if the identified maintaining contingency is inaccurate. Thus, the current study sought to extend the literature by evaluating whether the outcomes of multiple anecdotal assessments provided a basis for the development of effective, function-based treatments for problem behavior.
Participants and Setting

Assessments and treatments were conducted at a large, state-operated residential and training facility for adults with developmental disabilities. The anecdotal assessments were administered in quiet areas of the residents’ homes or at vocational sites.

Residents

The participants were 2 adults with developmental disabilities who have histories of SIB and aggression against others. These behaviors were identified by residential unit psychologists as target behaviors for formal behavior intervention. Beth was a Caucasian female, 45 years of age, who functioned in the profound range of mental retardation, and engaged in SIB. Her SIB was defined as using an open hand or closed fist to hit or slap her head, face, neck, or other body part. SIB occurred frequently and was moderately severe (230 times/month and caused redness). Rachel was a Caucasian female, 42 years of age, who functioned in the severe range of mental retardation, and engaged in SIB and physical aggression against others. For Rachel, SIB was defined as using an open hand or closed fist to hit herself anywhere on her body, hitting her head against objects, or biting herself on the hand, arm, or other body part. Rachel’s physical aggression was defined as hitting, kicking, biting, slapping, pushing/shoving, hair pulling, or scratching others. SIB and physical aggression against
others occurred frequently and was severe (i.e., more than 45 times/month, causing several instances of injury to herself and others).

Respondents

Each resident’s problem behavior was rated by 5 staff members who had worked directly with that resident for at least 6 months. All respondents had a minimum high school diploma or equivalent and worked directly with the residents as direct-contact staff, psychologists, or as a Qualified Mental Retardation Professional (QMRP).
CHAPTER 3
ANECDOTAL ASSESSMENT

Materials

The materials included 2 sets of MAS and FAST questionnaires for each resident (1 for the respondent and 1 for the interviewer), a sheet of paper on which the MAS rating scale and what each number represented were printed in large fourteen point font, and ball point pens.

MAS

The MAS was developed by Durand and Crimmins (1988) as an individualized assessment tool to identify contingencies of reinforcement maintaining aberrant behavior. The 16-item questionnaire was organized to provide information about 4 (possible) categories of reinforcing consequences: attention, tangible, escape, and sensory. Four questions were allocated to each category and respondents rated each question using a 7-point (0-6) Likert-type scale. A score of zero indicated that the behavior never occurred under the circumstances identified in the question; a score of 6 indicated that the behavior always occurred under the circumstances identified in the question; and scores between those values indicated intermediate ratings (e.g., a score of 3 indicated that the behavior occurred “sometimes” under the circumstances indicated by the question). A score was obtained for each category by summing the points from the 4 questions from that category. Categories were then rank-ordered according to the total scores, and the category receiving the most points was considered the hypothesized maintaining variable.
**FAST**

The FAST was developed by Iwata and DeLeon (1996) to identify variables that maintain aberrant behavior. Four possible categories of reinforcement were examined: 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 were allocated to each category and each question was answered in a Yes/No format. A point was awarded for every “Yes” response given by the respondent. The points were summed for each category, and the category receiving the most points was considered the hypothesized maintaining variable.

**Administration Procedures**

Graduate and undergraduate behavior analysis students were trained to administer the anecdotal assessments. The MAS and FAST were administered to each respondent individually. Both the interviewer and the respondent had copies of the assessments. Before the assessments were initiated, background information (e.g., name of resident, ID number of respondent, date, etc.) was entered on the forms by the interviewer. Definitions of the problem behaviors, including topographical descriptions, were developed for target behaviors identified in each resident’s behavior intervention program. 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 to the respondent by the interviewer and
respondents were instructed to reply to each question with a number from the Likert-scale. If the respondent indicated difficulty answering a question or asked for clarification, the interviewer repeated the question. No additional information or clarifications were provided. Upon completion of the assessments the interviewer thanked the respondent for participating, left the area, and scored the instruments.

Respondent Agreement Evaluation

Reliability of Scoring

All data from the MAS and FAST assessments were transduced by 2 graduate or undergraduate students trained in the administration and scoring of the instruments. Agreement was assessed on a question-by-question basis and was 100% for both instruments.

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. Because the categories of reinforcement varied slightly across assessments, general guidelines were established to assess agreement. For example, if MAS results indicated either attention or tangible as the hypothesized maintaining variable and FAST results indicated social positive reinforcement as the maintaining variable, agreement was scored. If MAS results indicated sensory reinforcement as the hypothesized maintaining variable and FAST results indicated either automatic positive or automatic negative reinforcement as the hypothesized maintaining variable, agreement was
scored. If MAS results indicated escape as the hypothesized maintaining variable and FAST results indicated social negative reinforcement as the hypothesized maintaining variable, agreement was scored. It was also necessary to accommodate within-respondent ties between categories of the hypothesized maintaining variables. If a respondent’s score on one instrument indicated that two categories of possible maintaining variables tied as the highest ranked category, both categories were compared with the highest ranked category from the other instrument to determine agreement across instruments. For example, if both sensory reinforcement and tangible reinforcement were ranked as the primary category on the MAS and social positive reinforcement was ranked as primary category from the FAST, then agreement was scored.

Agreement across Assessments (across Respondents)

Agreement across assessments was scored if 4 of the 5 respondents for each participant agreed on the hypothesized maintaining variable for both the MAS and FAST. Ties between categories of hypothesized maintaining variables were analyzed differently across assessments (across respondents). If a respondent’s score on one instrument indicated that two categories of possible maintaining variables tied as the highest ranked category, both categories were compared with the highest ranked category across the other respondents’ assessments to determine agreement within assessments (across respondents). For example, if both sensory and tangible were ranked as the primary category on the MAS for respondent 1 agreement was scored if other respondents ranked either sensory or tangible as the primary category. Overall
agreement across assessments (across respondents) was scored only when 4 out of 5 respondents agreed both within and across assessments. For example, if 4 (or 5) respondents scored tangible as the primary category on the MAS and social positive reinforcement as the primary category on the FAST, then agreement was scored. However, if 2 respondents scored tangible as the primary category and 3 respondents scored attention as the primary category on the MAS and all respondents scored social positive as the primary category on the FAST, then agreement was not scored.

Results

Table 1 displays primary categories of hypothesized maintaining variables listed across respondents for both residents. This table uses a scatterplot format to permit more detailed visual analyses of agreement across assessments within and across respondents. Yellow shading indicates within-respondent agreements, bold type indicates agreement across respondents, and green shading indicates agreement across instruments. These results show high levels of agreement across assessment and within respondents occurred for both residents. Only respondent 4’s score failed to produce within-rater, across-assessment correspondence; whereas, this rater’s MAS outcomes indicated that Rachel’s SIB was maintained by escape, the same rater’s FAST outcomes identified social positive reinforcement as the hypothesized maintaining contingency.

The MAS assessment identified attention as the hypothesized maintaining consequence for Beth’s SIB. These outcomes corresponded with those from the FAST, which identified social positive reinforcement in the form of attention or tangible items as
the likely maintaining consequence. The MAS assessment identified the receipt of tangible items as the hypothesized maintaining consequence for both of Rachel’s target behaviors. These outcomes corresponded with those from the FAST assessment, which identified social positive reinforcement in the form of attention or tangible items as the likely maintaining consequence.

In order to evaluate the utility of FAST and MAS outcomes to develop effective, function-based interventions for severe problem behavior, treatment procedures that corresponded to each participant’s anecdotal assessment outcomes were evaluated during a treatment analysis.
CHAPTER 4
TREATMENT ANALYSIS

Participants and Setting

Based on the outcomes of the anecdotal assessments, Beth and Rachel participated in the treatment analysis. Sessions for both residents were 10 min in length. One to three sessions were completed each day, at the same time each day, 4-5 times per week. Experimental sessions for Beth were conducted at a specialized facility for behavioral research and treatment (Behavior Analysis Resource Center [BARC]), located on the campus of her residence. Sessions were conducted in a 3.7-m by 2.4-m room, containing a table and two chairs. A one-way mirror was installed on one wall for unobtrusive observation and data collection. Initially, sessions for Rachel were conducted at the BARC facility; however, sessions were conducted on the patio of her home or vocational area starting with session 5. The patio attached to her home was 20 ft by 35 ft. A wooden fence surrounded the perimeter. A soda machine, a picnic table, a lawn chair, and a rocker glider chair were present on the home patio during each session. The vocational patio area was 50 ft by 75 ft. A chain linked fence surrounded the perimeter. Two tables, eight chairs, a lounge chair, and a bench were present on the vocational patio during each session.

Response Definitions and Measurement

Operational definitions from the participants’ formal behavior intervention programs were used during assessment and treatment. For Beth, SIB was defined as using an open hand or closed fist to hit or slap her head, face, neck, or other body part.
For Rachel, SIB was defined as using an open hand or closed fist to hit herself anywhere on her body, hitting her head against objects, or biting herself on the hand, arm, or other body part. Rachel’s physical aggression was defined as hitting, kicking, biting, slapping, pushing/shoving, hair pulling, or scratching others.

Data were collected on laptop personal computers with data collection software (Instant Data 0.8B ®). For both participants, data were collected on the frequency of target behaviors (SIB and aggression against others).

Interobserver agreement (IOA) data were collected simultaneously but independently during a portion of each of the participant’s sessions. IOA coefficients were calculated by dividing each session into 1-s bins, summing the number of bins in which 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. Beth’s IOA data were collected during 35% of baseline sessions and 53% of treatment sessions, and her IOA scores averaged 98% (range = 96.1% – 100%). Rachel’s IOA data were collected during 38% of baseline sessions and 47% of treatment sessions, and her IOA scores averaged 99.3% (range = 98.7% - 100%).

Experimental Designs

For Beth, the effects of a series of interventions were evaluated using a combination multielement and reversal design; for Rachel, treatment effects were evaluated using a multielement design. During each resident’s multielement analysis the order of conditions was designated by a coin toss. If one condition was selected twice consecutively the other condition was implemented next.
**Treatment Analysis: Beth**

Two different baseline conditions were conducted for Beth. During Condition 2 (Contingent Attention), Beth and the therapist were in the observation room together. Contingent on the target behavior, the therapist delivered 5 s of attention in the form of statements of concern and eye contact. During Condition 6 (Contingent Tangible), Beth and the therapist were seated in the room together. Immediately before each session, Beth was allowed noncontingent access to a magazine for approximately 10 s. After 10 s the therapist removed the item from Beth’s possession. Contingent on SIB, Beth received 10-s access to the magazine.

A series of interventions was implemented with Beth. During Condition 1 (Continuous Attention) Beth received continuous attention from the therapist in the form of casual conversation throughout the session. There were no programmed contingencies for SIB. Condition 3 (Continuous Attention + Time Out) procedures were identical to Condition 1 except that, contingent on SIB, the therapist withdrew attention for 10 s (i.e., a 10-s time out). Condition 4 (Continuous Attention + Magazines + Time Out) was identical to Condition 3 except that Beth had noncontingent access to a preferred tangible item (magazine) throughout the session. During Condition 5 (Noncontingent Magazines), Beth had continuous access to the tangible item while the therapist was present, but the therapist did not interact with Beth. There were no programmed consequences for SIB during Condition 5.

**Treatment Analysis: Rachel**

Prior to Condition 1 (Contingent Tangible), a paired stimulus preference
assessment was completed (Fisher et al., 1992). The preference assessment identified access to magazines as the most highly preferred item. The first four sessions of the multielement analysis were conducted in the BARC. Sessions were subsequently moved to the patio of Rachel’s home or vocational area because it was felt that there was a higher probability that Rachel would cause severe injury to herself or others in a small, enclosed environment vs. a larger, more open environment.

Immediately before each contingent tangible session, noncontingent access to the preferred item was provided for approximately 10 s. Then Rachel and the therapist either sat at the picnic table or walked around the patio (according to Rachel’s preference). The therapist held the preferred item in view at all times. Contingent on SIB or aggression, Rachel received access to the preferred item for approximately 10 s.

During condition 2 Rachel was given noncontingent and continuous access to preferred items throughout the session. The therapist was seated at a table and did not engage or reciprocate interaction with Rachel.

**Results**

*_Beth’s Results*_

Figures 1, 2, and 3 illustrate the results of the treatment analysis sessions for Beth. Figure 1 shows the count/session of SIB across each condition.

The first multielement analysis compared the effects of Condition 1 (Continuous Attention) relative to Condition 2 (Contingent Attention). During this analysis, measures of SIB varied widely across sessions, with a general increasing trend and little differentiation between Condition 1 and Condition 2 until the final 4 sessions, during
which measures in Condition 1 significantly exceeded measures in Condition 2. The mean number of SIB/session during Condition 2 was 15.8 (range = 0 – 51). During Condition 1 the mean number of responses was 55.7 (range = 0 - 134).

The second multielement analysis compared the effects of Condition 3 (Continuous Attention + Time Out) relative to Condition 2 (Contingent Attention). During this analysis, measures of SIB varied widely across sessions, with a generally decreasing trend and little differentiation between Condition 2 and Condition 3. The mean number of SIB/session during Condition 2 was 10.1 (range = 0 - 48). During Condition 3 the mean number of responses was 17.1 (range = 0 - 72).

Following the second multielement analysis, all subsequent conditions were presented in sequence, using a reversal design. This change was implemented due to the possibility that the failure to observe differentiation between conditions was due to problems establishing discriminative control due to rapid alternation between conditions. During the first presentation of Condition 2 (Contingent Attention) in the sequential design, measures of SIB were similar to those observed in the previous baseline conditions. The mean number of SIB/session during Condition 2 was 13.9 (range = 0 – 74).

During the first presentation of Condition 3 (Continuous attention + Time Out) in the sequential designs measures of SIB were consistent with the first presentation of Condition 3, showing an increasing trend over the course of 10 sessions. The mean number of SIB/session during the first presentation of Condition 3 was 20.2 (range = 0 - 61).
During the first presentation of Condition 4 (Continuous Attention + Magazines + Time Out) measures of SIB remained elevated for 6 sessions, then decreased to near zero levels for the remainder of the condition. The mean number of SIB/session during Condition 4 was 5.6 (range = 0 - 16).

During the second presentation of Condition 3 (Continuous Attention + Time Out) measures of SIB increased, replicating results from the first presentation of Condition 3. The mean number of SIB/session during the second presentation of Condition 3 was 57.8 (range = 9 - 107).

During the second presentation of Condition 4 (Continuous Attention + Magazines + Time Out) measures of SIB decreased to near zero levels, replicating the results from the first presentation of Condition 4. The mean number of SIB/session during the second presentation of Condition 4 was 2.4 (range = 0 - 10).

During the first presentation of Condition 5 (Noncontingent Magazines) measures of SIB decreased to near zero levels for 10 consecutive sessions. The mean number of SIB/session during the first presentation of Condition 5 was 1.1 (range = 0 - 4).

During the presentation of Condition 6 (Contingent Magazines) measures of SIB indicating an increasing trend. The mean number of SIB/session during the first presentation of Condition 6 was 25.3 (range 12 - 39).

During the second presentation of Condition 5 (Noncontingent Magazines) measures of SIB decreased to near zero levels, replicating the results from the first presentation of Condition 5. The mean number of SIB/session during the second presentation of Condition 5 was .75 (range 0 - 5).
Figure 2 shows the count/session of aggression against others (PAO) and aggression against property (AGP) across treatment analysis sessions. During the first multielement analysis, measures of PAO and AGP occurred at fairly low rates with little differentiation between Condition 2 and Condition 1. The mean number of PAO/session during Condition 2 was .21 times (range = 0 - 3). The mean number of PAO/session during Condition 1 was 0 (range = 0 – 0). The mean number of AGP/session during Condition 2 was .21 (range = 0-3). The mean number of AGP/session during Condition 1 was 1.5 times (range = 0 - 13).

During the second multielement analysis, measures of PAO and AGP continued to remain at zero or near zero levels with little differentiation between Condition 2 and Condition 3. The mean number of PAO/session during Condition 2 was 0 (range = 0 - 0). The mean number of PAO/session during Condition 3 was .41 (range = 0 – 10). The mean number of AGP/session during Condition 2 was .40 (range = 0 - 7). The mean number of AGP/session during Condition 3 was .41 times (range = 0 - 8).

During the first presentation of Condition 2 (Contingent Attention) in the sequential design, measures of PAO and AGP were variable and increased dramatically after the first 3 sessions. The mean number of PAO/session during Condition 2 was 2.3 (range = 0 – 13). The mean number of AGP/session during Condition 2 was 6.1 (range = 0 – 36).

During the first presentation of Condition 3 (Continuous attention + Time Out) in the sequential design measures of PAO and AGP decreased to zero levels.

During the first presentation of Condition 4 (Continuous Attention + Magazines + Time Out) measures of PAO remained at zero while measures of AGP increased after 2
sessions and decreased to zero over 9 sessions. The mean number of AGP/session during Condition 4 was 3.5 (range = 0 – 23).

During the second presentation of Condition 3 (Continuous Attention + Time Out) in the sequential design measures of PAO and AGP remained at zero.

During the second presentation of Condition 4 (Continuous Attention + Magazines + Time Out) measures of PAO and AGP remained at zero levels.

During the presentation of Condition 5 (Noncontingent Magazines) measures of PAO remained at zero while AGP increased slightly after 3 sessions. The mean number of AGP/session during Condition 5 was 1.2 (range = 0 - 4).

During the first presentation of Condition 6 (Contingent Magazines) measures of PAO were zero and measures of AGP increased slightly during the second session. The mean number of PAO per session during Condition 6 was .25 (range = 0 – 1).

During the second presentation of Condition 5 (Noncontingent Magazines) measures of PAO were zero and measures of AGP increased in the first session and decrease to zero for the remainder of the condition. The mean number of PAO per session during the second presentation of Condition 5 was .12 (range 0 – 1).

Figure 3 shows the duration (in min) of magazine engagement (contacting the magazine with her hands) during the first and second presentation of Condition 4 (Continuous Attention + Magazine + Time Out) and the first and second presentation of Condition 5 (Noncontingent Magazines). During the first presentation of Condition 4 (Continuous attention + Magazine + Time Out) measures of magazine engagement showed an increasing trend across the first 8 sessions and then stabilized for the
remainder of the condition. The mean duration of engagement was 7.76 min (range 1.40 min – 9.57 min).

During the first presentation of Condition 5 (Noncontingent Magazine) measures of magazine engagement were variable across session but showed an increasing trend overall. The mean duration of magazine engagement was 5.4 min (range .37 min – 9.25 min).

During the second presentation of Condition 4 (Continuous Attention + Magazine + Time Out) measures of magazine engagement appeared variable in comparison to the first presentation of Condition 4, and the average duration of magazine engagement was slightly lower. The mean duration of magazine engagement was 3.3 min (range = .07 min – 8.21 min).

During the second presentation of Condition 5 (Noncontingent Magazine) measures of magazine engagement were slightly variable across sessions, but on average rates were higher in comparison to the first presentation of Condition 5. The mean duration of magazine engagement was 7.6 min (range = 4.13 – 9.5 min).

Rachel’s Results

Figures 4, 5, and 6 illustrate the results of treatment analysis sessions for Rachel. Figure 4 shows the count/session of SIB and PAO in each condition.

During the multielement analysis conducted in the BARC, effects of Condition 2 (Noncontingent Tangible) were evaluated relative to Condition 1 (Contingent Tangible). During this analysis, measures of SIB and PAO showed an increasing trend in both
conditions. The mean number of responses during Condition 1 was 3 (range = 2 – 4). The mean number of responses during Condition 2 was 3 (range = 0 – 6).

During the multielement analysis conducted on the patio of her home or vocational site, the effects of Condition 2 (Noncontingent Tangible) were evaluated relative to Condition 1 (Contingent Tangible). During this analysis measures of PAO and SIB were higher in Condition 1 and showed a decreasing trend overall. The mean number of responses during Condition 1 was .83 (range = 0 - 14). The mean number or responses during Condition 2 was .19 (range = 0 – 4).

Figures 5 illustrates the count/session of PAO during Condition 1 (Contingent Tangible) and Condition 2 (Noncontingent Tangible). During the multielement analysis conducted in the BARC, effects of Condition 2 (Noncontingent Tangible) were evaluated relative to Condition 1(Contingent tangible). During this analysis, measures of PAO showed an increasing trend in Condition 2 and a decreasing trend in Condition 1. The mean number of responses during Condition 1 was .50 (range = 0 – 1). The mean number of responses during Condition 2 was .50 (range = 0 – 1).

During the multielement analysis conducted on the patio of her home or vocational area, measures of PAO were higher in Condition 1, but showed a decreasing trend overall. The mean number of responses during Condition 1 was .64 (range = 0 – 12). The mean number of responses during Condition 2 was .09 (range = 0 – 3).

Figure 6 illustrates the count/session of SIB during Condition 1 and Condition 2. During the multielement analysis conducted in the BARC, effects of Condition 2 (Noncontingent Tangible) were evaluated relative to Condition 1 (Contingent tangible). During this analysis, measures of SIB showed an increasing trend in Condition 2 and
Condition 1. The mean number of responses during Condition 2 was 2.5 (range = 0 – 5).
The mean number of responses during Condition 1 was 2.5 (range = 1 – 4).

During the multielement analysis conducted on the patio of her home or vocational area, measures of PAO were higher in Condition 1 and showed a decreasing trend overall. The mean number of responses during Condition 1 was .19 (range = 0 – 2). The mean number of responses during Condition 2 was .09 (range = 0 – 2).
In order to effectively treat problem behaviors it is often necessary to correctly identify the maintaining variable(s). The utility of a functional analysis assessment tool to correctly identify maintaining variables that lead to effective intervention is viewed by some as the most important measurement property if not the only important property (Floyd, Phaneuf, & Wilczynski, 2005; Hayes, Nelson, & Jarrett, 1987; Nelson-Gray, 2003). Conventional assessment methods include anecdotal assessments, descriptive assessments and experimental analyses; each has its advantages and disadvantages. For example, a commonly cited problem with descriptive assessments and experimental analyses is that both require extensive investments of time and effort to implement. Anecdotal assessments have been offered as an alternative, more efficient method to identify maintaining variables for problem behavior (Durand & Crimmins, 1988). However, the results of several studies have yielded inconclusive and sometimes conflicting results about reliability, validity, and overall utility of this type of assessment (Bihm, Kinlen, Ness, & Poindexter, 1991; Conroy, Fox, Bucklin, & Good, 1996; Durand, Crimmins, Caulfield, & Taylor, 1989; Goza & Ricketts, 1993; Kearney 1994; Newton & Sturmey, 1991; Sigafoos, Kerr, & Roberts, 1994; Thompson & Emerson, 1995; Zarcone, Rodgers, Iwata, Rourke, & Dorsey, 1991). Therefore, additional research is necessary to determine if anecdotal assessments can serve as quick and accurate means to assess aberrant behavior.

The purpose of the current study was to evaluate the overall agreement between multiple respondents on the primary function of aberrant behavior using the MAS and
FAST and to assess correspondence between the results of treatment procedures and anecdotal assessments. Correspondence between the anecdotal assessments and treatment outcomes was assessed by selecting residents whose outcomes resulted in a high level of across-rater agreement and comparing them to the treatment outcomes. This study attempted to extend the literature regarding validity of anecdotal assessments and their ability to correctly identify maintaining variables for aberrant behavior by evaluating effectiveness of interventions corresponding to anecdotal assessment outcomes.

Results from the 2 participants in this study indicated that anecdotal assessments were able to identify the general maintaining contingency of reinforcement for both participants (social positive reinforcement). For each participant, treatments based on assessment outcomes eventually produced substantial decreases in problem behavior. However, for Beth, MAS outcomes suggested that the form of social positive reinforcement maintaining her problem behavior was attention rather than access to tangible items. That is, although both FAST and MAS outcomes indicated that her problem behavior was maintained by a contingency of positive reinforcement, the MAS outcomes suggested that attention, per se, was the maintaining consequence. The tangible category received the second highest rating on the MAS with 3 of 5 respondents indicating that social positive reinforcement in the form of access to tangible items maintained problem behavior. The FAST assessment does not identify specific forms of potential maintaining consequences.

Beth’s initial treatment was based on the results of the MAS indicating that attention maintained her SIB. The results showed that attention-centered interventions
(Continuous Attention, Time Out) did not reduce SIB and, in fact, resulted in increases in SIB over baseline levels. Subsequently, introducing a preferred tangible item decreased SIB to near zero levels, and withdrawing the attention components did not result in an increase in SIB. In Condition 6, when SIB resulted in the delivery of the tangible items, there was an immediate increase in SIB. These outcomes, collectively, suggested that SIB was not maintained by attention but rather by access to a tangible item. Analysis of Beth’s aggression against property and aggression against others provided further insight into other potential maintaining variables (Figure 2). For example, during conditions when a tangible item (magazine or phone book) was present in the room (Baseline 1), rates of aggression against others and aggression against property increased, indicating that Beth’s SIB was potentially maintained by access to a tangible item. By contrast, when Beth was given noncontingent access to a magazine, rates of SIB decreased to near-zero levels. Finally, measures of the duration of engagement with the magazine (Figure 3) show that as the duration of magazine engagement increased, rates of SIB decreased. These outcomes were consistent with an account that Beth’s SIB was maintained by access to tangible items.

The results of the anecdotal assessments for Rachel indicated that SIB and aggression against others was maintained by social positive reinforcement in the form of access to a tangible item. The MAS indicated that access to a tangible item was the primary maintaining consequence for her SIB and aggression against others. The FAST also indicated that SIB and aggression against others were maintained by social positive reinforcement although, as previously noted, the FAST does not identify specific forms of potentially reinforcing consequences (e.g., attention, access to tangible items).
items). A treatment involving noncontingent availability of a preferred tangible item was
developed based on these results. Rates of SIB, aggression, and aggression against
others decreased to near zero levels when Rachel was given noncontingent access to
tangible items. By contrast, when tangible items were removed from Rachel's
possession and re-presented contingent on problem behavior she became aggressive
against others and engaged in SIB. These outcomes are consistent with those from the
anecdotal assessments indicating that her problem behavior was maintained by positive
reinforcement in the form of access to tangible items.

The results of this study should be considered in light of several limitations. For
example, the results of the treatment assessments strongly suggest but do not
unequivocally confirm the contingencies of reinforcement that maintained problem
behavior. For each participant, the final, ultimately effective, intervention consisted of
noncontingent availability of a preferred tangible item. Although this intervention
corresponds most directly to contingent availability of tangible items as a maintaining
consequence, access to preferred items also has been shown to be an effective
component of interventions aimed at decreasing behavior maintained by other forms of
reinforcement (e.g., Fisher, DeLeon, Rodriguez-Catter, & Keeney, 2004; Hanley,
Piazza, & Fisher, 1997). Subsequent research might attempt to use interventions that
more specifically target a single maintaining contingency. For example, showing that
the participants in the current study would emit effortful alternative behavior to produce
preferred items would have strengthened the current interpretations.

Another limitation was that there were only 2 participants in this study, both of
whose problem behavior appeared to be maintained by a common contingency (social
positive reinforcement). Additional participants may have provided more information regarding the utility of the anecdotal assessments in the identification of other maintaining contingencies (e.g., escape, automatic reinforcement) for aberrant behavior.

Another potential limitation of the current study was the use of both the MAS and the FAST assessments to indicate the function of the aberrant behavior. Although both assessments are designed to identify the operant function(s) of problem behavior, the outcomes are categorized differentially across assessments. For example, although both assessments evaluate contingencies of social positive reinforcement, the MAS separates social positive reinforcement into two separate categories: attention or tangible. The FAST does not identify specific forms of social reinforcement but rather assesses the general category of social positive reinforcement. The current results suggest that the assessments appear to be useful in identifying the general nature of maintaining contingencies (e.g., social positive reinforcement); however, the FAST does not permit a more fine-grained assessment of specific forms of reinforcement, and the current results cast some doubt about the utility of the MAS to identify those specific forms.

The absence of an assessment of specific forms of reinforcement in the FAST and current outcomes indicating that the MAS does not always correctly identify specific maintaining consequences suggests that additional assessments may be necessary to identify the particular stimuli to be manipulated during treatment of behavior disorders. For example, preference assessment may be a useful adjunct procedure to provide additional information about individual and/or idiosyncratic sensitivities to reinforcement.
(Lalli & Kates, 1998; Lindauer, Zarcone, Richman, & Schroeder, 2002). Subsequent research should investigate the utility of combining anecdotal and preference assessments to identify specific consequences maintaining problem behavior.

The current study extends previous research on anecdotal assessment by evaluating the effects of treatments for behavior disorders based directly on the outcomes of multiple administrations of anecdotal assessments. Although preliminary, these outcomes are promising, and suggest that multiple anecdotal assessments may represent a viable alternative approach to functional assessment. Future research might evaluate the utility of this approach in schools, where functional analyses and other assessments are even harder to conduct than in large facilities like DSS. This approach may also be useful in cases where problem behavior is particularly severe or when it is otherwise necessary to expedite assessment or conduct assessments that do not require the direct production and/or observation of the problem behavior itself.

Future research should continue to examine the potential utility of anecdotal assessments as an efficient alternative approach to functional assessment and treatment development (Hayes et al, 1987; Nelson-Gray, 2003).
Table 1

Resident-by-Resident Results, With the Primary Maintaining Variable Listed Across Respondents.

<table>
<thead>
<tr>
<th>Residents</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>Beth- SIB</td>
<td>Att</td>
<td>S+</td>
<td>Att</td>
<td>S+</td>
<td>Tan/Sen</td>
<td>S+</td>
</tr>
<tr>
<td>Rachel- SIB</td>
<td>Tan</td>
<td>S+</td>
<td>Tan</td>
<td>S+</td>
<td>Esc</td>
<td>S+</td>
</tr>
<tr>
<td>Rachel- PAO</td>
<td>Tan</td>
<td>S+ / S-</td>
<td>Tan</td>
<td>S+</td>
<td>Tan</td>
<td>S+</td>
</tr>
</tbody>
</table>

Note: On the MAS, Sen = sensory; Esc = escape; Att = Attention; Tan = tangible. On the FAST, S+ = social positive; S- = social negative; A+ = automatic positive; A - = automatic negative

Figure 1. Results of SIB from the treatment analysis sessions for Beth. Data are presented as count per session for each condition.
Figure 2. Results of PAO and AGP treatment analysis sessions for Beth. Data are represented as count per session for each condition.
**SESSION BY SESSION DURATION OF MAGAZINE ENGAGEMENT**

Figure 3. Results of magazine engagement from Condition 4 (Continuous Attention + Magazine + Time Out) and Condition 5 (Noncontingent Magazine) sessions for Beth. Data are presented in total duration of magazine engagement for these conditions.
Figure 4. Results of target behavior during treatment analysis sessions for Rachel. Data are presented as count per session for each condition.
Figure 5. Results of PAO from the treatment analysis session for Rachel. Data are presented as count per session for each condition.
Figure 6. Results of SIB for treatment analysis session for Rachel. Data are presented as count per session for each condition.
REFERENCES


