1998

Questions About Behavioral Function (QABF): A Behavioral Checklist for Functional Assessment of Aberrant Behavior.

Theodosia Renata Paclawskyj
Louisiana State University and Agricultural & Mechanical College

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

QUESTIONS ABOUT BEHAVIORAL FUNCTION (QABF): A BEHAVIORAL CHECKLIST FOR FUNCTIONAL ASSESSMENT OF ABERRANT BEHAVIOR

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Psychology

by
Theodosia Paclawskyj
B.A., The Johns Hopkins University, 1990
M.A., Louisiana State University, 1995
December, 1998

Acknowledgments

I would like to express my sincere gratitude to the following people who made it possible for me to achieve this stage in my career:

First, to Dr. Johnny Matson, for guiding me and inspiring me in my career and exemplifying what a mentor should be.

Next, to my dissertation committee, Dr. Katie Cherry, Dr. Drew Gouvier, Dr. David Reitman, and Dr. Geraldine Holmes, for providing me with valuable suggestions and encouragement.

Also, to Yemonja Smalls, Bruce Mortenson, Karen Tucker, and especially Karena Rush, without whose help these data could not have been collected.

Finally, to Beth Johnson, Roxanne Sykes, and Trish Kurtz, for the moral support that makes these kinds of endeavors possible.

Table of Contents

List of Tables

List of Figures

Abstract

Functional assessment is a tool in behavior therapy used to identify the functional relationships between a behavior of interest and the individual's environment. A functional assessment identifies variables that both occasion and maintain a particular behavior. Traditional functional assessment methodologies have relied on experimental techniques in which analog sessions are designed to mimic situations in the individual's environment. However, these techniques are time-consuming, require extensive training, and rely on the availability of numerous resources in the individual's setting. Development of a brief functional assessment checklist would circumvent these difficulties and make a significant contribution to applied behavioral psychology. The proposed study attempted to meet this need by providing the psychometric data for such a checklist: the Questions About Behavioral Function. These data include test-retest, interrater, and internal consistency statistics, factor analysis, and convergent validity. Reliability and factor analysis data were sound; however, convergent validity with analog functional analysis sessions requires additional work.

Introduction

Since Fuller (1949) first demonstrated the success of positive reinforcement in a client with profound mental retardation, behavior modification in persons with mental retardation has become a well-established area of research and practice (Matson & Coe, 1992). In the 1950's, basic laboratory research demonstrated the effectiveness of operant learning principles on the behavior of persons with mental retardation. Such research progressed into the application of behavioral interventions targeting both adaptive and maladaptive behaviors in the 1960's (Madle & Neisworth, 1990). The last two decades have yielded hundreds of research studies in these areas, proving that behavior modification with persons with mental retardation is both a highly efficacious treatment (American Psychological Association, Task Force on Promotion and Dissemination of Psychological Procedures, 1995) and the treatment of choice (Madle & Neisworth, 1990; Matson & Mulick, 1991).

In recent years, behavior modification research has been applied to a wide variety of problems in diverse settings. While such progress has led to an established technology, its application on a state-of-the-art level in the service sector still requires work. Clinicians in the field with large caseloads of clients may perceive research as intrusive and an interruption to immediate service needs. In contrast, basic experimental researchers may view the bureaucracy inherent in agency systems as prohibitive to strict research methodology (Matson & Mulick, 1991). To counter this impasse, applied research takes recent developments in theory and basic scientific knowledge and applies

them to problems of immediate need. Applied research therefore bridges the gap between research and practice (Keys, Fletcher, Holmes, & Schloss, 1989).

Overview of functional assessment

The technique of *functional assessment* represents a unique combination of basic and applied research. Functional assessment is the set of procedures which define the relationship between events in the environment and specific target behaviors (Iwata, Vollmer, & Zarcone, 1990; Linscheid, Iwata, & Foxx, 1996; Neef & Iwata, 1994; Sprague & Horner, 1995; Vollmer & Smith, 1996). Events in the environment may be external (social) or internal (physiological) variables. A functional assessment defines the topography of the behavior, its frequency and duration, the antecedent events that may occasion the behavior, the consequences that may maintain the behavior, and the appropriate treatment intervention (Lennox & Miltenberger, 1989; Sprague & Horner, 1995). This information is essential for planning a behavioral treatment in that it identifies reinforcers to be rescheduled or removed, critical situations to be targeted, and functionally equivalent behaviors that could be included into a training curriculum (Sturmey, 1994).

In short, functional assessment may improve the effectiveness and efficiency of behavioral treatment in most circumstances (Horner, 1994). Failure to use a functional assessment, however, may lead to delay in implementing effective treatment procedures (Lennox & Miltenberger, 1989), counter-therapeutic effects on the target behavior from an arbitrarily selected treatment (Solnick, Rincover, & Peterson, 1977), or unnecessary

exposure to aversive procedures (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982). Use of functional assessment in basic research has identified many of the fundamental processes underlying maladaptive behaviors such as self-injury and aggression. This information from basic research has led to the development of highly successful applied interventions (Vollmer & Smith, 1996).

Functional assessment has been described as the most effective tool for successful individualized behavioral interventions in persons with mental retardation (Horner, 1989; Laties & Mace, 1993; Repp, Felce, & Barton, 1998; Wacker, 1989). Functional assessment can predict which treatments will be effective prior to the actual implementation of a treatment intervention (Iwata et al., 1990). Several authors noted that functional assessment may provide more effective and humane interventions than historically-used aversive consequences and may result in greater generalization and maintenance of treatment effects (Axelrod, 1987; Mace, Lalli, Pinter-Lalli, & Shea, 1991; Malott, Whaley, & Malott, 1993). Current best practice requires a functional analysis of biological and social variables prior to treatment implementation (Arndorfer & Miltenberger, 1993; Association for Behavior Analysis, Task Force on the Right to Effective Behavioral Treatment, 1988; National Institutes of Health, Consensus Development Panel on Destructive Behaviors in Persons with Developmental Disabilities, 1989).

Didden, Duker, and Korzilius (1997) conducted a meta-analytic study on the effectiveness of treatments for problem behaviors in persons with mental retardation.

Examining variables such as functional level of the client, etiology of the behavior, intervention setting, duration of assessment and treatment, secondary handicaps, and pretreatment functional assessment, they determined that use of functional assessment was the only significant variable in the prediction of treatment success (defined as a clinically significant reduction in the target behavior), replicating the findings of Scotti, Evans, Meyer, and Walker (1991). Petersen and Martens (1995) found that the general use of functional assessment in both clinical and research settings has been increasing in the past decade.

In recent years, there have been both technological and conceptual refinements resulting in diverse methods of functional assessment (Sturmey, 1996). Recognizing the need for brief but effective assessment in applied settings, several researchers proposed the use of behavioral checklists as an efficient method by which to complete a functional assessment when resources are limited (Durand & Crimmins, 1988; Van Houten & Rolider, 1991; Weiseler, Hanson, Chamberlain, & Thompson, 1985). Such technology could significantly reduce the amount of time required to evaluate a client, which is a critical need given recent trends towards reduced service hours through managed care. Iwata et al. (1994) estimated that functional analyses currently require 1-2 weeks to complete.

Although the possibility of these checklists is promising and their use may be effective in some circumstances, their overall reliability and validity has been inadequate for wide-scale use to date (Sturmey, 1994). Matson et al. (1996) provided the most recent

attempt to construct a reliable and valid checklist with the development of the Questions About Behavioral Function (QABF). In this initial study, good internal consistency and split-half reliability data were reported as well as a factor structure consisting of five factors which corresponded to the subscales of the QABF.

The QABF is an instrument intended to decrease time, training, and resource needs far below those required for traditional functional analysis techniques. If the QABF can be used to identify hypotheses for problem behaviors most of the time in most situations, then it will be a useful tool developed from applied research for wide-scale clinical use.

Outline of the present study

The present study expands upon the initial psychometric data for the QABF by providing interrater reliability, test-retest reliability, construct validity, and concurrent validity data for the instrument. The literature review which follows outlines the historical development of functional assessment and the various methodologies of its use. Next, a rationale for the present study is provided in the context of previous methods to develop abbreviated assessment methods and the current best practice guidelines for conducting a functional assessment. Finally, reliability and validity data are summarized and discussed in the context of current research.

Literature Review

Terminology

The term functional assessment refers to the set of nonexperimental procedures used to identify possible functions of specific behaviors (Arndorfer & Miltenberger, 1993; Sturmey, 1994; Vollmer & Smith, 1996). In contrast, functional analysis refers to systematic procedures which can be used to experimentally manipulate environmental conditions in a controlled setting (Horner, 1994; Lennox & Miltenberger, 1989). This distinction will be observed throughout the following literature review.

Historical trends in functional assessment

Theory

Radical behaviorism, the scientific study of behavior outlined by B. F. Skinner (1953), differs from other psychological orientations in that it is a descriptive and observational science based on principles of inductive reasoning. Its specific focus is on behavior in its context; that is, the relationship between a behavior of interest, the conditions which precipitate the behavior, and the consequences which follow the behavior (Chiesa, 1992). As stated by Skinner (1953), the purpose of this science is prediction and control.

> "We undertake to predict and control the behavior of the individual organism. This is our 'dependent variable'-the effect for which we are to find the cause. Our 'independent variables'-the causes of behavior-are the external conditions of which behavior is a function. Relations between the two-the 'cause-and-effect relationships' in behavior-are the laws of science." (Skinner, 1953, p.35)

Skinner developed a concept of causation based on the work of Ernst Mach who substituted the expression "functional relation" for "cause and effect" (Chiesa, 1992). Rather than relying on the concept of "force" or an outside agency, Skinner (1953) defined a functional relation as a change in an independent variable effecting a change in a dependent variable. When empirical data on the variables are obtained, a functional relationship is inductively derived. When the behavior is successfully predicted and controlled, the functional relationship is confirmed. This process of identifying functional relations was labeled "functional analysis" (Delprato & Midgley, 1992).

Functional analysis yields what has been termed the "three-term contingency". That is, it specifies (1) the occasion upon which a response occurs, (2) the response itself, and (3) the reinforcing consequences (Skinner, 1969). Any of these stimuli may originate from objects in the environment, from the biological make-up of the organism, or from the behavior of others in the environment. However, while several consequences may follow a response, not all may be functionally related to the behavior (Bijou, Peterson, & Ault, 1968).

Baer, Wolf, and Risley (1968) concluded that the fundamental objective of behavioral research is to identify the relationship between a behavior of interest and the environment in which it occurs. Analysis of a behavior is achieved when the experimenter can demonstrate control over the behavior. Kanfer and Phillips (1970) used the acronym SORKC to describe functional analytic assessment in behavior therapy. Within this model, "S" represented the stimulus or antecedent to the behavior, "O" referred to

organismic variables, "R" referred to the target behavior, "K" referred to the contingency relations between the behavior and consequences (e.g., reinforcement schedules, durations, etc.), and "C" represented the consequences of the behavior, both social and nonsocial.

Application to problem behaviors

Radical behaviorism, as developed by Skinner, had its greatest impact on the habilitation of persons with mental retardation, particularly in the remediation of problem behaviors (Matson & Coe, 1992). Lowry and Sovner (1991) defined *problem behavior* as any pattern of action that is an inappropriate response to stimuli and results in a harmful effect to the individual or environment. Persons with severe to profound mental retardation and living in institutions display the highest prevalence of problem behaviors (Jacobson, 1982).

Problem behaviors result in five classes of consequences: physical, social, emotional, educational, and economic. Physical risks vary by the topography of the behavior but can include tissue and organ damage, infection, poisoning, pain, and disfigurement. Social effects include isolation from family and peers and more globally, the development of negative attitudes towards persons with mental retardation. This isolation may lead to co-morbid personality problems such as affective, explosive, or anxious behaviors. Caregivers managing these behaviors may have limited time to devote to educational programming, or may resort to placing the individual in a more restrictive educational setting. Finally, the additional staff needed to work with persons with problem behaviors

greatly increase the cost to society. A 1988 estimate suggested a cost of care exceeding $3 billion for persons with mental retardation and problem behaviors (National Institutes of Health, Consensus Development Panel on Destructive Behaviors in Persons with Developmental Disabilities).

Self-injury.

Carr (1977) described self-injury (SIB) as "the most dramatic and extreme form of chronic human psychopathology" (p. 800). Self-injury is more common in individuals with severe to profound mental retardation and may consist of head banging, self-hitting, self-scratching, self-biting, ingestion of inedible items, eye gouging, hair pulling, or other responses produced by the individual that result in injury to self (National Institutes of Health, Consensus Development Panel on Destructive Behaviors in Persons with Developmental Disabilities, 1989; Schroeder, 1991). Self-injury occurs in an average of 17% of persons with mental retardation, depending on the setting and population assessed (Schroeder, Rojahn, & Oldenquist, 1989). Self-injury frequently results in serious tissue damage, exclusion from social and educational activities, and may lead to a greater risk of vision and hearing impairments over the course of an individual's lifespan (Weiseler, Hanson, & Nord, 1995). Treatment of self-injury requires prompt and effective intervention (Johnson & Baumeister, 1978).

Aggression and property destruction.

Less serious for the individual but more dangerous to caregivers are destructive behaviors such as aggression and property destruction. These may include spitting,

kicking, biting, hair pulling, punching, throwing objects, tearing books or clothes, or breaking furniture (Mulick, Hammer, & Dura, 1991; National Institutes of Health, Consensus Development Panel on Destructive Behaviors in Persons with Developmental Disabilities, 1989). Schroeder et al. (1989) estimated the prevalence of aggressive behavior at 22% and property destruction to be approximately 15% among persons with mental retardation. Destructive behaviors may disrupt family functioning, result in school placement failure, and generally remove a person from the mainstream of society (Carr, Newsom, & Binkoff, 1980).

Assessment and treatment: Overview.

The potential severity of destructive behaviors necessitates a thorough yet efficient assessment to minimize risk to the client or others from delay in appropriate treatment. Skinner (1953) identified three conditions under which problem behaviors could develop. First, an individual might expose him/herself to aversive stimulation if by doing so he/she avoided a more aversive situation (negative reinforcement hypothesis). Next, an aversive stimulus might be paired with a positive reinforcer (positive reinforcement hypothesis). Finally, Skinner used the term "automatic reinforcement" to denote conditions of directly produced (e.g., physiological consequences such as the production of endogenous opioids) reinforcers.

In the past 30 years, researchers have identified an extensive set of variables in home, school, work, and community settings that affect problem behaviors (Sprague & Horner, 1995). However, interventions traditionally focused on punishing problem behaviors and

reinforcing desirable behaviors without examining the reinforcement history which had established the behavior pattern (Arndorfer & Miltenberger, 1993; Horner, 1989; Iwata et al. 1982/1994; Mace, 1994; Matson & Coe, 1992; Sprague & Horner, 1995). Treatments were designed to compete with whatever reinforcement conditions maintained the behavior. In many cases, this led to the development of default procedures for specific behaviors that relied on artificial reinforcement and aversive stimulation (e.g., differential reinforcement and time-out for aggression) (Mace, 1994). While some efforts proved successful, there were treatment failures (Carr, 1994). Lowry and Sovner (1991) labeled this approach as "management" rather than "treatment" of behavior problems.

Early studies in functional analysis

Lovaas, Freitag, Gold, and Kassorla (1965) conducted what may be considered the first functional analysis of environmental events which maintained SIB. They studied a nine-year old female with SIB diagnosed as schizophrenic for whom social approval was demonstrated to be a positively reinforcing event. In comparing social attention consequent to SIB, ignoring consequent to SIB, and a condition in which discriminative stimuli for behaviors incompatible with SIB were present, the authors found that the self-injury consistently increased in the attention condition, remained stable in the ignore condition, and consistently decreased in the incompatible behaviors condition.

This study was the first of several to demonstrate that self-injurious behavior served an operant function and that functional relationships could be examined (Bachman, 1972).

However, subsequent work tended to focus on the reinforcement of incompatible behaviors rather than function-based treatments to reduce self-injury.

Peterson and Peterson (1968) utilized differential reinforcement of other behavior (behavior not necessarily incompatible with self-injury) to reduce head-banging and head-slapping in an eight-year old institutionalized male. Similarly, Corte, Wolf, and Locke (1971) employed differential reinforcement of other behavior using edibles and mild food deprivation to decrease the self-injurious responses of two adolescents with profound mental retardation. However, other studies reported poor results in the use of differential reinforcement to treat self-injury (Measel & Alfieri, 1976; Young & Wincze, 1974).

At this time, self-injury as an operant response was considered to be a function of social conditions such as attention, affection, and abuse (Bachman, 1972). Therefore, the other vein of research focused on extinction as well as time-out from social reinforcement as a punishing consequence for self-injury. Extinction was effective in some instances (Jones, Simmons, & Frankel, 1974; Lovaas & Simmons, 1969) but not in others (Corte et al., 1971; Myers, 1975). Conflicting findings also were reported with the use of time-out. Solnick, Rincover, and Peterson (1977) reported the initial failure of time-out in reducing self-injury, tantrums, and spitting in two children with mental retardation. However, Hamilton, Stephens, and Allen (1967) utilized a chair time-out procedure (30 to 120 minutes in duration) following each occurrence of self-injury or aggression to successfully reduce the behavior in institutionalized individuals with severe mental retardation.

Tate and Baroff (1966) employed time-out from physical contact with staff to suppress the self-injury of a nine-year old male with autism. However, the high risk of injury for this participant led the authors to change to a punishment contingency of a 130-volt electric shock delivered by a cattle prod following each instance of self-injury. While implementing the punishment contingency, the authors added a social reinforcement schedule for all non-self-injurious responses. The combination of interventions decreased the self-injury by 97%. Using a similar intervention of electric shock consequent to high-risk self-injurious behavior and social reinforcement of appropriate behavior, Bucher and Lovaas (1968) eliminated the self-injury of two children with mental retardation. Lovaas and Simmons (1969) replicated this treatment package with three adolescents with severe to profound mental retardation and self-injury. Yeakel, Salisbury, Greer, and Marcus (1970) and Corte et al. (1971) also reported on the successful use of electric shock to suppress self-injury in their participants. Aversive stimulation appeared to be the only intervention consistently effective in reducing self-injury (Iwata et al., 1982/1994). However, professional and regulatory standards began requiring the use of least intrusive treatments and aversive techniques only when other interventions failed (May et al., 1975).

While most of the early research in problem behaviors tended to focus on consequences, some antecedent studies were conducted. Berkson and Mason (1964) studied the rate of stereotypic behavior in two conditions: a barren environment and an environment with toys available. They determined that the presence of toys decreased the

rates of stereotypy in a group of adults with mental retardation. Moseley, Faust, and Reardon (1970) expanded upon this study and developed analog sessions to investigate the effect of three environmental conditions on the rate of stereotypic behavior in children with mental retardation. These authors collected data in baseline (child alone in crib), toys only (4 toys available to child in crib), and social (experimenter interaction and toys offered to child) conditions. While their tabulated data showed considerable intrasubject variability, the authors chose to examine intersubject rates rather than deriving functional relationships for each subject and found a slight effect of toys and the experimenter on decreased rates of the behavior.

Formal functional analyses

Examining the trends of the previous decade, Carr (1977) observed that research on self-injury tended to focus on treatment, which, in many cases, was not successful. He determined that this lack of success may be due to multiple motivating variables for the behavior across participants, and that successful treatment outcome would depend on accurate identification of the motivating factor on an individual basis. Through a comprehensive review of existing research, Carr identified positive reinforcement in the form of social contact, negative reinforcement in the form of termination of an aversive stimulus, and positive internal reinforcement in the form of self-stimulation as viable hypotheses. Carr cited the need for single-subject research to shift the focus from group analysis (clouded by intersubject variability) to single case assessment of antecedent and consequent events affecting individual subjects. To promote such assessment, he

suggested a screening procedure to identify potential environmental variables affecting self-injury. First, a medical evaluation would assess for genetic or other disorders associated with self-injury (e.g., Lesch-Nyhan syndrome, otitis media). Next, observation of the behavior under conditions of attending to the behavior, withdrawing reinforcers, and being in the presence of adults would assess for potential social variables. Observation when demands or other aversive stimuli were presented would assess for possible escape variables. Finally, observation of the behavior under barren conditions would determine if self-stimulation was a motivator.

Carr et al. (1980) provided experimental validation of the effectiveness of such observations when they demonstrated escape from demands to be a reinforcer for the aggressive behavior of two children with severe to profound mental retardation. They validated the escape hypothesis first through a reversal design comparing rates of behavior in demand and no-demand conditions and then through treatment by reinforcement of compliance and escape extinction.

Also at this time, Hayes, Rincover, and Solnick (1980), in a review of publication trends in applied behavior analysis, noted that the field had reached a tendency towards simple replication of treatment effects rather than investigation of underlying conceptual issues. They labeled this a "structural" means of practicing applied behavior analysis in that it focused on descriptions of subjects, settings, and topographies. Hayes et al. (1980) and Birnbrauer (1979) argued that such a focus on structuralism was contrary to the functional nature of behaviorism (e.g., Skinner, 1953; Baer et al., 1968). Morris (1991)

described this tendency to rely on demonstrations of interventions rather than identification of functional relationships as the field becoming "technological to a fault" (p. 413).

However, in 1982, Iwata and colleagues developed the first comprehensive experimental methodology to evaluate the hypotheses proposed by Carr (1977). Noting the need for less intrusive interventions and the inconsistent success of arbitrarily determined treatments, these authors developed a series of analog sessions to experimentally determine the maintaining variables of self-injury so that treatment could match the function of the behavior. They observed the self-injurious behavior of nine participants with mental retardation in four conditions structured in a multi-element design: social disapproval, academic demand, unstructured play, and alone. In the social disapproval condition, the experimenter entered a therapy room with the subject and instructed the subject to play with toys while the experimenter "did some work". The experimenter provided attention only following an occurrence of self-injury. High rates in this condition would indicate that attention reinforced the behavior. In the academic demand condition, the experimenter presented academic tasks to the subject using a graduated 3-step prompting procedure. Social praise was delivered upon compliance and a 30-second break was given following each instance of self-injury. High rates in this condition would indicate negative reinforcement via escape maintained the behavior. The unstructured play condition served as a control in which no demands were presented, preferred toys were available, praise was given at least once per 30-seconds, and self-

injury was ignored. Little self-injury was expected in this condition. Finally, in the alone condition, the subject was observed alone in a room through a one-way mirror. High rates in this condition would indicate automatic, or internal reinforcement. Six of their nine participants displayed high rates in one condition, while the remaining three subjects showed an undifferentiated pattern of responding. The authors concluded that the occurrence of self-injury varied widely both between and within subjects, necessitating an individualized treatment approach based on functional analysis. This method of assessment is widely considered to be a significant advancement in the field of applied behavior analysis (Carr, 1994; Mace, 1994).

In subsequent developments, Repp et al. (1988) provided validity for the functional assessment process in demonstrating that only treatments corresponding to the function of a target behavior were effective in the reduction of that behavior. Day, Rea, Schussler, Larsen, and Johnson (1988) also used treatment outcome as a validity measure when they conducted analog functional analysis sessions across three conditions: alone, reinforcer withdrawal, and academic demands. They validated the results of their assessments on three participants by exposing each client to a treatment condition appropriate to the function of the behavior (e.g., functional communication training to request reinforcers for behavior maintained by access to tangible items). Following treatment implementation, the self-injury of each subject diminished in the treatment condition and returned to higher rates during control session probes. Finally, Durand and Crimmins (1988) included a

tangible condition into the analog assessments they conducted in their validation study of the Motivation Assessment Scale.

Repp, Singh, Olinger, and Olson (1990) observed that most studies of self-injurious behavior failed to report functional hypotheses. However, by 1993, approximately 70% of research studies were reporting some method of functional assessment in the treatment of aggression in persons with mental retardation (Hile & Desrochers, 1993). Iwata and colleagues (1994), in reviewing the functional analysis research they conducted in the previous decade, suggested that when treatment was matched to behavioral function the need for punishment-based procedures was greatly reduced. In a data set of 152 functional analyses, 95% showed differential responding in at least one condition. Escape was the function for the largest proportion of the sample (38%), followed by positive reinforcement (26%), automatic reinforcement (26%), multiple sources of control (5%), and uninterpretable (5%). In a similar group analysis, Vollmer, Marcus, Ringdahl, and Roane (1995) demonstrated differential responding in 85% of their subjects; however, their sample size was limited. Of 20 participants, 35% were responsive to positive reinforcement, 30% to negative reinforcement, and 30% to automatic reinforcement. Finally, Nguyen, Derby, Hagopian, Fisher, Thompson, and Owen (1996) found that in a sample of 71 participants, 38% responded to positive social reinforcement, 40% to negative reinforcement, 7% to tangible reinforcement, 7% to automatic reinforcement, and 11% to multiple sources of control. However, most of the participants were less than 11 years old; therefore, the generalizability of these findings is limited to younger children.

In an interesting reversal of perspective, Owens and Ashcroft (1982) noted that the effectiveness of default treatment interventions may reflect the number of distinct functional processes a behavior of interest may serve. For example, if an attention-based DRO treatment proves effective in 40% of cases of self-injury, then it may be possible that self-injury is maintained by attention in 40% of affected individuals.

With the development of functional assessment techniques, psychologists were able to determine that one behavior may, in fact, serve multiple functions for an individual depending on the environment (e.g., escape from academic demands and attention in isolated settings) (Favell, 1989; Rojahn & Schroeder, 1991). In contrast, topographically different behaviors (e.g., SIB and aggression) may serve the same functions for an individual (Sturmey, 1996). Derby et al. (1994) confirmed that separate functions for different behaviors could be assessed by separating data by topography.

Treatment of problem behaviors now focused on two issues: (1) weakening the relationship between the target response and its maintaining reinforcer and (2) strengthening the response-reinforcement relationship of a substitute adaptive behavior (Mace, 1994). Behavior therapy techniques expanded to include teaching functional replacement behaviors and environmental rearrangement in addition to modification of consequences (Sprague & Horner, 1995). Teaching of functional replacement behaviors now is specified as an individual right to treatment (Association for Behavior Analysis, Task Force on the Right to Effective Behavioral Treatment, 1988). Functional assessment was recognized as an important variable in defining curriculum content (Horner, Sprague,

& Flannery, 1993). Durand, Crimmins, Caulfield, and Taylor (1989) suggested that the reinforcers identified by functional analysis could be used to train other adaptive skills in what has been termed "functional equivalence training".

Summary

Functional analysis decreased the tendency to default to topography-based procedures by promoting assessment of environment-behavior interactions as the basis for treatment (Carr, 1994; Mace, 1994; Owens & Ashcroft, 1982). As a result, etiologic theories for problem behaviors include an understanding that multiple processes are involved and will differ on an individual basis (Belfiore & Dattilio, 1990).

Methodologies

The following section discusses the three available methodologies for conducting a functional assessment: direct observation, interviews and behavioral checklists, and formal functional analyses. Each method first is described and then the advantages and disadvantages are discussed in terms of required resources and degree of experimental control.

Direct observation

Direct observation methods are a foundation of behavior analysis (Baer et al., 1968), although the technology for conducting direct observations has become more refined over the years. Two commonly used formats for direct observations include the scatterplot and the antecedent-behavior-consequence (ABC) assessment (Lennox & Miltenberger, 1989).

Touchette, MacDonald, and Langer (1985) introduced the scatterplot assessment as a means to compare levels of the target behavior with time of day/activity variables. Each occurrence of the target behavior is plotted on a grid with time of day on the ordinate and consecutive days on the abscissa. Differentially high rates of the behavior can be observed by patterns on the temporal grid. Although the scatterplot does not provide information on all of the maintaining variables of the behavior, it provides useful preliminary information that may guide the direction of future assessment.

The ABC assessment involves recording each instance of the target behavior with the events that immediately preceded and followed it (Bijou, Peterson, & Ault, 1968; Kazdin, 1980). Repp et al. (1988) demonstrated that treatments based on ABC analyses were more effective.

Mace et al. (1993) described a multi-stage methodology for conducting an ABC assessment. First, the participant is monitored in an unstructured observation. Next, 3-5 categories each are developed for antecedent, response, and consequent events. Then a 10-sec partial interval recording system is used to note behaviors, antecedents, and consequences for a 15-60 minute structured observation by a pair of observers. Antecedents are recorded continuously while consequences are recorded if they occurred within 30-sec of the behavior.

Stereotypy Analysis.

Pyles, Riordan, and Bailey (1997) described a highly-detailed systematic observation system known as the Stereotypy Analysis for identifying environmental mediators to

stereotypic behavior. The Stereotypy Analysis identifies correlational relationships between specified environmental events and the relative level of stereotypy. A minimum of 70 one-minute observations are conducted and the presence or absence of prespecified environmental conditions (e.g., presence of demands, size of setting, time to next meal) are noted for each observation. If a condition is infrequently observed, the authors suggest repeating observations until all 19 conditions are observed on at least 15 occasions. A base rate of stereotypy is calculated across all observations. Next, the probability of stereotypy given each environmental condition is obtained. If the percentage of intervals for a given condition exceeds 10% of the base rate (authors' criterion), that condition is considered to be a potential environmental mediator for the stereotypy.

Pyles et al. (1997) examined the Stereotypy Analysis with five participants. Interrater agreement on occurrence of stereotypy averaged 95.2%. For three of the participants, an environmental indicator was suggested. The authors then conducted a treatment analysis with two participants in which treatment indicated by the Stereotypy Analysis yielded significant reduction in stereotypy levels when compared to a random treatment procedure.

Although the results of this study are promising, the extremely small sample size limits generalizability of the psychometric results. Replication with a larger sample size is warranted. In addition, validity assessment of this instrument should include all participants, regardless of the outcome. Participants with stereotypy not mediated by the environment should participate in the same comparison of indicated and nonindicated

treatments to rule out "false negatives" when identifying environmental maintaining variables.

Advantages of direct observation.

The primary advantage of direct observation is that one can observe the range of antecedent and consequent events in a manner more objective than verbal reports (Linscheid et al., 1996). The method has low cost relative to experimental techniques and can pinpoint time of day or other specific variables (Sprague & Horner, 1995). Mace (1994) noted that direct observation may often be used to identify idiosyncratic reinforcement conditions and provide an estimate of reinforcer schedules established in the natural environment.

Disadvantages of direct observation.

The major limitation of direct observation is that it yields correlational data which do not always correspond to experimental results (Mace, 1994; Sprague & Horner, 1995). An example of correlational data would be the child who head bangs and receives teacher attention each time. The hypothesis from this observation may be that attention maintains the behavior. However, the child may actually have a sinus headache and a treatment of teaching request for attention may not reduce the behavior. Such correlational errors may be due to observation of caregivers who are inconsistent in their responding to the target behavior (Lerman & Iwata, 1993).

A second difficulty is that observational methods can be difficult to standardize (Mace, 1994). For example, observers without experience using ABC assessments may record

global events unrelated to the target behavior rather than the specific events which are immediately related (Lennox & Miltenberger, 1989). The same events associated with the target behavior may also occur at other times and be correlated with nonoccurrence of the behavior (Iwata et al., 1990). Finally, behaviors that occur on a thin schedule of reinforcement may yield direct observations results that are difficult to interpret. Mace and Lalli (1991) and Lerman and Iwata (1993) found that formal descriptive analyses yielded results that were inconsistent with experimental sessions. These authors concluded that descriptive observations may be neither necessary nor adequate for identifying reinforcers for a target behavior. In general, direct observations are time-consuming, yield large quantities of data that may be difficult to interpret, and require extensive training (Sturmey, 1996).

Structured interviews and behavioral checklists

Functional Analysis Interview Form.

A complete behavioral interview should elicit information about the topography of the behavior, antecedent and consequent events, rate of the behavior, and setting events (Miltenberger & Fuqua, 1985). The Functional Analysis Interview Form (FAIF) is the only structured interview available to assess function of behavior. The FAIF is a component of the manual Functional assessment of problem behavior: A practical assessment guide published by O'Neill, Horner, Albin, Storey, and Sprague (1990). The FAIF consists of nine sections of open-ended questions and short forms in areas of (1) the problem behaviors, (2) potential ecological events, (3) events and situations that predict

occurrences of the problem behavior, (4) identifying the functions of the problem behavior, (5) the efficiency of the problem behavior, (6) the person's primary mode of communication, (7) functional alternative behaviors, (8) history of the problem behaviors, and (9) previous attempts at treatment. The interview requires 45-90 minutes for completion. O'Neill et al. emphasize that their manual is intended to be used with flexibility by clinicians with training in applied behavior analysis. Therefore, no psychometric data are available.

Although not an effective tool for measurement, the FAIF addresses areas of antecedent and setting events, potential treatment options, and several consequence variables, thus providing a wide-based and complex set of data that a trained clinician must interpret. What is not clear is if this broader set of data yields treatment interventions that prove to be more effective than those from other functional assessment methods (Sturmey, 1994). In short, some effort at establishing the reliability and validity of this instrument would better determine if its utility justifies its length and method of administration.

Motivation Analysis Rating Scale.

Weiseler et al. (1985) developed the Motivation Analysis Rating Scale (MARS), the first checklist to assess the functions of stereotypy and self-injury. The MARS is a six-item checklist rated on a 4-point Likert scale from 1 (almost never) to 4 (almost always). Three subscales are derived from the three pairs of items: Positive environmental consequences, Task escape/avoidance, and Self-stimulation. In the initial study, the

authors assessed 23 participants with self-injury and 37 with stereotypy. Interrater agreement on the primary function of the behavior was 73%. For those participants for whom raters were able to agree on a primary function, the authors observed 95% agreement between the MARS and naturalistic observations of antecedents and consequences to the target behavior.

No additional psychometric studies have been conducted on the MARS. This fact may be due to the inherent limitations of this instrument, given that it consists of only six items. Such a small number of items may compromise its psychometric integrity (Sturmey, 1994). Additionally, the high validity coefficient would be artificially inflated due to the sample selection of participants for whom congruent ratings were achieved.

Motivational Assessment Scale.

To provide an alternative to formal functional analyses, Durand and Crimmins (1988) developed the most extensively assessed checklist targeting functions of problem behavior: the Motivation Assessment Scale (MAS). The MAS is a 16-item questionnaire divided into 4 clinically-derived subscales addressing sensory, escape, tangible, and attention functions. Each item is rated on a 7-point Likert scale ranging from 0 (never) to 6 (always), and the entire scale may be completed within 5-10 minutes. The scores within each subscale are added and mean subscale scores are computed. The probable functions of the behavior are determined by rank ordering the subscale mean scores.

Durand and Crimmins (1988) selected a highly homogeneous sample for their initial psychometric studies on the MAS. Fifty children with developmental disabilities who

displayed self-injury at a rate of at least 15 behaviors per hour participated in the reliability assessment. Classroom teachers completed the MAS for each participant. Interrater reliability as computed by Pearson correlation coefficients for the individual questions was adequate to good (r = .66-.92); reliability coefficients for the subscale mean scores were good (r = .80-.95). Test-retest reliability for the individual items and the subscale means was high (r = .89-.98 and r = .92-.98, respectively).

Eight of these participants were included in the validation portion of the study. These participants were selected as the results of the MAS indicated a primary function in one of the four domains: attention, sensory, escape and tangible. Analog sessions were conducted and the results summarized in bar graphs that displayed the mean percent of intervals with self-injury in each condition. The primary functions identified by the MAS were correlated to the rankings from functional analyses. The overall correlation indicated a very high degree of correspondence (r = .99). In contrast, when the teachers were given descriptions of the four functions and asked to make their rankings by opinion, the correlation of opinion to the results of the functional analyses was low (r = .21). The authors concluded that the MAS was reliable, yielded information that could not have been generated from interviews, and was predictive of behavior in analog settings.

Later psychometric studies of the MAS offered variable results. Newton and Sturmey (1991) used a smaller sample (n=12) and found considerably lower interrater reliability for both the individual items (r = -.20-.70) and subscale scores (r = .25-.49). Internal consistency for the subscales ranged from alpha =.69-.91. These authors attributed their

findings to the inclusion of a range of behaviors other than self-injury, use of untrained direct care staff as raters, and the inclusion of participants with variable rates of behavior. Additionally, they observed variable inter-item correlations, suggesting that the clinically derived subscales may not be supported statistically.

Bihm, Kienlen, Ness, and Poindexter (1991) examined the factor structure of the MAS in a sample of 118 participants with severe and profound mental retardation and a range of problem behaviors. They found that the factor structure of the MAS generally supported the clinical subscales, although two items did not load on expected factors and two items did not load on any factor. The internal consistency was comparable to that obtained by Newton and Sturmey (1991) ($\underline{r}$ = .69-.81). Bihm et al. (1991) cited similar limitations to their study in that they assessed a range of behaviors with untrained direct care staff.

Zarcone, Rodgers, Iwata, Rourke, and Dorsey (1991) reported a failure to replicate the interrater reliability noted by Durand and Crimmins (1988). Using a sample of 55 participants with self-injury in both school and institutional settings, these authors found low correlations for interrater reliability. Pearson correlation coefficients for individual items ranged from $\underline{r}$ = -.51 to .55 and from $\underline{r}$ = -.80 to .99 for the subscales. This study also provided percent agreement data on individual items and subscales; percent agreement ranged from 8% to 59% for the items and 0% to 63% for the subscales. In sum, only 15% of the correlation coefficients and none of the percent agreement scores exceeded a minimum value of .80. In discussing these results, the authors suggested that ambiguity in wording as well as phrasing of items in a manner implying knowledge of

behavior analysis could lead to the discrepant findings in scale reliability. They concluded that verbal report measures could not replace direct observation of behavior.

However, Crawford, Brockel, Schauss, and Miltenberger (1992), in comparing the MAS to ABC observations and analog conditions, suggested that the MAS has potential utility as a cost-effective and abbreviated functional assessment tool. These authors concluded that using several assessment measures and looking for concordance in results would be best practice at this time. Yet in their assessment of interrater reliability, these authors found considerable variability between raters (r = -.306 to .974).

Singh et al. (1993) further assessed the internal validity of the MAS through factor analysis. These authors examined the results of the MAS with two population subgroups: an institutional sample with high rates of self-injury and a school sample with lower rates. Master's level psychologists or certified school teachers completed the rating forms. They identified a clinically significant four-factor solution corresponding to the clinically derived subscales with the institutional sample only. This sample displayed rates of behavior of at least 15 times per hour as in the original Durand and Crimmins (1988) study.

Sigafoos, Kerr, and Roberts (1994) identified another limiting condition to the MAS in their study assessing interrater reliability with aggressive behavior. They found highly variable Pearson correlation coefficients across pairs of raters (r = -.667 to .722) and generally low correlation coefficients across individual items (r = -.337 to .425). These

authors determined that the MAS was not a viable instrument in assessing aggressive behavior.

Finally, Spreat and Connelly (1996) submitted the MAS to several reliability assessments but stipulated that only the category scores required analysis, as the individual items were not to be used in clinical decision-making. They observed that although the MAS had fair internal consistency, its interrater reliability was insufficient. Specifically, the reliability of subscale scores ranged from $\underline{r} = .31$-$.57$ and the reliability of difference scores between subscales was inadequate ($\underline{r} = .37$ to $.78$). Spreat and Connelly concluded that the MAS was insufficient to use as an isolated assessment tool and that revision of the scale with the addition of items would be warranted to improve upon its psychometric properties.

In summary, although the MAS was intended to serve as an inexpensive and efficient substitute for analog functional analysis techniques, its psychometric properties are insufficient to support such wide-scale use. Given above research, the MAS appears to have reliability and validity only in circumstances where the persons completing the rating are well-trained in behavior management and the participants display high rates of self-injury. In terms of scale construction, the MAS targets only four consequences of behavior and does not include antecedent or setting events. Assessing behaviors that are not motivated by these consequences may contribute to the unreliable psychometrics of this instrument (Sturmey, 1994). Finally, in the original study of convergent validity (Durand & Crimmins, 1988), participants were not randomly selected. The selection of

participants based on clear-cut results on the MAS limits the generalizability of findings to participants with one clear primary function to their behavior.

Functional Analysis Checklist.

The Functional Analysis Checklist (FAC) (Van Houten & Rolider, 1991) is a 15-item functional assessment measure addressing the physical environment, adjunct behaviors, transitions between activities, escape from demand, and positive reinforcement. Van Houten and Rolider caution that this measure is not intended to be all-inclusive as variables such as illness and discomfort are not addressed. Rather, they suggest that the FAC be used to develop hypotheses regarding the target behavior to be assessed through experimental analysis.

Sturmey (1994) reported poor interrater and test-retest reliability on this instrument when a sample of 30 participants with a wide range of behaviors and behavior frequencies was used. Sturmey (1994) hypothesized that, as with the MAS, replication of the psychometric properties of this instrument may indicate that it is applicable only to trained informants reporting on high frequency behaviors.

Evaluation of Existing Checklists.

None of the current functional assessment checklists possess adequate or replicable psychometric properties. However, at least some of these instruments seem reliable in certain circumstances; that is, when trained informants are used and high-frequency behaviors are assessed. The earlier measures (i.e., the MARS and the MAS) assessed consequences of behavior only. The newer measures (i.e., the FAIF and FAC) assess a

broader range of variables; however, they provide complex data that do not lead directly to treatment. Clearly, further research in this area is needed before an effective measure applicable across a range of behaviors and settings is identified.

Advantages of interviews and checklists.

Rating scales and interviews are the most cost-effective and efficient methods of functional assessment (Sprague & Horner, 1995). Clinicians can administer rating scales to large groups of people in a minimal amount of time (Sturmey, 1994). Administration requires considerably less training than other methods of functional assessment (Sturmey, 1996). Finally, interviews elicit the greatest range of information; that is, they identify establishing operations and antecedents, idiosyncratic consequences, reinforcement schedules, and other contingencies that other methods do not always assess (Sprague & Horner, 1995).

Disadvantages of interviews and checklists.

These methods are the most efficient means of generating hypotheses of functional relations; however, the results may correspond poorly to experimental data (Sprague & Horner, 1995; Sturmey, 1994). Currently, interview methods are not psychometrically sound and therefore difficult to compare to other methods. With checklists, the range of variables assessed can be limited depending on the instrument (Sturmey, 1996).

With both methods, the examiner does not observe the behavior directly but must rely on the informant's recollection, which may not always be accurate (Kazdin, 1980; Linscheid et al., 1996). Informants must be able to accurately discriminate and report a

client's behavior in behavior analytic terms (Sturmey, 1996). Rating scales in general may be subject to errors of leniency or severity, central tendency, and the halo effect (Pedhazur & Pedhazur-Schmelkin, 1991). Errors of leniency or severity imply a tendency to give ratings that are consistently too high or too low, while errors of central tendency suggest avoidance of extreme categories. With the halo effect, a rater's general impressions bias ratings on more specific and distinct aspects of the construct being assessed.

Experimental functional analyses

Bijou et al. (1968) stated that descriptive studies can only describe relationships between events; experimental studies are required confirm functional relationships. From simple reversal designs demonstrating functional control (e.g., Lovaas et al., 1965) to the well-established multi-element design of Iwata et al. (1982/1994), experimental functional analyses are considered to be the most valid method of identifying functional relations (Favell, 1989; Martin & Pear, 1996; Van Houten & Rolider, 1991; Wacker, 1989). The present section will discuss the advantages and disadvantages of the Iwata et al. (1982/1994) design and the refinements in methodology that developed in the present decade of research.

Advantages of functional analyses.

Functional analyses provide the only empirical demonstration of functional control and the most direct and reliable match between function and treatment (Mace, 1994; Sprague & Horner, 1995; Van Houten & Rolider, 1991). They demonstrate the relative

independence of several problem behaviors, and predict the level of change when treatment is withdrawn (Van Houten & Rolider, 1991). The validity of this approach has been established for self-injurious behavior (Day et al., 1988), aggression (Slifer, Ivancic, Parrish, Page, & Burgio, 1986), stereotypy (Sturmey, Carlsen, Crisp, & Newton, 1988), and disruption (Carr & Durand, 1985).

Disadvantages of functional analyses.

This technique is the most costly to administer and requires the most extensive training; therefore, it is not easily replicated by teachers or other caregivers (Lennox & Miltenberger, 1989; O'Neill et al., 1990; Sprague & Horner, 1995; Sturmey, 1995). Administration may require from one to two weeks of several sessions per day (Iwata et al., 1994). Next, as control over environmental variables is difficult to achieve in the natural environment, functional analysis requires the arrangement of analog sessions that correspond to environmental situations such as academic demands, play, etc. However, these techniques may not identify unique conditions in the natural settings. Consequently, the results can be difficult to generalize and may not be ecologically valid (Mace, 1994; Mace, Lalli, & Pinter-Lalli, 1991; Sturmey, 1995). Low frequency behaviors in particular may not be seen during sessions, especially when the frequency is as low as once per day or less. When behaviors are also of high intensity and risk (e.g., sexual assault), a functional analysis may not be ethically appropriate (Sturmey, 1995; Vollmer & Smith, 1996).

In the area of learning, new functions can become acquired as a result of the assessment (Linscheid et al., 1996; Pyles et al., 1997). Behaviors with an extensive past learning history often acquire complex, multifunctional properties (Rojahn & Schroeder, 1991), but only a limited range of variables can be investigated (Linscheid et al., 1996; Sturmey, 1995).

Several methodological problems are associated with functional analyses. First, carry-over effects may exist within the multi-element format. That is, a client may not respond to the change between conditions, leading to undifferentiated results (McGonigle, Rojahn, Dixon, & Strain, 1987; Sturmey, 1995). Next, session termination criteria may be too lenient or too conservative, resulting in patterns of behavior that do not correspond to the natural environment (Sturmey, 1995). For example, if session termination criteria are defined as the client displaying redness or other sign of injury, but the client is interrupted and restrained upon the first occurrence of self-injury in the natural environment, the response pattern in the functional analysis will not be ecologically valid. Finally, non-targeted behaviors may require a consequence to be provided in the session as well, especially if they are part of a response chain which eventually results in the target behavior (Sturmey, 1995). For example, a client may repeatedly get out of his seat and attempt to leave the session room; if these behaviors are ignored as per the protocol, subsequent aggression may increase to a rate beyond that in the natural environment. Sprague and Horner (1992) confirmed that a decrease in one member of a response class produced an increase in another.

Rojahn and Schroeder (1991) concluded that analog functional analyses may yield results that are as vague as without the procedure when a behavior is under complex, multifunctional control. While the benefits in validity of conducting functional analyses generally outweigh the disadvantages, many of the limitations could be circumvented if brief yet valid assessments could be identified (Vollmer, Iwata, Zarcone, Smith, & Mazaleski, 1993).

Recent methodological refinements.

To counter carry-over effects and discrimination problems with some participants, Iwata, Duncan, Zarcone, Lerman, and Shore (1994) suggested a sequential test-control methodology. In their procedure, each test condition is presented sequentially as in typical reversal designs. However, within each phase a control condition also is run in a multi-element format with the test condition. The control condition remains the same across all test conditions. This method clarified the previous undifferentiated findings in two of three of their participants, although it proved to be less time efficient.

Northup et al. (1991), citing the need for abbreviated procedures when time is a consideration in outpatient settings, developed a brief functional analysis procedure that could be accomplished within a 90-minute evaluation period. The analog sessions developed by Iwata et al. (1982/1994) served as the basis for structuring the assessment. Each condition was conducted for 10 minutes. Then, the condition with the highest percent interval of aggression was repeated in a contingency reversal where the consequence was given only for the emission of an appropriate replacement behavior (e.g.,

stating, "Come here please," for attention-maintained behavior). The authors concluded that this brief functional analysis was an effective abbreviated procedure.

Vollmer, Iwata, Zarcone, et al. (1993) noted two possible limitations to the brief assessment format. First, similar rates of behavior may be observed across two conditions suggesting multiple control; however, the behavior actually may have been increasing in one condition and decreasing in the other. Next, extinction bursts within a condition may yield high rates of behavior and be misinterpreted as indicating a source of control. Therefore, these authors suggested examining within-session trends (i.e., minute-by minute rates of behavior within each condition) to accurately explore such fluctuations. They compared the results of brief within-session analyses to extended functional analyses and found that within-session analyses effectively predicted the source of control in 3 out of 4 subjects (the fourth subject demonstrated possible cyclical patterns in the extended analysis although both analyses suggested multiple control).

Research into functional analysis has broadened into many avenues. Functional analysis has been applied to ADHD symptoms (Northup et al., 1997), autistic symptoms (Ivanick, Helsinger, & Harris, 1996), bizarre speech (Mace & Lalli, 1991), breath holding (Kern, Mauk, Marder, & Mace, 1995), drug ingestion (Chapman, Fisher, Piazza, & Kurtz, 1993), elopement (Piazza, Hanley, Bowman, Ruyter, Lindauer, & Saiontz, 1997), eye-poking (Lalli, Livezey, & Kates, 1996), feeding problems (Munk & Repp, 1994), food refusal (Maenpaa, Lockwood, & Williams, 1996), hand mouthing (Goh, Iwata, Shore, DeLeon, Lerman, Ulrich, & Smith, 1995), inappropriate social behavior (Frea & Hughes,

1997), noncompliance (Rortvedt & Miltenberger, 1994), pica (Lockwood, Borden, & Williams, 1996; Piazza, Hanley, & Fisher, 1996), and has utilized parents as therapists (Cooper, Wacker, Sasso, Reimers, & Donn, 1990).

Additional concerns have included improving the accuracy of the assessment techniques, developing more effective behavioral interventions, making the assessment techniques feasible for wide-scale use, and making the technique cost-effective (Sprague & Horner, 1995). Matson and Coe (1992) called for the use of functional assessment to better assess the contingencies that occur early in the development of psychiatric disorders in persons with dual diagnosis. Other avenues include identifying more specific variables maintaining problem behavior (e.g., social avoidance, respondent conditioning, or biological reinforcers) (Carr, 1994; Taylor, Ekdahl, Romanczyk, & Miller, 1994). Techniques in general need to be able to include the variety of antecedent, consequence, and setting event variables that can influence the problem behavior of one individual (Carr, 1994; Horner, 1994; Repp, 1994).

Best practice

Several authors have called for the use of multiple assessment measures, reasoning that concordance in results would provide the most valid conclusions on functional relationships. They suggest a multistage methodology beginning first with a behavior checklist that would provide the data for a hypothesis of functional relationships. This hypothesis could be tested through experimental analog sessions and ultimately verified through treatment intervention, generalization, and maintenance (Crawford et al., 1992;

Mace et al., 1991; Sturmey, 1995). Lennox and Miltenberger (1989) and Mace et al. (1991) stressed that while experimental methods provide the most rigorous methodology, direct observation and rating scales are still essential in that they contribute to a multi-modal assessment and minimally, provide more benefit than treatment selection without any method of functional assessment. Even if these indirect methods identify the function in fewer cases than analog assessments, they would still reduce the amount of time before treatment is implemented (Vollmer et al., 1995). They also are more ethically appropriate for high risk behaviors and behaviors of low frequency (Mace et al., 1991). Direct observation is useful for identifying idiosyncratic events that may adversely affect functional analysis results (e.g., the presence or absence of materials in the assessment room, therapist variables, etc.) (Carr, Yarbrough, & Langdon, 1997). It would be worthwhile to explore how each functional assessment strategy contributes to the design of effective treatments technology (Hiles & Desrochers, 1993).

With specific regard to experimental methods, Vollmer, Marcus, Ringdahl, and Roane (1995) identified a comprehensive, sequential methodology that allows a clinician to proceed from least time-consuming and intrusive methods to more complex procedures when needed. They suggested beginning with brief assessments interpreted via minute-by-minute analysis (i.e., Vollmer, Iwata, Zarcone, et al., 1993). When these assessments prove inconclusive, an extended multielement design is used. If this also proves inconclusive, extended alone sessions are conducted. If the behavior persists, an automatic function may be assumed; if the behavior extinguishes, social contingencies are

assumed to be in place. Finally, a reversal design is used if unclear results persist to account for possible difficulty in discriminating between conditions by the participant (e.g., Vollmer, Iwata, Duncan, et al., 1993).

Questions About Behavioral Function (QABF)

The QABF (Matson & Vollmer, 1995) is a questionnaire designed to identify the variables maintaining problem behavior in persons with mental retardation. These variables include social attention, escape, tangible reinforcement, physical discomfort, and nonsocial reinforcement. This is a broader range of variables than previously addressed by other measures. In particular, under-investigated variables such as social avoidance (Taylor, Ekdahl, Romanczyk, & Miller, 1994) and physical discomfort (Carr, 1994; Lowry & Sovner, 1991) are included to make the scale more comprehensive than previous instruments.

The QABF is administered to an informant familiar with the client's behavior for at least six months. Results of the QABF can serve as an initial basis for intervention, although the authors state that the intervention may need to be modified after extended functional analysis sessions or treatment inefficacy. The QABF is intended for use in situations where practitioners do not have the time, resources, or training to administer formal functional analyses. While previous functional assessment measures yielded mixed psychometric outcomes, the QABF has been prepared in a straight forward manner that more directly assesses the behavior of interest (e.g., "engages in the behavior to

escape from work"). Minimal ambiguity within the items may lead to better reliability and validity for the instrument (Fowler, 1988).

The QABF is a 25-item questionnaire scored on a 4-point Likert scale from (0) never to (3) often. If an item does not apply to the behavior (e.g., "engages in the behavior when alone" and the behavior is aggression), the "N/A" option may be endorsed. When the scale is completed, the total number of items endorsed as occurring at least "rarely" within each subscale are circled on a score summary sheet. In addition, a total score for each subscale is computed by adding the individual item scores within that subscale. Item endorsement and total score are used to indicate potential behavioral functions. If several items are high on item endorsement, the total scores are used to differentiate between possible functions. When there is no discernible scoring outcome, the functional assessment is determined to be undifferentiated; that is, the behavior may be occurring at high/low rates across several situations.

Matson et al. (1996) presented the initial psychometric data for this instrument. Data were collected at a state institution for persons with mental retardation. Participants (n=462) ranged in age from 13-86 and were predominantly in the profound range of mental retardation. The types of behaviors assessed varied but included typical problem behaviors such as SIB, aggression, and property destruction. Internal reliability and validity data were assessed. Coefficient alpha and Guttman split-half reliability coefficients were very acceptable ($\underline{r}$=.86 and $\underline{r}$=.91, respectively). An exploratory factor analysis with varimax rotation yielded five factors accounting for 74.5% of the variance

and corresponding to the five hypothesized subscales. These initial psychometric data are promising, but further data are required to substantiate that the QABF is a reliable and valid tool for functional assessment.

Rationale

Functional assessment techniques have made a substantial impact on improving outpatient-based behavioral treatment (Wacker et al., 1994). The need for efficient methods of conducting functional assessments in varied settings is an essential future direction for the field (Carr, 1994). Of particular significance is the need for methods requiring minimal time, effort, and training on the part of the clinician (Horner, 1994; Sturmey, 1994).

Of the functional assessment methodologies, experimental functional analysis has the greatest precision and experimental rigor. However, it is time-consuming, requires considerable expertise, and assesses a limited range of variables. Although descriptive measures may be less definitive, they are also less intrusive and require less time. Development of a reliable and valid behavioral checklist for functional assessment would significantly contribute to applied work in behavior modification (Vollmer & Smith, 1996).

In the present research, we attempt to address this need by providing a range of psychometric data for the QABF. First, to determine the stability of the instrument across raters and over time, interrater and test-retest reliability were assessed. Next, to substantiate the initial factor solution identified by Matson et al. (1996), a second exploratory factor analysis was conducted. Finally, to provide an estimate of convergent validity, the results of the QABF were compared to experimental functional analyses and the Motivation Assessment Scale (Durand & Crimmins, 1988) with a larger sample size

than in previous research. The proposed data analysis plan is described in greater detail in the sections that follow.

Method

Site

All of the data were collected at Pinecrest Developmental Center (PDC) in Pineville, Louisiana. A total of 730 clients reside at PDC. Table 1 outlines the demographic characteristics of this facility. These data were collected as a part of ongoing clinical services provided to these clients and were given to developmental center staff for incorporation into existing Behavior Treatment Plans. Institutional Review Board approval for this project was obtained under a previously approved research proposal entitled: "Norming Psychological Assessment Battery for Treatment Plans". Individual consent for participation in experimental functional analyses was obtained from the parent or legal guardian of each participant in this phase of the study (see Appendix B for consent form).

Reliability

Reliability data were collected first as adequate reliability is an essential precursor to later validity studies (Crocker & Algina, 1986). As the internal reliability of the QABF previously was assessed by Matson et al. (1996), the following study attempted to replicate these data and to provide additional external reliability data of test-retest and interrater statistics.

Table 1: Demographic characteristics of Pinecrest Developmental Center (n=730)

Demographic	Percentage
Age	
0-21	5.07
22-45	48.90
46-65	34.38
66+	11.64
Gender	
male	55.75
female	44.25
Race	
Caucasian	76.58
African American	23.29
Native American	.13
Level of mental retardation	
Mild	2.47
Moderate	5.48
Severe	15.48
Profound	76.58

Test-retest

Data from 34 participants were included in this portion of the study. The following information was obtained for each participant: age, gender, race, level of mental retardation, and target behavior. Table 2 outlines the demographic characteristics of this sample.

A trained examiner (LSU graduate student) administered the QABF to an informant who had known the client for at least 6 months. The same examiner readministered the QABF to the same informant between 1-3 weeks after the initial administration. This interval was chosen to ensure that the QABF was stable over time and did not reflect the participant's behavior over a short interval.

Pearson Product-moment correlation coefficients were computed to determine the pattern of agreement for the total and subscale scores (Sattler, 1992). For the individual items, Spearman rank order correlation coefficients were calculated as the Likert ratings were ordinal in nature (Fowler, 1988; Siegel & Castellan, 1988). In addition, to determine the exact agreement between items during separate administrations, the percent agreement was computed (Sturmey, 1996). To correct for chance levels of agreement, Cohen's Kappa was calculated for the individual items, subscale and total scale scores (Cohen, 1960; Sattler, 1992).

Table 2: Demographic characteristics of test-retest sample (n=34)

Demographic	Percent of sample
Age	
22-45	50.0
46-65	43.8
66+	6.2
Gender	
male	82.4
female	17.6
Race	
Caucasian	79.4
African American	20.6
Level of mental retardation	
Mild	2.9
Moderate	8.8
Severe	11.8
Profound	76.5

table 2 continued

Target behavior	
self-injury	32.4
aggression	11.8
property destruction	2.9
tantrums/verbal aggression	29.4
stereotypy	5.9
pica	8.8
stealing	5.9
elopement	2.9

Interrater

Twenty-three additional participants were used for this portion of the study. The same demographic information collected for the test-retest sample was compiled for this sample (see Table 3). While this sample consisted of all males, the proportion of target behaviors remained similar to that of the test-retest sample and gender has not been associated with differential functional analysis outcome.

Table 3: Demographic characteristics of interrater sample (n=23)

Demographic	Percent of sample
Age	
0-21	8.7

table 3 continued

22-45	60.2
46-65	17.2
66+	12.9
Gender	
male	100.0
Race	
Caucasian	78.3
African American	21.7
Level of mental retardation	
Mild	4.3
Moderate	13.0
Severe	26.1
Profound	47.8
Unspecified	8.7
Target behavior	
self-injury	26.1
aggression	21.7
property destruction	4.3
tantrums/verbal aggression	30.4

table 3 continued

stereotypy	4.3
pica	4.3
stealing	4.3
rectal digging	4.3

A trained examiner (LSU graduate student) administered the QABF to an informant who knew the client for at least 6 months. The same examiner readministered the QABF to a different informant within 1 month of the first administration.

As with the test-retest sample, Pearson correlation coefficients were computed for the total subscale scores, and Spearman rank order correlation coefficients were calculated for the individual items. Again, as these correlation coefficients assess degree of association only, the actual agreement between raters for individual items was calculated by percent agreement (Sturmey, 1996; Zarcone et al., 1991) and by Cohen's Kappa for the individual items, subscale and total scores.

Internal reliability

Completed QABF scales for 243 additional participants were utilized to conduct the internal reliability studies and a second exploratory factor analysis. Demographic characteristics for this sample are listed in Table 4.

Table 4: Demographic characteristics of internal consistency and factor analysis sample (n=243)

Demographic	Percent of sample
Age	
0-21	5.2
22-45	53.2
46-65	28.0
66+	11.6
Gender	
male	62.4
female	37.6
Race	
Caucasian	81.2
African American	18.8
Level of mental retardation	
Mild	3.3
Moderate	12.2
Severe	13.5
Profound	69.4
Target behavior	
self-injury	14.7

table 4 continued

aggression	9.8
tantrums/verbal aggression	36.4
stereotypy	23.2
pica	5.3
stealing	2.8
elopement	1.2
rectal digging	1.6

Data for the target behavior were unavailable for 12 subjects (4.9%). As with the previous assessments, a trained examiner (LSU graduate student or PDC Associate to a Psychologist) administered the QABF to an informant who knew the client for at least 6 months.

To estimate internal consistency, Cronbach's alpha was calculated for the test as a whole and for the individual subscales. Cronbach's alpha assesses the proportion of variation in the cases that is explained by the items (Cramer, 1994; Crocker & Algina, 1986).

To further assess scale consistency, split-half reliability was calculated on this sample using the Spearman-Brown formula. This statistic assesses the degree of correlation between two halves of the test when corrected for a reduced number of items (Pedhazur & Pedhazur-Schmelkin, 1991; Sattler, 1992).

Validity

Construct validity: factor analysis

Construct validity assesses the extent to which a test measures a particular psychological construct. One method of assessing construct validity is through a factor analysis, which determines whether the intercorrelations of items can be accounted for by a set of underlying factors within the measure (Sattler, 1992).

The most widely used method of extracting factors is known as Principal Axis Factoring (PAF). This method of assessing item intercorrelations explains the common variance shared by the items excluding error variance. To maximize the probability that each item will load predominantly on one factor, the matrix of item intercorrelations is subjected to transformation, or rotation. When the proposed factors are not hypothesized to be intercorrelated, an orthogonal rotation (varimax or quartimax) is used. More specifically, the varimax rotation is appropriate when no general factor is suggested (Pedhazur & Pedhazur-Schmelkin, 1991).

As previously stated, completed QABF scales for 243 participants were utilized to conduct a second exploratory factor analysis. This sample size is appropriate as it is approximately 10 times larger than the item pool (Pedhazur & Pedhazur-Schmelkin, 1991). The results of this analysis were used to replicate the initial findings of Matson et al. (1996) who identified a 5-factor solution using PAF with varimax rotation.

Convergent validity: Correlations with other measures

Convergent validity refers to the degree of correlation between maximally different methods assessing the same construct (Campbell & Fiske, 1959). The convergent validity of the QABF was assessed through comparison to experimental functional analyses and the Motivation Assessment Scale (MAS) (Durand and Crimmins, 1988).

Thirteen participants identified by PDC staff as needing a formal functional analysis were included in this portion of the study. This sample size was selected as it is similar to that of Durand and Crimmins (1988), who reported on the convergent validity of the MAS. Demographic characteristics for this sample are presented in Table 5.

Table 5: Demographic characteristics of convergent validity sample

Demographic	Percent of sample
Age	
22-45	69.3
46-65	30.7
Gender	
male	46.2
female	53.8
Race	
Caucasian	92.3
African American	7.7

table 5 continued

Level of mental retardation	
Severe	7.7
Profound	92.3
Target behavior	
self-injury	61.5
aggression	23.1
tantrums/verbal aggression	7.7
stereotypy	15.4

Informed consent for participation in functional analysis sessions was obtained from the parent or legal guardian of each participant. Each participant was included in a formal functional analysis methodology as outlined in Vollmer et al. (1995). That is, each session condition first was conducted once and the data analyzed based on minute-by-minute responding (see Appendix C for data sheet). If a pattern had emerged, this would have been replicated by repeating the conditions in which differential rates of responding were evident. However, in each case, the initial session results were unclear; therefore, the conditions were continued in a multi-element format. When the results remained unclear for one participant, extended alone sessions were conducted to determine if the behavior persisted in the absence of social contingencies.

Two to three LSU graduate students conducted the functional analysis sessions. Each student was trained in functional analysis procedures and achieved at least 80% interobserver agreement with another trained student on a functional analysis conducted prior to this study. These students were blind to the results of the QABF and MAS until the experimental sessions were complete.

Data were collected using one-minute intervals over a total of 10 minutes. The observers recorded the total frequency at which the behavior occurred within each interval except in two instances when the behavior was so rapid that occurrence/nonoccurrence during 30-second intervals was scored.

A total of 257 ten-minute sessions were required to complete the functional analyses for the 13 participants. Reliability was obtained for 54.1% of sessions. When the frequency measure was used, reliability for the session was computed by dividing the smaller frequency by the larger frequency within each interval, multiplying by 100%, and averaging across intervals. When the occurrence/nonoccurrence measure was used, reliability was calculated by dividing the number of agreements by the number of intervals and multiplying by 100% (Vollmer et al., 1995). Reliability averaged 95.7% with a range of 61.7-100%.

The following conditions were assessed in the experimental functional analyses: (1) attention, (2) demand, (3) tangible, (4) alone/ignore, and (5) toy play. However, a preferred item was not identified for two participants; therefore, the tangible condition

was not run in these cases. Also, the alone condition was not conducted for the three participants who displayed aggression as the target behavior.

Each condition was conducted in an empty day room located in the client's home or workshop area. The therapist for the session sat next to or behind the client as appropriate for the condition. One or two observers collected data from the far side of the room and did not interact with the client until the session was terminated. Each session had the following termination criteria: (1) occurrence of injury (redness or bleeding) to the client or therapist, (2) the therapist determines he/she cannot safely manage the client, or (3) the client falls asleep during the session. No session met termination criteria for each of the 13 participants.

The attention, toy play, and demand conditions were conducted in the manner described by Iwata et al. (1982). In the attention condition, one therapist sat near the client and read a magazine. The therapist provided verbal and physical attention only on the occurrence of the target behavior. In the toy play condition, the therapist provided attention once every 30 seconds, ignored any target behaviors, and made preferred objects available to the client. In the demand condition, the therapist presented functional demands (e.g., folding towels, stacking cups) to the client once every 30 seconds using a 3-step hierarchy of verbal prompts, modeling, and physical prompts. The therapist allowed the client to escape the demand for 30 seconds contingent on a target behavior.

In the alone/ignore condition, the therapist sat behind the client or out of the client's line of sight. The therapist ignored target behaviors. Finally, in the tangible condition, the

therapist gave the client a highly preferred object to interact with for two minutes prior to the start of the session. Then, the therapist removed the item once every 30 seconds and returned it contingent on any target behavior.

One examiner (LSU graduate student) administered the QABF and the MAS to an informant familiar with the client for at least 6 months. This examiner was blind to the results of the experimental sessions until all assessment was completed for that client. The QABF and MAS were administered during or up to one month after the experimental sessions were conducted.

The maintaining variables for the problem behavior as assessed by functional analysis sessions were determined using the criteria developed by Hagopian et al. (1997). The general procedure of these criteria is to use the toy play sessions as the method by which to quantitatively determine which experimental sessions are differentiated. Two criterion lines are drawn on the functional analysis graph: one at approximately one standard deviation (SD) above the mean and one approximately 1SD below. Differentiation occurs when the majority of the data points from a condition fall above the upper criterion line. More specific guidelines are described for the interpretation of upward and downward trends, low magnitude of effects, low-rate behavior, automatic reinforcement, and multiple maintaining variables. Twelve possible interpretations result: (1) undifferentiated, (2) maintained by attention, (3) maintained by escape from demands, (4) maintained by tangible reinforcement, (5) maintained by automatic reinforcement, (6) maintained by attention and escape), (7) maintained by attention and tangible reinforcement, (8)

maintained by tangible reinforcement and escape, (9) maintained by automatic reinforcement and escape, (10) maintained by automatic reinforcement and attention, (11) maintained by automatic and tangible reinforcement, and (12) maintained by attention, tangible reinforcement, and escape.

Percent agreement on interpretation was computed by dividing the number of exact agreements on function by the number of agreements and disagreements. Percent agreement between the two raters was 92.3%.

As the QABF includes a section on physical illness or discomfort not assessed by experimental functional analyses, each client's record was reviewed to determine if a physician identified a medical diagnosis (e.g., eczema, otitis media) as a contributing factor to the client's problem behavior. In each medical chart, the following sections were examined: active illnesses, major problems, minor problems, progress notes for the past year, and physician's orders. In no case did a physician identify a physical illness variable as active and contributing to the client's maladaptive behavior. Therefore, in the interpretation of data, physical illness was ruled out as a maintaining variable for the maladaptive behaviors of the 13 participants.

Primary functions as identified by the QABF were determined by the methodology described in the user's manual (Matson & Vollmer, 1995). A scoring profile (available with the measure) was completed for each QABF. If one subscale received a greater total score than the other four subscales, that subscale was labeled as the primary maintaining

variable. If two subscales had identical total scores, then both were identified as primary maintaining variables.

Primary functions as identified by the MAS were determined by following the methodology of Durand and Crimmins (1988). That is, the mean scores were computed for each subscale and rank ordered. The primary ranking was selected for convergent validity calculations; in the case of tie scores, both functions were identified as primary.

As the subscale scores on the QABF and MAS both yielded ordinal data, Spearman rank-order correlation coefficients were computed between the scores. As the interpretation of the analog sessions produced categorical data, Cramer's V, a nonparametric measure of association, was considered as a possible means of assessing the correlation between the analog sessions and the QABF and MAS. However, Cramer's V is a statistic derived from the chi square value for a set of variables; therefore, the conditions for chi square must be met (Siegel & Castellan, 1988). The present data did not meet all of the conditions in that more than 20% of the cells in the contingency tables for each pair of variables had expected frequencies less than 5. As a result, Cramer's V was not valid. Kappa also could not be computed as symmetric two-way tables with identical variables could not be produced. Therefore, only total percent agreement on function was calculated by dividing the number of exact agreements by the number of agreements plus disagreements.

In summary, several results were anticipated. First, the QABF would be stable over time and across informants. Next, past assessment of internal consistency and factor

structure could be replicated. Finally, correspondence between the QABF and analog sessions would be greater than that between the analog sessions and the MAS and the QABF and the MAS. Lower correlations were anticipated with the MAS for two reasons. First, the QABF items were constructed to be less ambiguous than those of previous checklists, making the QABF less likely to be redundant with the content of past instruments (Matson & Vollmer, 1995). Second, past research has demonstrated that the instrument is not psychometrically robust; therefore, it would seem unlikely that the MAS would correlate highly with analog functional analysis sessions.

Results

Reliability

Test-retest

To assess the stability of the individual QABF items over time, three measures of reliability were assessed. First, Spearman rank-order correlation coefficients were computed for each item as the item values were ordinal in nature (Fowler, 1988; Siegel & Castellan, 1988). Next, total agreement between items upon separate administrations was calculated through percent agreement (Sturmey, 1996). Finally, Cohen's kappa was computed when feasible for each item to correct for chance levels of agreement (Sattler, 1992). Kappa could not be computed when the contingency tables were not symmetric. That is, in some cases, the range of score values did not match between administrations (e.g., 0, 1, 3 vs. 0, 1, 2, and 3). Values for each statistic are provided below in Table 6.

Spearman rank-order correlations on the whole were high, with 76% of item statistics exceeding the minimally acceptable value of 0.8 (Nunnally, 1978; Sturmey, 1996). Correlations ranged from 0.646-1.000. Similarly, total percent agreement was good, with 96% of the items exceeding a minimum 80% (Rojahn & Schroeder, 1991). Total agreement ranged from 69.57-95.65%. The available kappa values also were adequate, ranging from 0.642-1.000, with 83% exceeding a minimum value of 0.7 (Sattler, 1992).

Table 6: Test-retest statistics for individual items (n=34)

Item	Spearman	Percent agreement	Kappa
1. Engages in the behavior to get attention.	.956**	91.18	–

2. Engages in the behavior to escape work or learning table 6 continued	.646**	88.24	.728*
situations.			
3. Engages in the behavior as a form of "self stimulation".	.830**	88.24	.812*
4. Engages in the behavior because he/she is in pain.	.998**	94.12	.887*
5. Engages in the behavior to get access to items such as preferred toys, food, or beverages.	.782**	82.35	.646*
6. Engages in the behavior because he/she likes to be reprimanded.	1.000**	100.00	1.000*
7. Engages in the behavior when asked to do something (get dressed, brush teeth, work, etc.).	.800**	91.18	.744*
8. Engages in the behavior even if he/she thinks no one is in the room.	.929**	88.24	.811*
9. Engages in the behavior more frequently when he/she is ill.	.893**	97.06	.910*
10. Engages in the behavior when you take something away from him/her.	.681**	64.71	--
11. Engages in the behavior to draw attention to him/herself.	.987**	91.18	.825*
12. Engages in the behavior when he/she does not want to do something.	.832**	88.24	.833*
13. Engages in the behavior because there is nothing else to do.	.947**	91.18	.867*
14. Engages in the behavior when there is something bothering him/her physically.	1.000**	100.00	1.000*
15. Engages in the behavior when you have something he/she wants.	.729**	82.35	.686*
16. Engages in the behavior to try to get a reaction from you.	.937**	94.12	.880*
17. Engages in the behavior to try to get people to leave him/her alone.	.849**	88.24	.751*
18. Engages in the behavior in a highly repetitive manner, ignoring his/her surroundings.	.871**	82.35	--
19. Engages in the behavior because he/she is physically uncomfortable.	1.000**	100.00	1.000*
20. Engages in the behavior when a peer has something he/she wants.	.875**	82.35	.642*
21. Does he/she seem to be saying "come see me" or "look at me" when engaging in the behavior?	.998**	97.06	--

table 6 continued

22. Does he/she seem to be saying "leave me alone" or "stop asking me to do this" when engaging in the behavior?	.774**	88.24	.775*
23. Does he/she seem to enjoy the behavior, even if no one is around?	.931**	88.24	--
24. Does the behavior seem to indicate to you that he/she is not feeling well?	1.000**	97.06	--
25. Does he/she seem to be saying "give me that (toy item, food item)" when engaging in the behavior?	.754**	88.24	--

*significant at the .001 level, **significant at the .01 level

To assess the stability over time of the subscales and test as a whole, Pearson product-moment correlation coefficients were computed. High reliability was obtained for each subscale and total score. Values for each subscale and total score are provided in Table 7.

Table 7: Test-retest subscale and total score statistics (n=34)

Subscale	Pearson correlation
Attention	.988*
Escape	.832*
Nonsocial	.934*
Physical	.990*
Tangible	.795*
TOTAL SCALE	.867*

* significant at the .01 level

Interrater reliability

Interrater reliability was calculated using the same statistics as for test-retest reliability. That is, Spearman rank-order correlation coefficients, total percent agreement, and Cohen's kappa were computed for individual items while Pearson product-moment correlation coefficients were obtained for subscale and total scores.

Table 8 provides reliability statistics for the individual items. Spearman rank order correlation coefficients were slightly lower than for test-retest, ranging from -0.095-1.000, with 52% of items exceeding a minimum of 0.8. Total agreement also was slightly lower on the whole, ranging from 69.57-95.65% with 56% of the items exceeded 80% agreement. Kappa values were adequate, ranging from .427-.921, with 41% exceeding a minimum of 0.7.

Table 8: Interrater statistics for individual items (n=23)

Item	Spearman	Percent agreement	Kappa
1. Engages in the behavior to get attention.	.811**	73.91	.596*
2. Engages in the behavior to escape work or learning situations.	.473***	73.91	--
3. Engages in the behavior as a form of "self stimulation".	.938**	86.96	--
4. Engages in the behavior because he/she is in pain.	.641**	78.26	--
5. Engages in the behavior to get access to items such as preferred toys, food, or beverages.	.732**	82.61	.738*
6. Engages in the behavior because he/she likes to be reprimanded.	-.095	82.61	--
7. Engages in the behavior when asked to do something (get dressed, brush teeth, work, etc.).	.954**	82.61	.733*

table 8 continued

8. Engages in the behavior even when he/she thinks no one is in the room.	.931**	82.61	.749*
9. Engages in the behavior more frequently when he/she is ill.	.869**	86.96	.524*
10. Engages in the behavior when you take something away from him/her.	.893**	73.91	.613*
11. Engages in the behavior to draw attention to him/herself.	.790**	78.26	.633*
12. Engages in the behavior when he/she does not want to do something.	.848**	82.61	--
13. Engages in the behavior because there is nothing else to do.	.947**	91.30	.826*
14. Engages in the behavior when there is something bothering him/her physically.	.942**	78.26	.579*
15. Engages in the behavior when you have something he/she wants.	.633**	82.61	.623*
16. Engages in the behavior to try to get a reaction from you.	.799**	69.57	.427**
17. Engages in the behavior to try to get people to leave him/her alone.	.716**	78.26	.672*
18. Engages in the behavior in a highly repetitive manner, ignoring his/her surroundings.	.988**	95.65	.921*
19. Engages in the behavior because he/she is physically uncomfortable.	.981**	86.96	.728*
20. Engages in the behavior when a peer has something he/she wants.	.744**	69.57	--
21. Does he/she seem to be saying "come see me" or "look at me" when engaging in the behavior?	.716**	82.61	--
22. Does he/she seem to be saying "leave me alone" or "stop asking me to do this" when engaging in the behavior?	.648**	69.57	.517*
23. Does he/she seem to enjoy the behavior, even if no one is around?	.923**	86.96	.777*
24. Does the behavior seem to indicate to you that he/she is not feeling well?	1.000**	95.65	--
25. Does he/she seem to be saying "give me that (toy item, food item)" when engaging in the behavior?	.664**	78.26	.637*

* significant at the .001 level, ** significant at the .01 level, ***significant at the .05 level

Interrater reliability for the subscale and total scores is provided in Table 9. Pearson product-moment correlation coefficients were acceptable for each subscale and the total score.

Table 9: Interrater subscale and total score statistics (n=23)

Subscale	Pearson correlation
Attention	.860*
Escape	.826*
Nonsocial	.987*
Physical	.945*
Tangible	.790*
TOTAL SCALE	.792*

*significant at the .01 level

Internal consistency

To assess internal consistency, coefficient alpha was computed using the larger sample (n=243) of participants. Coefficient alpha provides an estimate of reliability based on the average correlation between items as well as the total number of items. As such, it is an essential measure of reliability as it can reflect measurement error from inadequate sampling of content (Nunnally, 1978). Coefficient alpha for the present study was computed for the individual subscales and the test as a whole and is summarized in Table 10. Coefficient alpha was very high for each subscale, but lower for the test as a whole.

Table 10: Internal consistency (n=243)

Subscale	Alpha
Attention	.9004
Escape	.9266
Nonsocial	.9297
Physical	.9138
Tangible	.9078
TOTAL SCALE	.6010

Item	Alpha if item deleted
1. Engages in the behavior to get attention.	.5697
2. Engages in the behavior to escape work or learning situations.	.5799
3. Engages in the behavior as a form of "self stimulation".	.6306
4. Engages in the behavior because he/she is in pain.	.5844
5. Engages in the behavior to get access to items such as preferred toys, food, or beverages.	.5741
6. Engages in the behavior because he/she likes to be reprimanded.	.5961
7. Engages in the behavior when asked to do something (get dressed, brush teeth, work, etc.).	.5773
8. Engages in the behavior even if he/she thinks no one is in the room.	.6454
9. Engages in the behavior more frequently when he/she is ill.	.5840
10. Engages in the behavior when you take something away from him/her.	.5685

table 10 continued

11. Engages in the behavior to draw attention to him/herself.	.5679
12. Engages in the behavior when he/she does not want to do something.	.5712
13. Engages in the behavior because there is nothing else to do.	.6054
14. Engages in the behavior when there is something bothering him/her physically.	.5762
15. Engages in the behavior when you have something he/she wants.	.5676
16. Engages in the behavior to try to get a reaction from you.	.5863
17. Engages in the behavior to try to get people to leave him/her alone.	.5872
18. Engages in the behavior in a highly repetitive manner, ignoring his/her surroundings.	.6370
19. Engages in the behavior because he/she is physically uncomfortable.	.5850
20. Engages in the behavior when a peer has something he/she wants.	.5762
21. Does he/she seem to be saying "come see me" or "look at me" when engaging in the behavior?	.5819
22. Does he/she seem to be saying "leave me alone" or "stop asking me to do this" when engaging in the behavior?	.5806
23. Does he/she seem to enjoy the behavior, even if no one is around?	.6482
24. Does the behavior seem to indicate to you that he/she is not feeling well?	.5961
25. Does he/she seem to be saying "give me that (toy item, food item) when engaging in the behavior?	.5757

To further assess internal consistency, the Spearman-Brown statistic was computed to assess degree of consistency between halves of the test. The Spearman-Brown correlation coefficient corrected for unequal length was .6002.

The lower values for alpha for the total score and for the Spearman-Brown correlation coefficient were anticipated as the QABF was not designed to measure a homogeneous construct, but rather a heterogeneous grouping of 5 potential functions of problem behavior.

Validity

Factor Analysis

To confirm the heterogeneity of the QABF, an exploratory factor analysis was conducted using data from the 243 participants included in the analysis of internal consistency. Principal Axis Factoring (PAF) with varimax rotation yielded 5 factors whose labels corresponded to the subscales of the QABF, replicating the results of Matson et al. (1996). These factors accounted for 76.1% of the variance shared by the test items. Table 11 outlines the factor structure and item loadings.

Table 11: Factor structure (n=243)

Factor 1: Escape

(28.6% of variance)

Item	Loading
2. Engages in the behavior to escape work or learning situations.	.8534
7. Engages in the behavior when asked to do something (get dressed, brush teeth, work, etc.).	.8395
12. Engages in the behavior when he/she does not want to do something.	.8856
17. Engages in the behavior to try to get people to leave him/her alone.	.8100
22. Does he/she seem to be saying "leave me alone" or "stop asking me to do this" when engaging in the behavior?	.8884

table 11 continued

Factor 2: Tangible

(15.2% of variance)

Item	Loading
5. Engages in the behavior to get access to items such as preferred toys, food, or beverages.	.8229
10. Engages in the behavior when you take something away from him/her.	.7734
15. Engages in the behavior when you have something he/she wants.	.8879
20. Engages in the behavior when a peer has something he/she wants.	.7890
25. Does he/she seem to be saying "give me that (toy item, food item)" when engaging in the behavior?	.9018

Factor 3: Physical

(14.0% of variance)

Item	Loading
4. Engages in the behavior because he/she is in pain.	.8876
9. Engages in the behavior more frequently when he/she is ill.	.8352
14. Engages in the behavior when there is something bothering him/her physically.	.8514
19. Engages in the behavior because he/she is physically uncomfortable.	.8824
24. Does the behavior seem to indicate to you that he/she is not feeling well?	.8601

Factor 4: Nonsocial

(10.8% of variance)

table 11 continued

Item	Loading
3. Engages in the behavior as a form of "self stimulation".	.8491
8. Engages in the behavior even if he/she thinks no one is in the room.	.8065
13. Engages in the behavior because there is nothing else to do.	.8290
18. Engages in the behavior in a highly repetitive manner, ignoring his/her surroundings.	.8206
23. Does he/she seem to enjoy the behavior, even if no one is around?	.8177

Factor 5: Attention

(7.6% of variance)

Item	Loading
1. Engages in the behavior to get attention.	.8637
6. Engages in the behavior because he/she likes to be reprimanded.	.6204
11. Engages in the behavior to draw attention to him/herself.	.8993
16. Engages in the behavior to try to get a reaction from you.	.8651
21. Does he/she seem to be saying "come see me" or "look at me" when engaging in the behavior?	.8824

Table 12: Factor loadings of items across factors

Item	Factor 1:	Factor 2:	Factor 3:	Factor 4:	Factor 5:
1	.0524	.2290	.0537	-.1641	.8637
2	.8534	.0658	.0525	-.1900	.0185
3	-.1945	-.1307	-.1065	.8491	-.1290

table 12 continued

4	.0287	.0531	.8876	-.0431	.0560
5	.0249	.8229	.0256	-.2012	.1295
6	-.0020	-.0727	.0167	-.0416	.6204
7	.8395	.0742	.0716	-.1388	.0213
8	-.3020	-.2187	-.0443	.8065	-.1599
9	.0848	-.0452	.8352	-.0427	.1190
10	.2023	.7734	.0236	-.0905	.0485
11	.0544	.2028	.0756	-.1539	.8993
12	.8856	.1328	.0822	-.1593	-.0017
13	-.0351	-.0678	-.0111	.8290	-.0324
14	.1098	.1267	.8514	.0361	.0360
15	.0456	.8879	.0290	-.0913	.0890
16	.0019	.0491	-.0055	-.1381	.8651
17	.8100	.0009	.0398	-.1880	.0165
18	-.2401	-.2028	-.0079	.8206	-.1842
19	.0317	.0362	.8824	.0285	-.0062
20	.0346	.7890	.0014	-.0997	.0867
21	-.0660	.1056	.0326	-.1042	.8824
22	.8884	.0490	.0363	-.1438	-.0359
23	-.2956	-.2007	-.0348	.8177	-.2421

table 12 continued

24	.0144	-.1106	.8601	-.1604	-.0469
25	.0155	.9018	-.0167	-.1886	.0576

Convergent validity

Analog functional analysis sessions were completed for 13 participants. In each case, the conditions were run in a multi-element format. For most participants, this format yielded definitive results. However, for participant 8, extended alone sessions were conducted when the multi-element format proved inconclusive.

Participant 11 was included in this study although the majority of his functional analysis was completed in 1997. Two factors led to his inclusion: (1) no consistent intervention had been implemented for his aggression at the time of this study, and (2) past research demonstrated the stability of problem behaviors over time in adults with mental retardation living in institutions (Reid, Ballinger, Heather, & Melvin, 1984). Therefore, additional functional analysis sessions were conducted which replicated the results of the earlier sessions. The QABF and MAS also were administered at this time.

Figures 1-4 display the functional analysis results for participants 1-13 grouped by functional analysis outcome. In Figures 1 and 2, the outcomes for three participants with an escape function are displayed. That is, for each participant, escape from demands proved to be the primary maintaining variable for their problem behaviors according to the

interpretation criteria of Hagopian et al. (1997). Similarly, in Figures 3 and 4, the outcomes for four participants whose behavior had an automatic function are displayed. Figures 5 and 6 show three participants who had more than one operant function to their behavior, while the graphs in Figures 7 and 8 are undifferentiated.

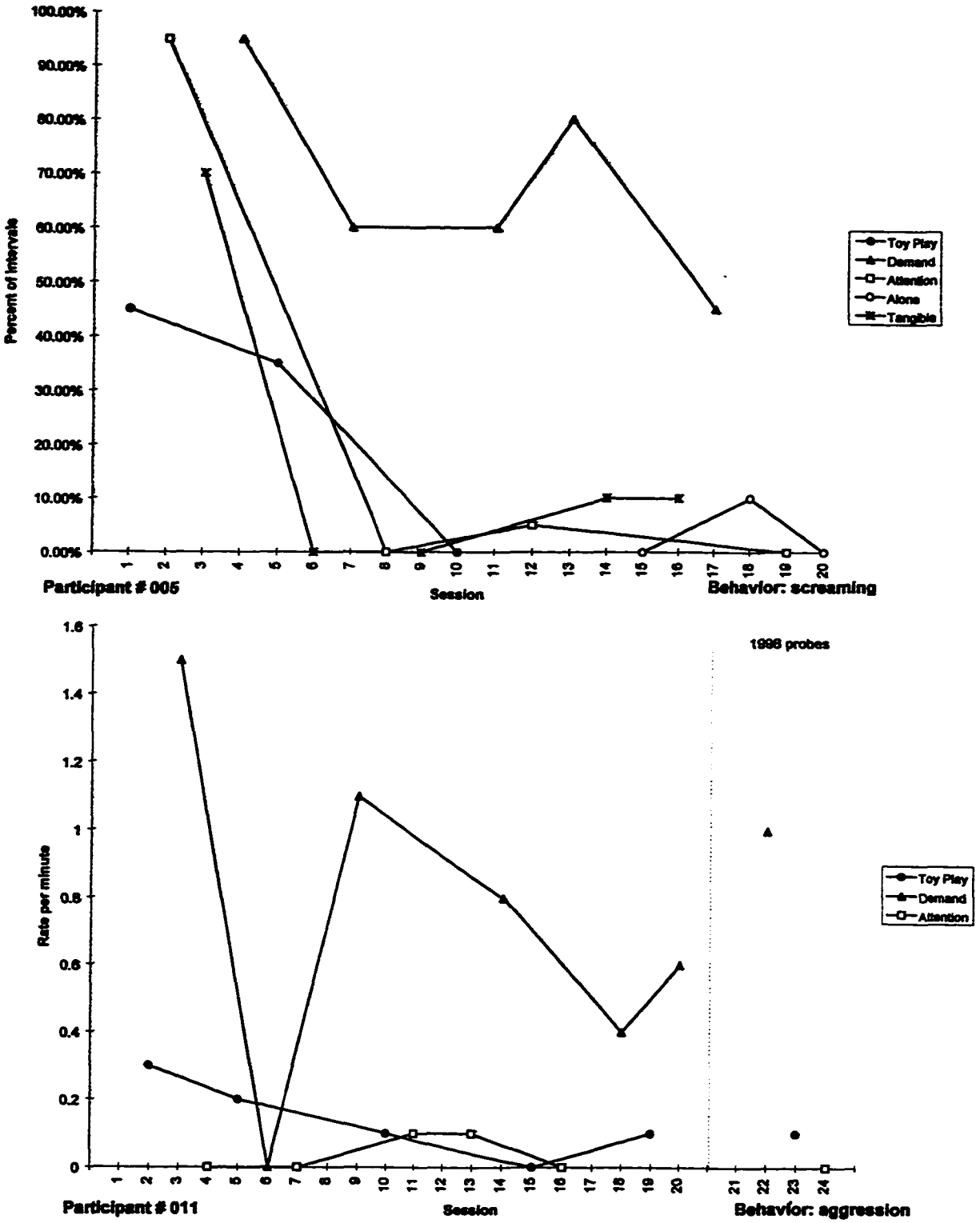


Figure 1: Single function analyses: Escape function (participants 5, 11)

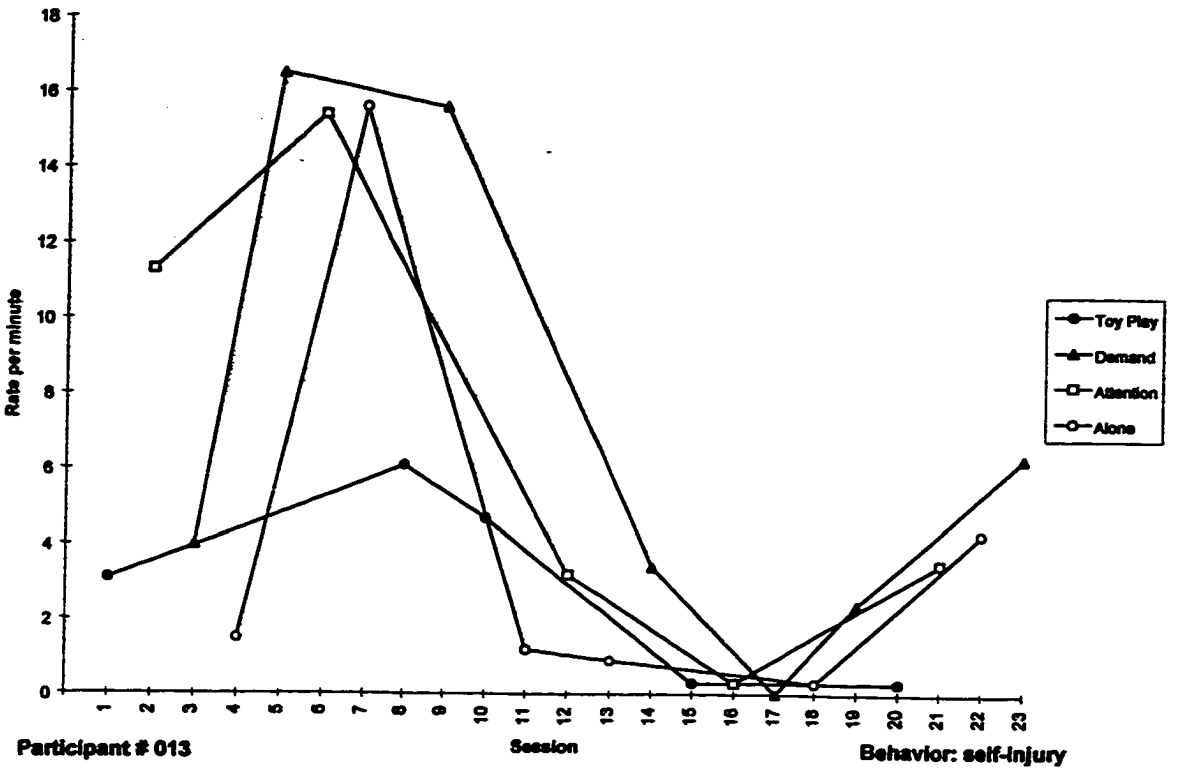


Figure 2: Single function analyses: Escape function (participant 13)

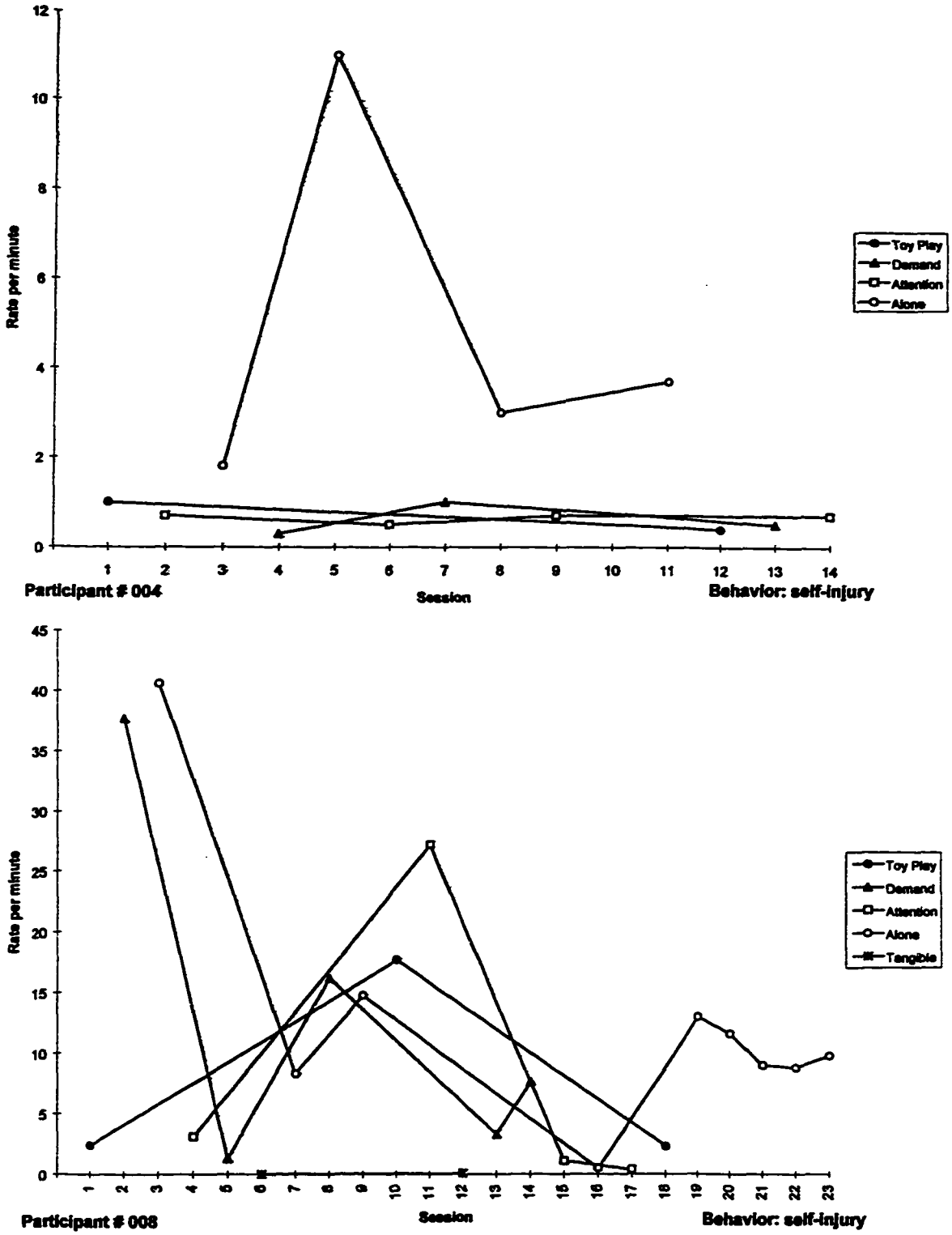


Figure 3: Single function analyses: Automatic function (participants 4, 8)

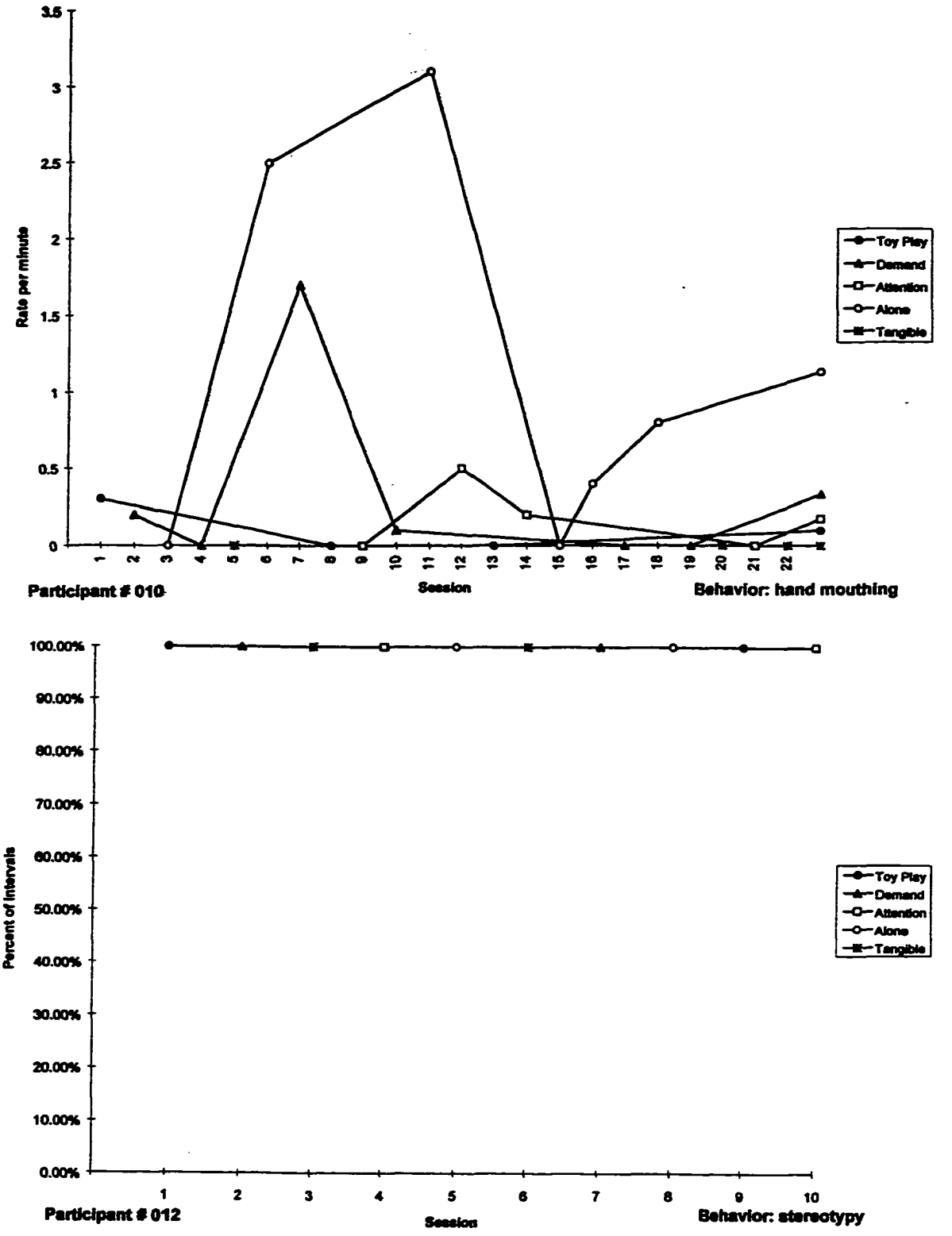


Figure 4: Single function analyses: Automatic function (participants 10, 12)

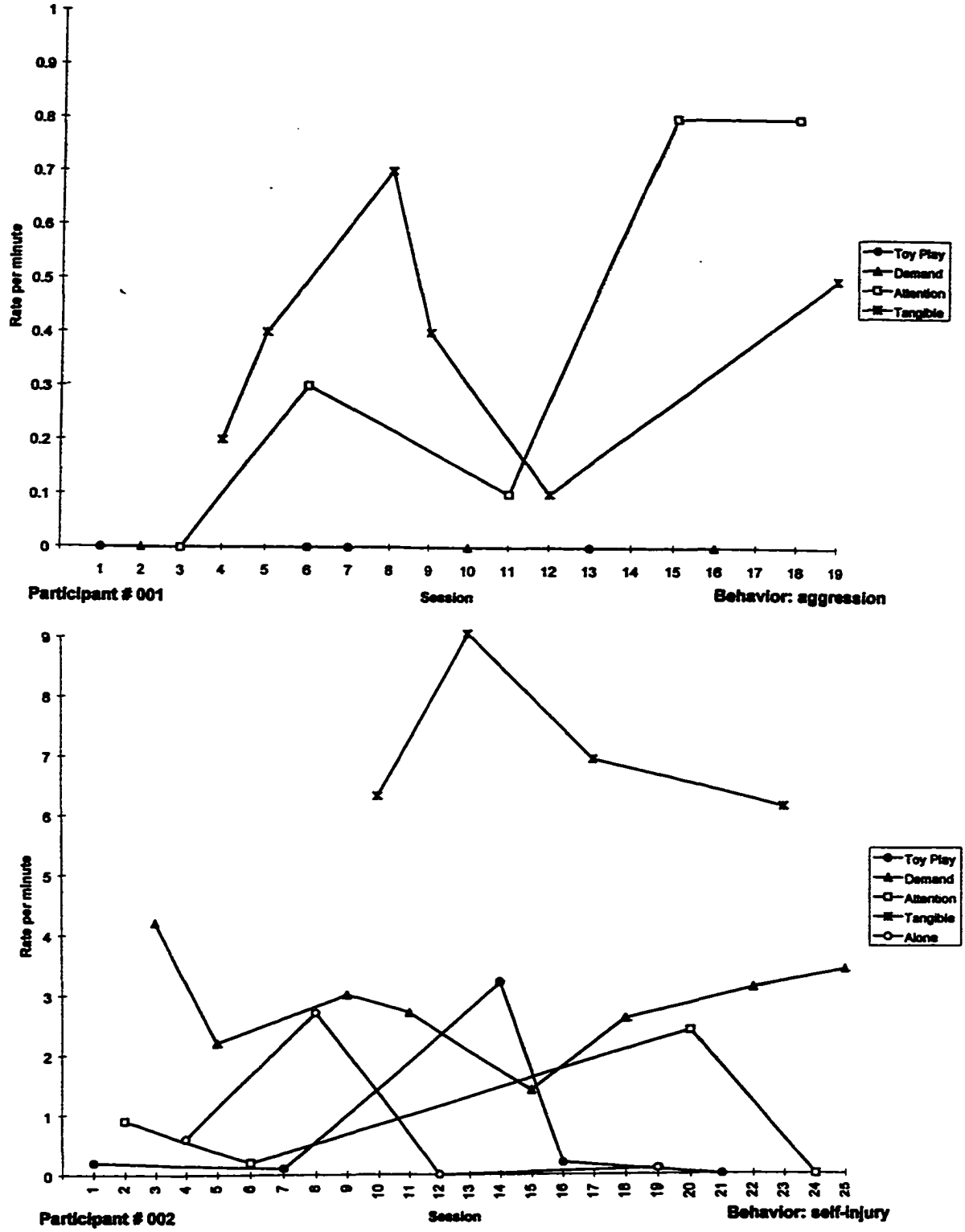


Figure 5: Multiply-controlled functional analyses (participants 1, 2)

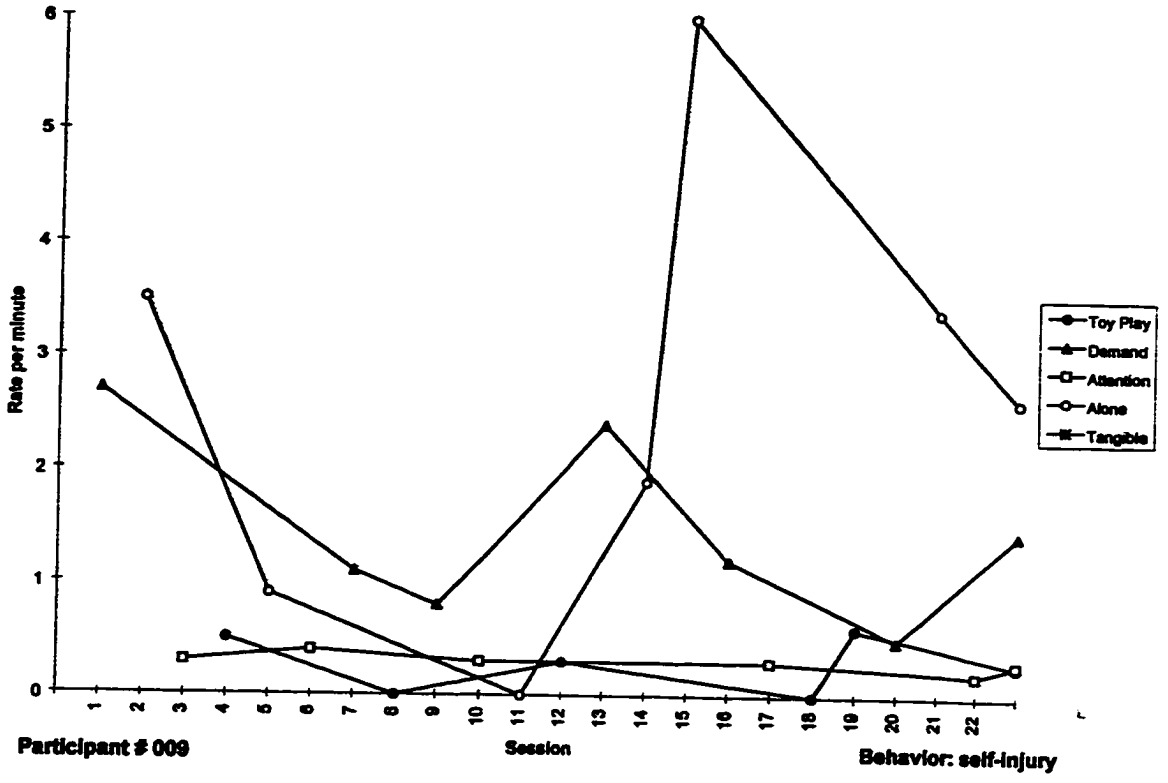


Figure 6: Multiply-controlled functional analyses (participant 9)

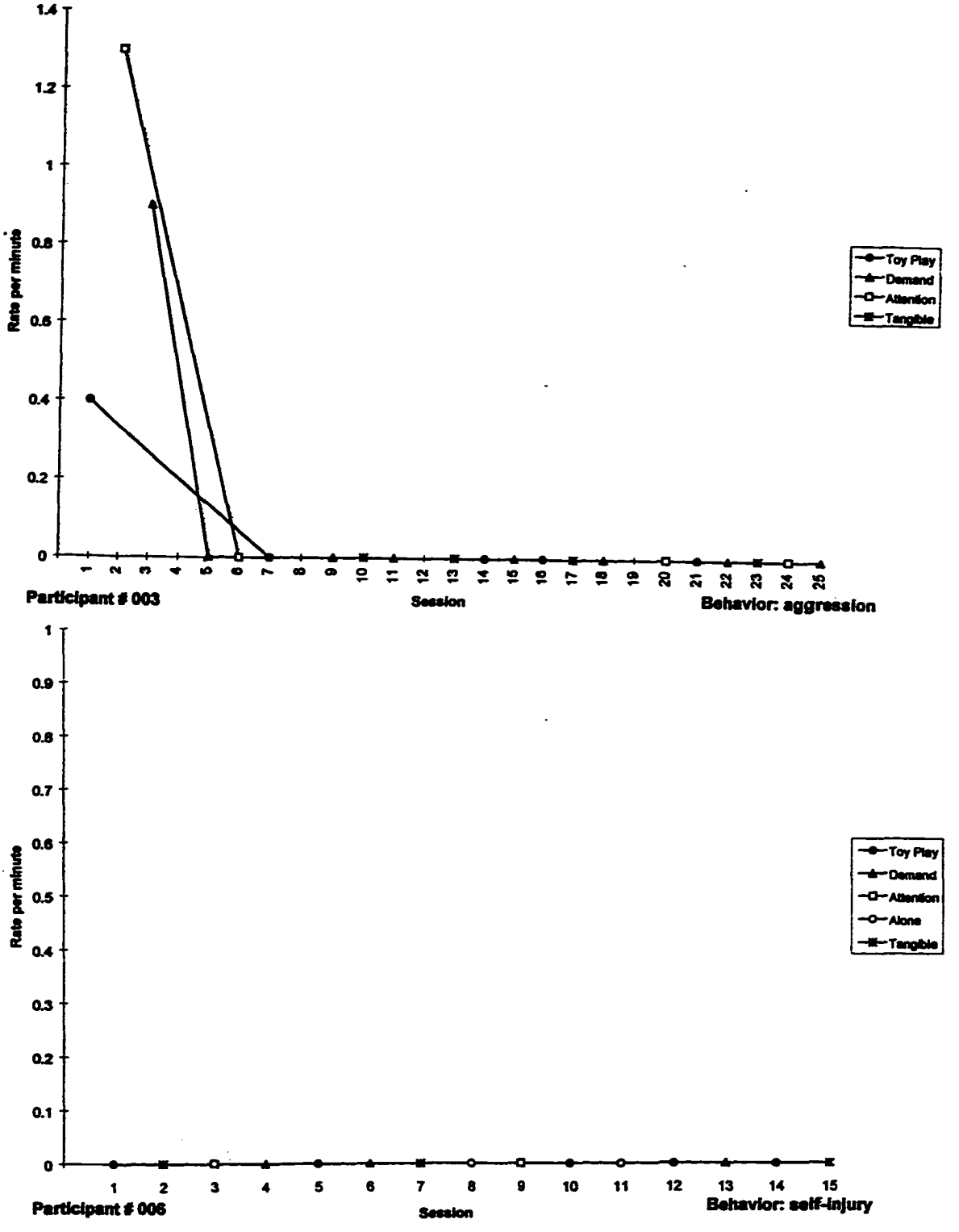


Figure 7: Undifferentiated functional analyses (participants 3, 6)

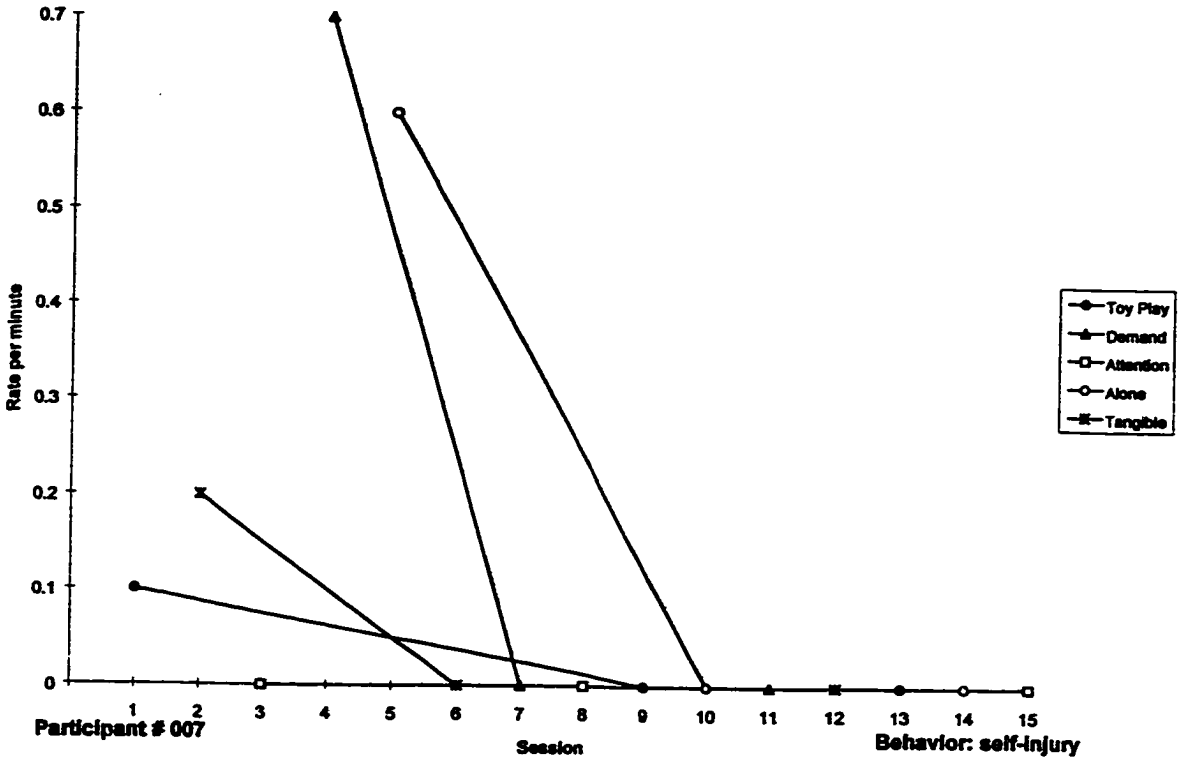


Figure 8: Undifferentiated functional analyses (participant 7)

The graphs for participants 3, 6, and 7 were labeled as undifferentiated. In each case, there were insufficient data to suggest a primary function. However, a review of behavioral records for the three participants from the three months prior to the functional analyses indicated that participant 3 displayed 0 incidents of aggression during that time, participant 6 exhibited an average of 0.9 incidents of aggression per day, and participant 7 displayed an average of 0.03 self-injurious behaviors per day. Therefore, for these three participants, the target behaviors were of sufficient intensity to warrant referral for a complete functional analysis, but the behaviors were too infrequent to be assessed by this method.

Table 13 compares the functions of the target behavior for each participant as identified by the analog sessions, the QABF, and the MAS.

Table 13: Identified behavioral functions

Participant	Function(s) from analog sessions	Function(s) from MAS	Function(s) from QABF
1	attention, tangible	attention, tangible	attention, tangible
2	tangible, escape	tangible	tangible, escape
3	undifferentiated	attention	attention
4	automatic	automatic	automatic
5	escape	automatic	automatic
6	undifferentiated	tangible	automatic, tangible
7	undifferentiated	automatic	automatic

table 13 continued

8	automatic	escape	automatic
9	automatic, escape	escape	escape
10	automatic	automatic	escape
11	escape	escape	escape
12	automatic	escape	escape
13	escape	automatic	escape

For the 13 participants, the QABF and the analog sessions agreed in 46.2% of cases, the MAS and analog sessions in 30.8% of cases, and the QABF and MAS in 61.5% of cases.

Moderate correlations were observed between the QABF and the MAS. In some cases, parallel subscales were correlated; in others (e.g., QABF Nonsocial, MAS Tangible) conceptually different subscales appeared to be related. Table 13 outlines the Spearman rank-order correlation coefficients between the MAS and the QABF.

Table 14: Spearman correlations between the MAS and the QABF

	MAS-Sensory	MAS-Escape	MAS-Attention	MAS-Tangible
QABF-Attention	.247	-.278	.512	.410
QABF-Escape	.466	.508	-.135	.126
QABF-Nonsocial	.794**	.038	.269	.657*
QABF-Physical	.796**	.049	.338	.815**
QABF-Tangible	.540	-.192	.435	.857**

table 14 continued

*significant at the .01 level, **significant at the .001 level

These correlations, however, need to be interpreted in the context of a small sample size. With an *n* of 13, a correlation would have to exceed approximately 0.475 to achieve statistical significance at the .05 level (Glass & Hopkins, 1996).

Discussion

The present study examined five psychometric properties of the QABF: test-retest reliability, interrater reliability, internal consistency, factor analysis, and concurrent validity. Reliability data were sound, as was a factor solution that corresponded to the clinical subscales of the QABF. Convergent validity data, however, require further assessment. Based on these results, the following conclusions can be drawn.

External Reliability

First, the QABF has good stability over time. Teat-retest reliability for the individual items and the subscale scores was high across three measures of agreement: Spearman rank-order correlations, total percent agreement, and Cohen's kappa. Therefore, raters' impressions remained consistent between two administrations 1-3 weeks apart.

Similarly, different raters typically agreed in their impressions of a client's behavior. Interrater reliability was generally high although slightly lower than that for test-retest reliability.

Each subscale of the QABF had consistently high external reliability. This is in contrast to the variable results observed in studies of the reliability of the MAS, the most widely available functional assessment checklist (e.g., Crawford et al., 1992; Newton & Sturmey, 1991; Sigafoos, Kerr, & Roberts, 1994; Zarcone et al., 1991). These studies citing poor reliability suggested that the use of untrained staff as informants, the inclusion of participants with variable rates of behaviors, and the inclusion of participants with behaviors other than self-injury resulted in lower reliability values than desirable.

However, in the present study, informants were direct care staff with minimal training in behavior analysis, and the participants displayed a variety of behaviors as well as various rates of behaviors. The participants did, however, all reside in a state institution.

The present results for the external reliability of the QABF are promising. If the high values for test-retest and interrater reliability can be replicated in the institutional setting, then the QABF may be a reliable instrument with this population. Additional replication with other populations (e.g., school-age children with developmental disabilities) would be needed to generalize the results across the population of individuals with developmental disabilities.

Internal Reliability and Factor Analysis

Coefficient alpha values for the individual subscales were high, suggesting that the subscales were homogeneous in content. These values replicated the initial results of Matson et al. (1996). Split-half reliability was lower than expected but may have been due to the small number of items that resulted when the test was divided into halves. Internal consistency for the scale as a whole was somewhat low but expected as the scale was constructed to tap five unrelated variables.

Subsequent factor analysis replicated the original 5-factor solution identified by Matson et al. (1996). Each item loaded onto a factor corresponding to a hypothesized subscale, and a considerable proportion of the variance shared by test items was accounted for (76.1%).

As a whole, the internal consistency and factor structure demonstrate that the QABF consists of five underlying factors that are statistically significant and clinically meaningful. These factors correspond to the five potential functions of problem behavior as assessed by the QABF (attention, escape, nonsocial, physical, and tangible), and are supported by high internal consistency among the items of each subscale.

As with the assessment of external reliability, the participants for the factor analysis and assessment of internal consistency displayed a range of target behaviors beyond self-injury alone. Despite this variability, the factor structure remained consistent in this and the previously studied sample (Matson et al., 1996) of participants residing in a state institution. Therefore, the QABF appears to have statistically significant and clinically meaningful factors when assessing the functions of problem behaviors in individuals residing in institutions.

Convergent Validity

Convergent validity was assessed through comparison of the QABF to analog functional analysis sessions. In addition, to determine the degree of validity of the QABF in relation to the MAS, comparisons also were made between the MAS and analog sessions and the QABF and the MAS. The assessment of convergent validity resulted in a lower agreement between the QABF and the analog functional analysis sessions (46.2%) than would be ideal. Agreement between the MAS and the analog sessions was still lower (30.8%), while agreement between the QABF and the MAS was the highest (61.5%). It would appear that the two checklists tap similar content domains, but do not correspond

well to analog sessions. These results are discussed in the context of limitations of the small sample size, inclusion of subjects with undifferentiated functional analysis results, methodology for conducting convergent validity studies in this area, and implications for use of the QABF.

It is difficult to determine the accuracy in the level of agreement between the QABF and the other measures due to the small sample. With an *n* of 13, it was not possible to compute Cramer's V, a correlation coefficient for nominal data (Siegel & Castellan, 1988) as a chi square value could not be calculated with this sample size. Similarly, the Spearman rank-order correlation coefficients between the QABF and the MAS are difficult to interpret as for an *n* of 13, a correlation would have to exceed a minimum value of 0.475 in order to be significant at the .05 level. The small sample size also widens the 95% confidence interval for these values as smaller sample sizes tend to have confidence intervals with a large range which reduces the generalizability of the results (Glass & Hopkins, 1996). Therefore, for more robust correlations between the QABF and other functional assessment measures, a larger sample size is required.

Another factor affecting the current results is the inclusion of subjects with undifferentiated functional analysis outcomes who were determined to have low frequency behaviors. Low frequency but high intensity behaviors are underinvestigated in current functional analysis research (Sturmey, 1995). In this study, 23% of the participants fit such criteria. In each case, the functional analysis failed to capture a sufficient sample of the behavior to permit identification of a primary function. In fact, if the data from the

three participants with undifferentiated results were not included in the validity assessment, the percent agreement between the QABF and the analog sessions would increase to 60%, the percent agreement between the MAS and the analog sessions would increase to 40%, and the agreement between the QABF and the MAS would remain stable at 60%.

Therefore, the inclusion of participants with undifferentiated behavior has a significant impact on the convergent validity portion of this study. This issue was not addressed by Durand and Crimmins (1988) who conducted the only other study examining the convergent validity of a functional analysis checklist. In their validity assessment of the MAS, they pre-selected 8 participants for whom the MAS identified a single, primary function. Next, they conducted analog sessions and summarized the results by graphically displaying the mean percent of intervals in which the behavior occurred in each condition. Finally, they correlated the primary functions as identified by the MAS and by the bar graphs and obtained a correlation of 0.99.

Two significant limitations to this methodology preclude the acceptance of such a high validity coefficient. First, the participant sample only included those with clear single-function outcomes on the MAS, thereby not addressing such functional analysis outcomes as multiply-controlled behaviors (e.g., Derby et al., 1994) and undifferentiated results (e.g., Hagopian et al., 1997). The present study selected participants based on clinical need only, thereby potentially including the range of functional analysis outcomes. Given the discrepancy between the validity results of the present study and that of Durand and

Crimmins (1988), it may be that functional analysis checklists are most valid when the behavior is maintained by a single variable.

The second major limitation of the Durand and Crimmins (1988) study that may have resulted in an inflated validity coefficient is the use of bar graphs summarizing mean values in each condition. Best practice in single-subject research is to present all of the available data to allow other investigators to determine first-hand if trends in variability affect conclusions about the outcome. Intrasubject averaging carries considerable risk in that the variability of the behavior becomes masked (Barlow & Hersen, 1984). For example, in the case of participant 003, behaviors were initially observed during the functional analysis and then decrease to 0 for the remainder of the assessment. The downward trend resulted in the analysis being labeled "undifferentiated" according to the criteria of Hagopian et al. (1997). Averaging the data would have suggested an attention function for the behavior. It is not clear to what extent this issue of variability affected the results of Durand and Crimmins (1988).

In summary, in using more objective means of participant selection and interpretation of functional analysis data, the present study yielded convergent validity results that were lower than anticipated and were significantly affected by 3 participants with undifferentiated functional analysis outcomes. More specifically, these participants displayed behaviors that were low frequency and could not be adequately assessed by a formal functional analysis. However, both the QABF and the MAS suggested possible primary functions in these cases. Future research with the QABF could be conducted to

determine if the function(s) suggested by the QABF in these cases were valid. An alternative method of validity assessment would be to conduct a comparison of treatment outcomes between a treatment suggested by the results of the QABF and one that was contraindicated by the results. If only the suggested treatment resulted in a reduction of the target behavior, then the hypothesized behavioral function could be confirmed.

Conclusions

There is a definite need for more efficient methods of conducting functional assessments as extended functional analysis sessions are not feasible across clinical settings (Horner, 1994; Van Houten & Rolider, 1991). More efficient methods are required to meet the needs of service providers who do not have the resources available to conduct extended analog sessions. In the present study, the total time required to conduct the functional analyses was 43.0 hours (preparation time excluded), whereas the total time to administer the QABFs was 4.3 hours (assuming a 20-min administration time), a ten-fold difference. For clinicians with large caseloads, this represents a significant decrease in the amount of time needed to provide essential clinical intervention.

With the exception of the limitations to the convergent validity data, the QABF has sound psychometric data to support further investigation of its role in conducting functional assessments. The goal of such work would be to determine the limiting conditions of its use; that is, in what cases would it be a valid method of functional assessment and in what situations would the results call for further analysis with analog sessions. First, as previously discussed, the QABF may be useful in the functional

assessment of low frequency behavior. Next, there may be certain outcomes that are more likely to be valid and could lead a clinician directly to treatment. For example, a larger convergent validity study may demonstrate that single-function outcomes (e.g., attention, escape, tangible, nonsocial, or physical) have good validity and require no further assessment. Another possibility is that certain functions may yield more valid results than others (e.g., escape and nonsocial functions are more likely to be accurate than physical discomfort). Third, a criterion needs to be established to determine what the score cut-offs should be on the QABF for determining multiply-controlled behavior. In the present study, only exact agreement between subscale scores was used to assign more than one function to a behavior and in no case were the scores close enough to make the interpretation ambiguous. However, at the present time there are insufficient data to determine if a minor difference in subscale scores on the QABF (e.g., 1-2 points) warrants an interpretation of the results as multiply-controlled. Again, a large scale study is required to make such a conclusion. Such a study would require a sample size such that each of the 12 possible functional analysis outcome categories was represented in the analysis. In this manner, the limiting conditions for the use of the QABF could be determined. Although the QABF cannot substitute for analog functional analysis sessions in all cases, there is clear potential for it within a hierarchical model of functional assessment. That is, when the limiting conditions of its use are identified, it can be incorporated as a screening tool to determine whether further functional analysis is needed. The psychometric properties are sufficiently stable to begin such research.

References

American Psychological Association, Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training and dissemination of empirically-validated psychological treatments: Report and recommendations. The Clinical Psychologist, 48, 3-24.

Arndorfer, R. E., & Miltenberger, R. G. (1993). Functional assessment and treatment of challenging behavior: A review with implications for early childhood. Topics in Early Childhood Special Education, 13, 82-105.

Association for Behavior Analysis, Task Force on the Right to Effective Behavioral Treatment (1988). The right to effective behavioral treatment. Journal of Applied Behavior Analysis, 21, 381-384.

Axelrod, S. (1987). Functional and structural analyses of behavior: Approaches leading to reduced use of punishment procedures? Research in Developmental Disabilities, 8, 165-178.

Bachman, J. A. (1972). Self-injurious behavior: A behavioral analysis. Journal of Abnormal Psychology, 80, 211-224.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 1, 91-97.

Barlow, D. H., & Hersen, M. (1984). Single case experimental designs: Strategies for studying behavior change, 2nd ed. Boston: Allyn & Bacon.

Belfiore, P. J., & Dattilio, F. M. (1990). The behavior of self-injury: A brief review and analysis. Behavioral Disorders, 16, 23-31.

Berkson, G., & Mason, W. A. (1964). Stereotyped movements of mental defectives: IV. The effects of toys and the character of the acts. American Journal of Mental Deficiency, 68, 511-524.

Bihm, E. M., Kienlen, T. L., Ness, M. E., & Poindexter, A. R. (1991). Factor structure of the Motivation Assessment Scale for persons with mental retardation. Psychological Reports, 68, 1235-1238.

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

Birnbrauer, J. S. (1979). Applied behavior analysis, service, and the acquisition of knowledge. The Behavior Analyst, 2, 15-21.

Bucher, B., & Lovaas, O. (1968). Use of aversive stimulation in behavior modification. In M. Jones (Ed.). Miami symposium on the prediction of behavior, 1967: Aversive stimulation. Coral Gables, FL: University of Miami Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

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

Carr, E. G. (1994). Emerging themes in the functional analysis of problem behavior. Journal of Applied Behavior Analysis, 27, 393-399.

Carr, E. G., & Durand, V. M. (1985). Reducing problem behaviors through functional communication training. Journal of Applied Behavior Analysis, 18, 11-126.

Carr, E. G., Newsom, C. D., & Binkoff, J. A. (1980). Escape as a factor in the aggressive behavior of two retarded children. Journal of Applied Behavior Analysis, 13, 101-117.

Carr, E. G., Yarbrough, S. C., & Langdon, N. A. (1997). Effects of idiosyncratic stimulus variables on functional analysis outcomes. Journal of Applied Behavior Analysis, 30, 673-686.

Chapman, S., Fisher, W., Piazza, C. C., & Kurtz, P. F. (1993). Functional assessment and treatment of life-threatening drug ingestion in a dually diagnosed youth. Journal of Applied Behavior Analysis, 26, 255-256.

Chiesa, M. (1992). Radical behaviorism and scientific frameworks. American Psychologist, 47, 1287-1299.

Cohen, J. (1960). A coefficient for nominal scales. Educational and Psychological Measurement, 68, 409-412.

Cooper, L. J., Wacker, D. P., Sasso, G. M., Reimers, T. M., & Donn, L. K. (1990). Using parents as therapists to evaluate appropriate behaviors of their children: Application to a tertiary diagnostic clinic. Journal of Applied Behavior Analysis, 23, 285-296.

Corte, H. E., Wolf, M. M., & Locke, B. J. (1971). A comparison of procedures for eliminating self-injurious behavior of retarded adolescents. Journal of Applied Behavior Analysis, 4, 201-213.

Cramer, D. (1994). Introducing statistics for social research. London: Routledge.

Crawford, J., Brockel, B., Schauss, S., & Miltenberger, R. G. (1992). A comparison of methods for the functional assessment of stereotypic behavior. Journal of the Association for Persons with Severe Handicaps, 17, 77-86.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Orlando, FL: Harcourt Brace Jovanovich.

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

Delprato, D. J., & Midgley, B. D. (1992). Some fundamentals of B. F. Skinner's behaviorism. American Psychologist, 47, 1507-1520.

Derby, K. M., Wacker, D. P., Peck, S., Sasso, G., DeRaad, A., Berg, W., Asmus, J., & Ulrich, S. (1994). Functional analysis of separate topographies of aberrant behavior. Journal of Applied Behavior Analysis, 27, 267-278.

Didden, R., Duker, P. C., & Korzilius, H. (1997). Meta-analytic study on treatment effectiveness for problem behaviors with individuals who have mental retardation. American Journal on Mental Retardation, 101, 387-399.

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

Durand, V. M., Crimmins, D. B., Caulfield, M., & Taylor, J. (1989). Reinforcer assessment I: Using problem behavior to select reinforcers. Journal of the Association for Persons with Severe Handicaps, 14, 113-126.

Favell, J. E. (1989). Environmental foundations of destructive behavior. In Treatment of destructive behaviors in persons with developmental disabilities (NIH Publication No. 91-2410, pp. 38-41). Bethesda, MD: Author.

Fowler, F. J. (1988). Survey research methods. Newbury Park, CA: Sage Publications.

Frea, W. D., & Hughes, C. (1997). Functional analysis and treatment of social-communicative behavior of adolescents with developmental disabilities. Journal of Applied Behavior Analysis, 30, 701-704.

Fuller, P. R. (1949). Operant conditioning of a vegetative human organism. American Journal of Psychology, 62, 587-590.

Glass, G. V., & Hopkins, K. D. (1996). Statistical methods in education and psychology, 3rd ed. Boston: Allyn & Bacon.

Goh, H. L., Iwata, B. A., Shore, B. A., DeLeon, I. G., Lerman, D. C., Ulrich, S. M., & Smith, R. G. (1995). An analysis of the reinforcing properties of hand mouthing. Journal of Applied Behavior Analysis, 28, 269-284.

Hagopian, L. P., Fisher, W. W., Thompson, R. H., Owen-DeSchryver, J., Iwata, B. A., & Wacker, D. P. (1997). Toward the development of structured criteria for interpretation of functional analysis data. Journal of Applied Behavior Analysis, 30, 313-325.

Hamilton, J., Stephens, L., & Allen, P. (1967). Controlling aggressive and destructive behavior in severely retarded institutionalized residents. American Journal of Mental Deficiency, 71, 852-856.

Hayes, S. C., Rincover, A., & Solnick, J. V. (1980). The technical drift of applied behavior analysis. Journal of Applied Behavior Analysis, 13, 275-285.

Hiles, M. G., & Desrochers, M. N. (1993). The relationship between functional assessment and treatment selection for aggressive behaviors. Research in Developmental Disabilities, 14, 265-274.

Horner, R. H. (1989). Recommendations based on community support for people with developmental disabilities and destructive behavior. In Treatment of destructive behaviors in persons with developmental disabilities (NIH Publication No. 91-2410, pp. 64-69). Bethesda, MD: Author.

Horner, R. H. (1994). Functional assessment: Contributions and future directions. Journal of Applied Behavior Analysis, 27, 401-404.

Horner, R. H., Sprague, J. R., & Flannery, K. B. (1993). Building functional curricula for students with severe intellectual disabilities and severe problem behaviors. In R. Van Houten & S. Axelrod (Eds.). Behavior analysis and treatment. New York: Plenum Press. (pp. 47-71).

Ivanick, C., Helsinger, S., & Harris, T. (1996). The functional assessment of autistic symptoms. Poster presented at the 22nd Annual Convention of the Association for Behavior Analysis, San Francisco, CA.

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

Iwata, B. A., Duncan, B. A., Zarcone, J. R., Lerman, D. C., & Shore, B. A. (1994). A sequential, test-control methodology for conducting functional analyses of self-injurious behavior. Behavior Modification, 18, 289-306.

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

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

Jacobson, J. W. (1982). Problem behavior and psychiatric impairment within a developmentally disabled population: Behavior frequency. Applied Research in Mental Retardation, 3, 121-139.

Johnson, W. L., & Baumeister, A. A. (1978). Self-injurious behavior. Behavior Modification, 2, 465-487.

Jones, F. H., Simmons, J. Q., & Frankel, F. (1974). An extinction procedure for eliminating self-destructive behavior in a 9-year old autistic girl. Journal of Autism and Childhood Schizophrenia, 4, 214-250.

Kanfer, F. H., & Phillips, J. S. (1970). Learning foundations of behavior therapy. New York: Wiley.

Kazdin, A. E. (1980). Behavior modification in applied settings. Homewood, IL: Dorsey.

Kern, L., Mauk, J. E., Marder, T. J., & Mace, F. C. (1995). Functional analysis and intervention for breath holding. Journal of Applied Behavior Analysis, 28, 339-340.

Keys, C. B., Fletcher, R. J., Holmes, P. A., & Schloss, P. J. (1989). Bridging the two cultures of dual diagnosis: Research and service. Research in Developmental Disabilities, 10, 277-284.

Lalli, J. S., Livezey, K., & Kates, K. (1996). Functional analysis and treatment of eye-poking with response blocking. Journal of Applied Behavior Analysis, 29, 129-132.

Laties, V. G., & Mace, F. C. (1993). Taking stock: The first 25 years of the Journal of Applied Behavior Analysis. Journal of Applied Behavior Analysis, 26, 513-525.

Lennox, D. B., & Miltenberger, R. G. (1989). Conducting a functional assessment of problem behavior in applied settings. Journal of the Association for Persons with Severe Handicaps, 14, 304-311.

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

Linscheid, T. R., Iwata, B. A., & Foxx, R. M. (1996). Behavioral assessment. In J. W. Jacobson and J. A. Mulick (Eds.). Manual of diagnosis and professional practice in mental retardation. Washington DC: American Psychological Association.

Lockwood, K., Borden, D. L., & Williams, D. E. (1996). Functional analysis and treatment of life-threatening pica with differential reinforcement and response interruption. Poster presented at the 22nd Annual Convention of the Association for Behavior Analysis, San Francisco, CA.

Lovaas, O. I., Freitag, G., Gold, V. J., & Kassorla, I .C. (1965). Experimental studies in childhood schizophrenia: Analysis of self-destructive behavior. Journal of Experimental Child Psychology, 2, 67-84.

Lovaas, O. I., & Simmons, J. Q. (1969). Manipulation of self-destruction in three retarded children. Journal of Applied Behavior Analysis, 2, 143-157.

Lowry, M., & Sovner, R. (1991). The functional significance of problem behavior: A key to effective treatment. The Habilitative Mental Healthcare Newsletter, 10, 59-62.

Mace, F. C. (1994). The significance and future of functional analysis methodologies. Journal of Applied Behavior Analysis, 27, 385-392.

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

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

Mace, F. C., Lalli, J. S., Pinter-Lalli, E., & Shea, M. C. (1993). Functional analysis and treatment of aberrant behavior. In R. Van Houten & S. Axelrod (Eds.). Behavior analysis and treatment. New York: Plenum Press. (pp. 75-99).

Maddle, R. A., & Neisworth, J. T. (1990). Mental retardation. In A. S. Bellack, M. Hersen, & A. E. Kazdin (Eds.). International handbook of behavior modification and therapy (2nd ed.). New York: Plenum Press. (pp. 731-762).

Maenpaa, M., Lockwood, K., & Williams, D. E. (1996). Functional analysis and treatment of meal refusal and severe weight loss. Poster presented at the 22nd Annual Convention of the Association for Behavior Analysis, San Francisco, CA.

Malott, R. W., Whaley, D. L., & Malott, M. E. (1993). Elementary principles of behavior. (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Martin, G., & Pear, J. (1996). Behavior modification: What it is and how to do it. (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Matson, J. L., & Coe, D. A. (1992). Applied behavior analysis: Its impact on the treatment of mentally retarded emotionally disturbed people. Research in Developmental Disabilities, 13, 171-189.

Matson, J. L., & Mulick, J. A. (1991). Handbook of Mental Retardation (2nd edition). New York: Pergamon Press.

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

Matson, J., Vollmer, T., Paclawskyj, T., Smiroldo, B., Applegate, H., & Stallings, S. (1996). Questions About Behavioral Function (QABF): An instrument to assess functional properties of problem behaviors. Poster presented at the 22nd Annual Convention of the Association for Behavior Analysis, San Francisco, CA.

May, J. G., Risley, T. R., Twardosz, S., Friedman, P., Bijou, S., Wexler, D. et al. (1975). Guidelines for the use of behavioral procedures in state programs for the retarded. NARC Monograph. Mental Retardation Research, 1.

McGonigle, J. J., Rojahn, J., Dixon, J., & Strain, P. S. (1987). Multiple treatment interference in the alternating treatments design as a function of the intercomponent interval length. Journal of Applied Behavior Analysis, 20, 171-178.

Measel, C. J., & Alfieri, P. A. (1976). Treatment of self-injurious behavior by a combination of reinforcement for incompatible behavior and overcorrection. American Journal of Mental Deficiency, 81, 147-153.

Miltenberger, R. G., & Fuqua, R. W. (1985). Evaluation of a training manual for the acquisition of behavioral assessment interviewing skills. Journal of Applied Behavior Analysis, 18,
323-328.

Morris, E. K. (1991). Deconstructing "technological to a fault". Journal of Applied Behavior Analysis, 24, 411-416.

Mosely, A., Faust, M., & Reardon, D. M. (1970). Effects of social and nonsocial stimuli on the stereotyped behavior of retarded children. American Journal of Mental Deficiency, 74, 809-811.

Mulick, J. A., Hammer, D., & Dura, J. R. (1991). Assessment and management of antisocial and hyperactive behavior. In J. L. Matson & J. A. Mulick (Eds.), Handbook of mental retardation (2nd ed., pp. 397-412). New York: Pergamon Press.

Munk, D. D., & Repp, A. C. (1994). Behavioral assessment of feeding problems of individuals with severe disabilities. Journal of Applied Behavior Analysis, 27, 241-250.

Myers, D. V. (1975). Extinction, DRO, and response cost procedures for eliminating self-injurious behavior: A case study. Behavior Research and Therapy, 13, 189-191.

National Institutes of Health, Consensus Development Panel on Destructive Behaviors in Persons with Developmental Disabilities. (1989). Treatment of destructive

behaviors in persons with developmental disabilities (NIH Publication No. 91-2410). Bethesda, MD: U.S. Department of Health and Human Services.

Neef, N. A., & Iwata, B. A. (1994). Current research on functional analysis methodologies: An introduction. Journal of Applied Behavior Analysis, 27, 211-214.

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

Nguyen, J. H., Derby, K. M., Hagopian, L., Fisher, W., Thompson, R., & Owen, J. (1996). An extended epidemiological analysis on the functions of aberrant behavior. Poster presented at the 22nd Annual Convention of the Association for Behavior Analysis, San Francisco, CA.

Northup, J., Jones, K., Broussard, C., DiGiovanni, G., Herring, M., Fusilier, I., & Hanchey, A. (1997). A preliminary analysis of interactive effects between common classroom contingencies and methylphenidate. Journal of Applied Behavior Analysis, 30, 121-126.

Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., & DeRaad, A. (1991). A brief functional analysis of aggressive and alternative behavior in an outclinic setting. Journal of Applied Behavior Analysis, 24, 509-522.

Nunnally, J. C. (1978). Psychometric theory. New York: McGraw-Hill Book Company.

O'Neill, R. E., Horner, R. H., Albin, R. W., Storey, K., & Sprague, J. R. (1990). Functional analysis of problem behavior: A practical assessment guide. Sycamore, IL: Sycamore.

Owens, R. G., & Ashcroft, J. B. (1982). Functional analysis in applied psychology. British Journal of Clinical Psychology, 21, 181-189.

Pedhazur, E. & Pedhazur Schmelkin L. (1991). Measurement, design, and analysis: An integrated approach. Hillsdale, NJ: Lawrence Erlbaum Associates.

Peterson, F. M., & Martens, B. K. (1995). A comparison of behavioral interventions reported in treatment studies and programs for adults with developmental disabilities. Research in Developmental Disabilities, 16, 27-41.

Peterson, R., & Peterson, L. (1968). The use of positive reinforcement in the control of self-destructive behavior in a retarded boy. Journal of Experimental child Psychology, 6, 351-360.

Piazza, C. C., Hanley, G. P., Bowman, L. G., Ruyter, J. M., Lindauer, S. E., & Saiontz, D. M. (1997). Functional analysis and treatment of elopement. Journal of Applied Behavior Analysis, 30, 653-672.

Piazza, C. C., Hanley, G. P., & Fisher, W. W. (1996). Functional analysis and treatment of cigarette pica. Journal of Applied Behavior Analysis, 29, 437-450.

Pyles, D. A. M., Riordan, M. M., & Bailey, J. S. (1997). The Stereotypy Analysis: An instrument for examining environmental variables associated with differential rates of stereotypic behavior. Research in Developmental Disabilities, 18, 11-38.

Reid, A. H., Ballinger, B. R., Heather, B. B., & Melvin, S. J. (1984). The natural history of behavioural symptoms among severely an profoundly mentally retarded patients. British Journal of Psychiatry, 145, 289-293.

Repp, A. (1994). Comments on functional analysis procedures for school-based behavior problems. Journal of Applied Behavior Analysis, 27, 409-411.

Repp, A. C., Felce, D., & Barton, L. E. (1988). Basing the treatment of stereotypic and self-injurious behaviors on hypotheses of their causes. Journal of Applied Behavior Analysis, 21, 281-289.

Repp, A., Singh, N., Olinger, E., & Olson, D. (1990). The use of functional analyses to test causes of self-injurious behavior: Rationale, current status and future directions. Journal of Mental Deficiency Research, 34, 95-105.

Rojahn, J., & Schroeder, S. R. (1991). Behavioral assessment. In J. L. Matson & J. A. Mulick (Eds.), Handbook of mental retardation (2nd ed., pp. 240-259). New York: Pergamon Press.

Rortvedt, A. K., & Miltenberger, R. G. (1994). Analysis of a high-probability instructional sequence and time-out in the treatment of child noncompliance. Journal of Applied Behavior Analysis, 27, 327-330.

Sattler, J. M. (1992). Assessment of children, 3rd ed. San Diego: Jerome M. Sattler, Publisher, Inc.

Schroeder, S. R. (1991). Self-injury and stereotypy. In J. L. Matson & J. A. Mulick (Eds.), Handbook of mental retardation (2nd ed., pp. 382-396). New York: Pergamon Press.

Schroeder, S. R., Rojahn, J., & Oldenquist, A. (1989). Treatment of destructive behaviors among people with mental retardation and developmental disabilities: Overview of the problem. In Treatment of destructive behaviors in persons with developmental disabilities (NIH Publication No. 91-2410, pp. 125-171). Bethesda, MD: NIH.

Scotti, J. R., Evans, I. M., Meyer, L. H., & Walker, P. (1991). A meta-analysis of intervention research with problem behavior: Treatment validity and standards of practice. American Journal on Mental Retardation, 96, 233-256.

Siegel, S., & Castellan, N. J. (1988). Nonparametric statistics for the behavioral sciences, 2nd ed. New York: McGraw-Hill, Inc.

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

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

Skinner, B. F. (1953). Science and human behavior. New York: The Free Press.

Skinner, B. F. (1969). Contingencies of reinforcement. Englewood Cliffs, NJ: Prentice Hall.

Slifer, K. S., Ivancic, M. T., Parrish, J. M., Page, T. J., & Burgio, L. D. (1986). Assessment and treatment of multiple behavior problems exhibited by a profoundly retarded adolescent. Journal of Behavior Therapy and Experimental Psychiatry, 17, 203-213.

Solnick, J. V., Rincover, A., & Peterson, C. R. (1977). Some determinants of punishing effects of time-out. Journal of Applied Behavior Analysis, 10, 415-424.

Sprague, J. R., & Horner, R. H. (1992). Covariation within functional response classes: Implications for treatment of severe problem behavior. Journal of Applied Behavior Analysis, 25, 735-745.

Sprague, J. R., & Horner, R. H. (1995). Functional assessment and intervention in community settings. Mental Retardation and Developmental Disabilities Research Reviews, 1, 89-93.

Spreat, S., & Connelly, L. (1996). Reliability analysis of the Motivation Assessment Scale. American Journal on Mental Retardation, 100, 528-532.

Sturmey, P. (1994). Assessing the functions of aberrant behaviors: A review of psychometric instruments. Journal of Autism and Developmental Disorders, 24, 293-304.

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

Sturmey, P. (1996). Functional analysis in clinical psychology. New York: John Wiley & Sons.

Sturmey, P., Carlsen, A., Crisp, A. G., & Newton, J. T. (1988). A functional analysis of multiple aberrant responses: A refinement and extension of Iwata et al.'s methodology. Journal of Mental Deficiency Research, 32, 31-46.

Tate, B. G., & Baroff, G. S. (1966). Aversive control of self-injurious behavior in a psychotic boy. Behavior Research and Therapy, 4, 281-287.

Taylor, J. C., Ekdahl, M. M., Romanczyk, R. G., & Miller, M. L. (1994). Escape behavior in task situations: Task versus social antecedents. Journal of Autism and Developmental Disorders, 24, 331-344.

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

Van Houten, R., & Rolider, A. (1991). Applied behavior analysis. In J. L. Matson & J. A. Mulick (Eds.), Handbook of mental retardation (2nd ed., pp. 569-585). New York: Pergamon Press.

Vollmer, T. R., Iwata, B. A., Duncan, B. A., & Lerman, D. C. (1993). Extensions of multielement functional analyses using reversal-type designs. Journal of Developmental and Physical Disabilities, 5, 311-325.

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). Within-session patterns of self-injury as indicators of behavioral function. Research in Developmental Disabilities, 14, 479-492.

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

Vollmer, T. R., & Smith, R. G. (1996). Some current themes in functional analysis research. Research in Developmental Disabilities, 17, 229-249.

Wacker, D. P. (1989). Evaluation of functional communication training: An analysis of active treatment components. In Treatment of destructive behaviors in persons with developmental disabilities (NIH Publication No. 91-2410, pp. 70-72). Bethesda, MD: Author.

Wacker, D. P., Berg, W. K., Cooper, L. J., Derby, K. M., Steege, M. W., Northup, J., & Sasso, G. (1994). The impact of functional analysis methodology on outpatient clinic services. Journal of Applied Behavior Analysis, 27, 405-407.

Weiseler, N. A., Hanson, R. H., Chamberlain, T. P., & Thompson, T. (1985). Functional taxonomy of stereotypic and self-injurious behaviors. Mental Retardation, 23, 230-234.

Weiseler, N. A., Hanson, R. H., & Nord, G. (1995). Investigation of mortality and morbidity associated with severe self-injurious behavior. American Journal on Mental Retardation, 100, 1-5.

Yeakel, M. H., Salisbury, L. L., Greer, S. L., & Marcus, L. F. (1970). An appliance for autoinduced adverse control of self-injurious behavior. Journal of Experimental Child Psychology, 10, 159-169.

Young, J. A., & Wincze, J. P. (1974). The effects of reinforcement of compatible and incompatible alternative behaviors on the self-injurious and related behaviors of a profoundly retarded female adult. Behavior Therapy, 5, 614-623.

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

Appendix A

QABF Items Grouped by Subscale

Attention

1. Engages in the behavior to get attention.

6. Engages in the behavior because he/she likes to be reprimanded.

11. Engages in the behavior to draw attention to him/herself.

16. Engages in the behavior to try to get a reaction from you.

21. Does he/she seem to be saying "come see me" or "look at me" when engaging in the behavior?

Escape

2. Engages in the behavior to escape work or learning situations.

7. Engages in the behavior when asked to do something.

12. Engages in the behavior when he/she does not want to do something.

17. Engages in the behavior to try to get people to leave him/her alone.

22. Does he/she seem to be saying "leave me alone" or "stop asking me to do this" when engaging in the behavior?

Non-social

3. Engages in the behavior as a form of "self stimulation".

8. Engages in the behavior even if he/she thinks there is no one in the room.

13. Engages in the behavior because there is nothing else to do.

18. Engages in the behavior in a highly repetitive manner, ignoring his/her surroundings.

23. Does he/she seem to enjoy the behavior, even if no one is around?

Physical

4. Engages in the behavior because he/she is in pain.

9. Engages in the behavior more frequently when he/she is ill.

14. Engages in the behavior when there is something bothering him/her physically.

19. Engages in the behavior because he/she is physically uncomfortable.

24. Does the behavior seem to indicate to you that he/she is not feeling well?

Tangible

5. Engages in the behavior to get access to items such as preferred toys, food, or beverages.

10. Engages in the behavior when you take something away from him/her.

15. Engages in the behavior when you have something he/she wants.

20. Engages in the behavior when a peer has something he/she wants.

25. Does he/she seem to be saying "give me that (toy item, food item)" when engaging in the behavior?

Appendix B

Consent to Participate in Assessment and Treatment of Behavior Problems

Client's Name: Date:

The purpose of this letter is to seek your permission for XXXX to participate in an evaluation to assess his/her problem behaviors for the development of appropriate treatment procedures. The assessment process is described below. Please complete the last page and return this entire document to the address above.

1. SPECIFIC ISSUE

XXXX currently engages in self-injurious behavior/aggressive behavior in the form of xxxx. This behavior occurs xxxx times per day with an episode lasting xxxx minutes. Staff report that XXXX has injured him/herself in the past, causing bruises, scratches, cuts, _____.

Direct care staff report that XXXX has exhibited these behaviors over the past years and that they are long term and relatively unchanged. XXXX has not received medication for behavior management. Previous behavior strategies used include Differential Reinforcement of Appropriate Behavior (DRA), relaxation, environmental engineering, and redirection procedures. Limited documentation hinders evaluation of the effectiveness of these procedures.

2. BEHAVIOR ASSESSMENT PROCEDURES:

We will try a variety of things to find out what is causing this problem behavior. Trained psychological staff will conduct interviews and complete questionnaires with the people who work with XXXX. XXXX will be observed several times and in several places. Informal observations will be made. If these natural observations do not tell us why XXXX is doing this behavior, we will have to attempt more formal methods.

One way is to try to watch XXXX at several different times of the day during different activities. This is called descriptive analysis. When the behavior occurs, trained staff will respond in different ways. If the behavior becomes a risk, it will be immediately stopped. The goal is to try to determine the situations in which the behavior occurs most so that we can develop appropriate interventions to decrease the behavior.

However, sometimes the individual will not do the behavior if he or she is watched. The observers may be distracting. To develop a Behavior Treatment Plan, we have to see the behavior to know what is causing it. If XXXX does not do the behavior

when watched, it may be necessary to bring XXXX to a therapy room and try to do the same tasks in a more controlled setting. This procedure is called functional analysis.

During a functional analysis, we will try to copy the natural setting as much as possible to see if XXXX engages in the behavior when he/she is performing a task, given attention, or not given attention. Before doing a functional analysis, a risk assessment will be done to determine conditions for ending a functional analysis (e.g., if SIB/Aggression escalates). Trained staff will determine if XXXX is at risk at any time. Sessions are stopped immediately when necessary. These procedures allow us to determine what is causing XXXX to harm him/herself or engage in aggression and thus better able us to reduce these harmful behaviors.

3. CLIENT'S STATUS IN REGARD TO THIS ISSUE:

XXXX is a xx-year old, xxxx retarded individual who is non-interdicted. Mr./Ms./Mrs. XXXX is his/her correspondent.

4. PURPOSE OF BEHAVIOR ASSESSMENT:

The behavior assessment procedures described above should allow us to develop a more appropriate intervention to decrease the problem behavior XXXX is exhibiting. Without such as assessment XXXX may continue to engage in the problem behaviors; thus, the above procedures have been found to be beneficial to our clients in the past.

5. QUESTIONS/CONCERNS:

If you have any questions, concerns, or desire further information please contact Dr. Randy Logan or Karena Rush, M.A. at 318-641-2258.

6. RIGHT TO REFUSE:

XXXX and XXXX have the right to refuse by checking I DISAGREE and signing. Also, if consent is given, it may be withdrawn at any time.

_____ I AGREE _____ I DISAGREE

Signatures:

_____ _____

Client's Name Date Parent/Guardian Date

Appendix C

Functional Analysis Analogue Assessment Data Collection Sheet

NAME:_____ **DATE:** _____

Therapist_____ **Observer**_____

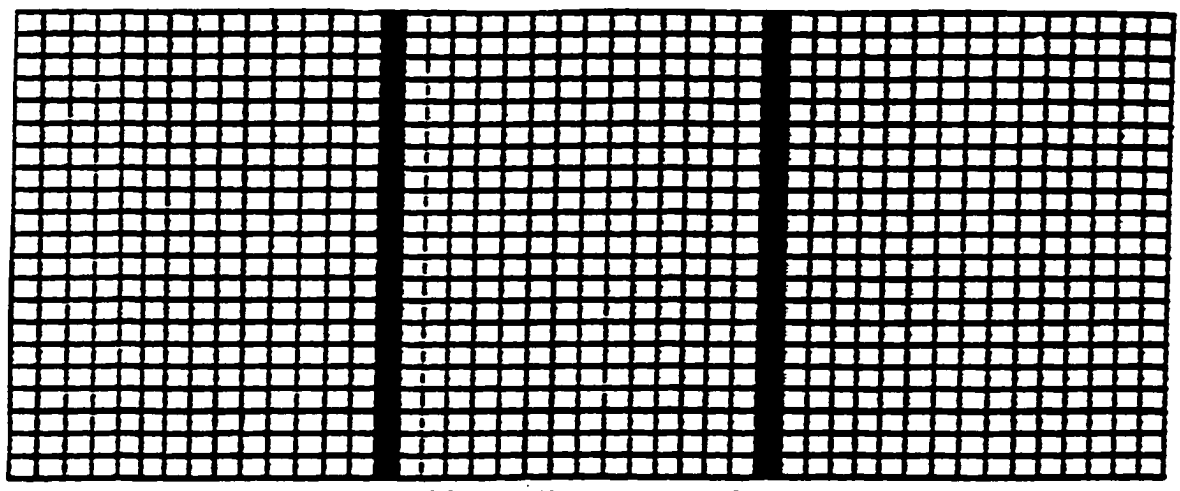
Target Behavior:_____

Condition_____ **Condition**_____ **Condition**_____

	Tally	Frq
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

	Tally	Frq
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

	Tally	Frq
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		



MINUTES WITHIN SESSIONS

Vita

Theodosia Paclawskyj received both her Master of Arts and Doctor of Philosophy degrees from the Department of Psychology at Louisiana State University in Baton Rouge, Louisiana. Prior to this, she completed a Bachelor of Arts degree in Public Health at the Johns Hopkins University in 1990 and then functioned as a behavior therapist working with children with mental retardation before beginning graduate school. Theodosia currently is starting a post-doctoral fellowship at the Kennedy Krieger Institute with the department of Behavioral Psychology in Baltimore, Maryland. Her professional interests include the assessment and treatment of severe behavior disorders and dual diagnosis, program evaluation, and teaching.

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Theodosia Paclawskyj

Major Field: Psychology

Title of Dissertation: Questions About Behavioral Function (QABF): A Behavioral Checklist for Functional Assessment of Aberrant Behavior

Approved:

Major Professor and Chairman

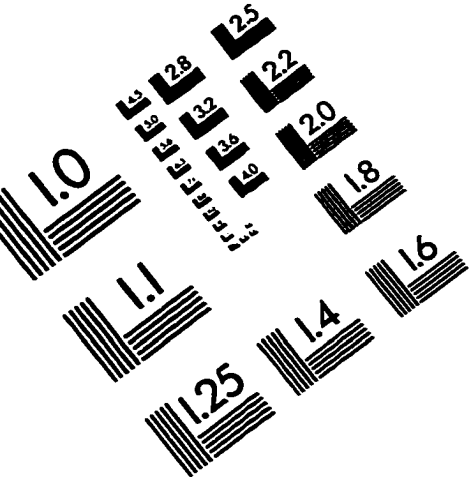
Dean of the Graduate School

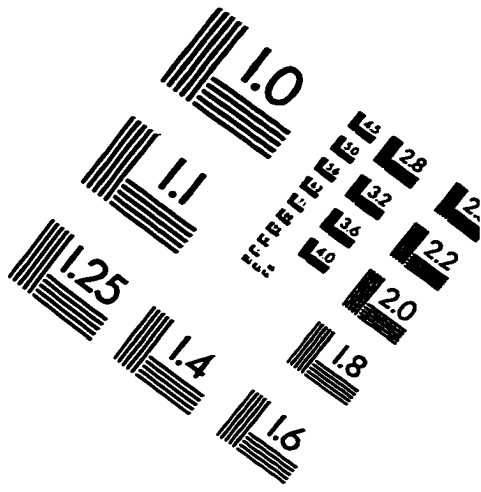
EXAMINING COMMITTEE:

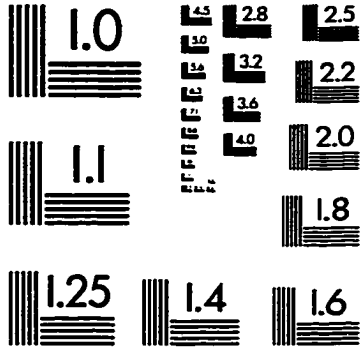
Date of Examination:

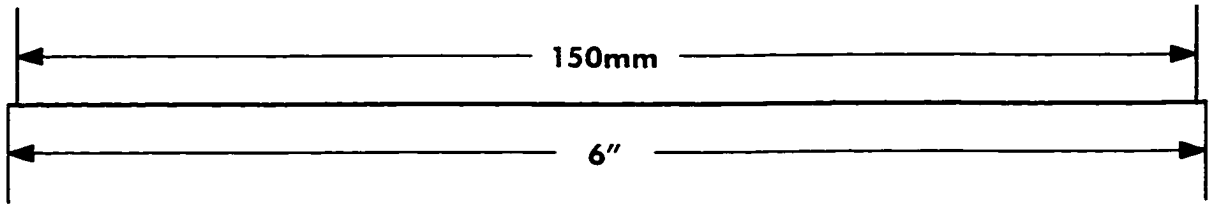
August 18, 1998

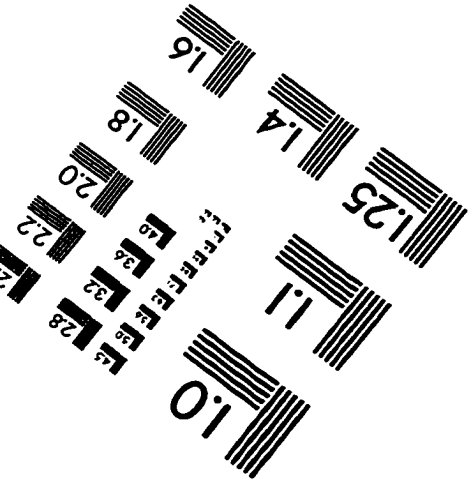
IMAGE EVALUATION TEST TARGET (QA-3)

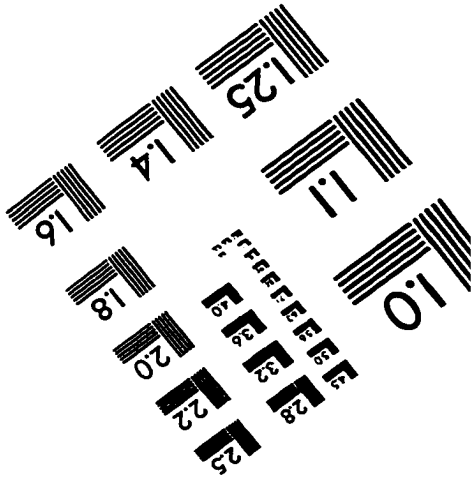












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989