

Questions about behavioral function (QABF): a behavioral checklist for functional assessment of aberrant behavior

Theodosia R. Paclawskyj^a, Johnny L. Matson^{b,*},
Karena S. Rush^a, Yemonja Smalls^b, Timothy R. Vollmer^c

^a*The Kennedy Krieger Institute and The Johns Hopkins School of Medicine, Baltimore, MD
21218, USA*

^b*Louisiana State University, Baton Rouge, LA 70803, USA*

^c*University of Florida*

Abstract

Functional assessment is a method to identify the relationships between a behavior of interest and an individual's environment. Traditional methods for functional assessment have relied on experimental techniques in which analog sessions are designed to replicate conditions in the individual's environment. However, these techniques can be time-consuming, require advanced training, and rely on the availability of extensive resources in the individual's setting. Development of a brief functional assessment checklist would circumvent these difficulties and meet clinical needs for efficient assessment methods. The current study provides psychometric data for the Questions About Behavioral Function. These data include test–retest, inter-rater, and internal consistency. © 2000 Elsevier Science Ltd. All rights reserved.

1. Introduction

Functional assessment has proven useful for identifying environmental causes of maladaptive behaviors (Sturme, 1996). Traditionally, experimental functional analyses were used, but they have proven to be very labor intensive. As a

* Corresponding author. Tel.: +1-225-388-4104; fax: +1-225-388-4125.

E-mail address: johnmatson@home.com (J. Matson).

result, researchers have been searching for more efficient means of obtaining this type of data. Various scales have been published for some time, but unfortunately no psychometric data were collected (e.g., Functional Analysis Interview Form (O'Neil et al., 1990). Conversely, the Motivation Assessment Scale (Durand & Crimmins, 1988) has been studied but the psychometric qualities have proven to be inconsistent, and generally low inter-rater and internal consistency reliability coefficients have been reported (Bihm et al., 1991; Crawford et al., 1992; Newton & Sturmey, 1991; Sigafoos et al., 1994; Spreat & Connelly, 1996; Zarcone et al., 1991). Nonetheless, these researchers made valuable contributions by developing first attempts at time-efficient functional analysis methods.

Given the potential value of a checklist to determine behavior function, further efforts at scale development seem warranted. As a result the Questions About Behavioral Function (QABF) (Matson & Vollmer, 1995) was designed for persons with mental retardation. Maintaining variables included in this scale are social attention, escape, tangible reinforcement, physical discomfort, and nonsocial reinforcement, and constitutes a broader range of variables than previously addressed by functional analysis checklists. In particular, under-investigated variables such as social avoidance (Taylor et al., 1994) and physical discomfort (Carr, 1994; Lowry & Sovner, 1991) are included to make the scale more comprehensive than previous instruments.

Matson et al. (1996) presented the initial psychometric data for the QABF. 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 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 ($r = 0.86$ and $r = 0.91$, 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.

The QABF has also been found to be successful in identifying the function of self-injury, aggression, or stereotypies in 84% of 398 people (Matson, Bamburg, Cherry, & Paclawskyj, 1999). Further, treatments developed from the QABF identified functions, resulted in significantly more effective treatments for these problems than treatments that were not tailored to antecedent functions identified with the QABF.

In the present research, we provide a range of psychometric data for the QABF. In experiment 1 we attempted to determine the stability of the instruments across rates and over time, inter-rater and test–retest reliability were assessed. Experiment 2 was designed to substantiate the initial factor solution identified by Matson et al. (1996), a second exploratory factor analysis was conducted.

2. Method

2.1. Experiment 1

2.1.1. Subjects

Data from 34 participants were included for the test–retest portion of the study. Participants were predominantly male (82.4%) functioning in the profound range of mental retardation (76.5%). The remainder functioned in the severe range of mental retardation. Subjects' target behaviors included self-injury, aggression, property destruction, tantrums/verbal aggression, stereotypy, pica, stealing, and elopement.

Data were collected at a large state developmental center. Data were collected as part of ongoing clinical services provided to these clients and were given to developmental center staff for incorporation into existing behavior plans.

Twenty-three additional participants (total $N = 57$) were included for the inter-rater portion of the study. This sample consisted of males functioning predominantly in the severe (26.1%) and profound (47.8%) levels of mental retardation, with the same target behaviors as above.

2.1.2. Procedure

A graduate student examiner administered the QABF to a direct care staff member who had known the client for at least 6 months. The same examiner readministered the QABF to the same informant between 1 and 3 weeks after the initial administration.

2.1.3. Results

To assess stability of the individual QABF items over time, three measures of reliability were assessed: Spearman rank–order correlation coefficients were computed for each item, total agreement between items upon separate administrations was calculated through percent agreement, and Cohen's Kappa was computed when feasible for each item to correct for chance levels of agreement (kappa could not be computed when the contingency tables were not symmetric).

Spearman rank–order correlations on the whole were high, with 76% of item statistics exceeding the minimally acceptable value of 0.8 (Nunally, 1978; Sturme, 1996). Correlations ranged from .646 to 1.0. Similarly, total percent agreement was good, with 96% of the items exceeding a minimum of 80% (Rojahn & Schroeder, 1991). Total agreement ranged from 69.57 to 95.65%. The available kappa values were also adequate, ranging from .642 to 1.0, with 83% exceeding a minimum value of .7 (Sattler, 1992).

To assess the stability over time for subscales and the QABF as a whole, Pearson product–moment correlation coefficients were computed. High reliability was obtained for each subscale and total score (range = 0.795–0.990, $p < .01$).

Inter-rater 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.

Spearman rank order correlations were slightly lower for test–retest, ranging from -0.095 to 1.0 , with 52% of items exceeding a minimum of 0.8 . Total agreement also was slightly lower on the whole, ranging from 69.57 to 95.65% , with 56% of items exceeding 80% agreement. Kappa values were adequate, ranging from 0.427 to 0.921 , with 41% exceeding a minimum of 0.7 . Pearson product–moment correlations were acceptable for each subscale and the total score (range = 0.79 – 0.987 , $p < .01$).

2.1.4. Discussion

The QABF appears to have good stability over time. Test retest reliability for the individual items and the subscale scores were 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 to 3 weeks apart. Similarly, different raters tended to agree in their impressions of a client's behavior. Inter-rater reliability was generally high although slightly lower than that for test–retest coefficients. Each subscale of the QABF had consistently high external reliability. This finding is in contrast to the variable results observed in studies of reliability of the Motivation Assessment Scale, the most widely available functional assessment checklist (e.g., Crawford et al., 1992; Newton & Sturmey, 1991; Sigafoos et al. 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 and had displayed their problem behaviors over a period of years.

2.2. Experiment 2

2.2.1. Subjects/procedure

Completed QABF scale for 243 additional participants were used to conduct the internal reliability studies and a second exploratory factor analysis (subsequent to the original analysis reported by Matson et al., 1996). The majority of this sample was male (62.4%) and diagnosed with profound mental retardation (69.4%). Target behaviors included self-injury, aggression, property destruction, tantrums/verbal aggression, stereotypy, pica, stealing, elopement, and rectal digging. Data for the target behavior were unavailable for 12 subjects (4.9%). As with the previous assessments, a trained examiner administered the QABF to an informant who knew the client for at least 6 months.

Table 1
Factor loadings of items across factors

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
1	0.052	0.229	0.053	−0.164	0.863
2	0.853	0.065	0.052	−0.190	0.018
3	−0.194	−0.130	−0.106	0.849	−0.129
4	0.028	0.053	0.887	−0.043	0.056
5	0.024	0.822	0.025	−0.201	0.129
6	−0.002	−0.072	0.016	−0.041	0.620
7	0.839	0.074	0.071	−0.138	0.021
8	−0.302	−0.218	−0.044	0.806	−0.159
9	0.084	−0.045	0.835	−0.042	0.119
10	0.202	0.773	0.023	−0.090	0.048
11	0.054	0.202	0.075	−0.153	0.899
12	0.885	0.132	0.082	−0.159	−0.001
13	−0.035	−0.067	−0.011	0.829	−0.032
14	0.109	0.126	0.851	0.036	0.036
15	0.045	0.887	0.029	−0.091	0.089
16	0.001	0.049	−0.005	−0.138	0.865
17	0.810	0.001	0.039	−0.188	0.016
18	−0.240	−0.202	−0.007	0.820	−0.184
19	0.031	0.036	0.882	0.028	−0.006
20	0.034	0.789	0.001	−0.099	0.086
21	−0.066	0.105	0.032	−0.104	0.882
22	0.888	0.049	0.036	−0.143	−0.035
23	−0.295	−0.200	−0.034	0.817	−0.242
24	0.014	−0.110	0.860	−0.160	−0.046
25	0.015	0.901	−0.016	−0.188	0.057

2.2.2. Results

To assess internal consistency, coefficient alpha was computed using the sample of 243 participants. Coefficient alpha for the present study was computed for the individual subscales and the test as a whole. Coefficient alpha was very high for each subscale (range = 0.900–0.928), but lower for the test as a whole (0.601).

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, corrected for unequal length was 0.600. The lower values for alpha for the total score and for the Spearman–Brown correlation 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.

To confirm the heterogeneity of the QABF, an exploratory factor analysis was conducted. Principal Axis Factoring 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 1 has data on outlines, the factor structure and item loadings.

2.2.3. Discussion

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 five-factor solution identified by Matson et al. (1996). Each item loaded onto a factor corresponding to a hypothesized subscale, and accounted for a considerable portion of the variance shared by test items (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, tangible) and are supported by high internal consistency among the items of each subscale.

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

3. General discussion

There is a need for more efficient methods of conducting functional assessments, as extended functional analysis sessions are often 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 analog sessions, except perhaps in very extreme cases. Total administration time for each QABF was approximately 20 min, which is substantially less time than that reported in most studies utilizing analog functional analyses. For clinicians with large caseloads this represents a significant decrease in the amount of time needed to provide essential clinical intervention.

The present psychometric data support the soundness of the QABF and its role in conducting functional assessment. Analyses of test-retest reliability, inter-rater reliability, and internal consistency indicate that the measure is viable for conducting functional analyses in this population. Further, through factor analysis we revealed the presence of five factors that correlate with the behavioral functions assessed by the QABF (Matson et al., 1996). Thus, the QABF appears

to be a promising substitute for much more labor extensive methods in the identification of environmental antecedents of problem behaviors.

References

- 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. *Psychol Rep* 68, 1235–1238.
- Carr, E. G. (1994). Emerging themes in the functional analysis of problem behavior. *J Appl Behav Anal*, 27, 393–399.
- Crawford, J., Brockel, B., Schauss, S., & Miltenberger, R. G. (1992). A comparison of methods for the functional assessment of stereotypic behavior. *JASH* 17, 77–86.
- Durand, V. M., & Crimmins, D. B. (1988). Identifying the variables maintaining self-injurious behavior. *J Autism Develop Dis*, 18, 99–117.
- Horner, R. H. (1994). Functional assessment: contributions and future directions. *J App Behav Anal*, 27, 401–404.
- Lowry, M. & Sovner, R. (1991). The functional significance of problem behavior: A key to effective treatment. *Habilitative Mental Healthcare Newsletter*, 10, 59–62.
- Matson, J. L., Bamburg, J. W., Cherry, K. E., & Paclawskyj, T. R. (1999). A validity study on the Questions About Behavioral Function (QABF) scale: predicting treatment success for self-injury, aggression, and stereotypies. *Res Dev Disabil*, 20, 142–160.
- Matson, J. L. & Vollmer, T. R. (1995). *User's Guide: Questions About Behavioral Function (QABF)*. Baton Rouge, LA: Scientific Publishers.
- Matson, J. L., Vollmer, T. R., Paclawskyj, T. R., Smiroldo, B. B., Applegate, H. R., & 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.
- Newton, J. T. & Sturmey, P. (1991). The Motivational Assessment Scale: Inter-rater reliability and internal consistency in a British sample. *J Ment Defic Res*, 35, 472–474.
- Nunally, J. C. (1978). *Psychometric Theory*. New York: McGraw–Hill Book Company.
- O'Neil, 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.
- 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.
- Sattler, J. M. (1992). *Assessment of Children* (3rd ed.). San Diego: Jerome M. Sattler, Publisher, Inc.
- Sigafoos, J., Kerr, M., & Roberts, D. (1994). Inter-rater reliability of the Motivational Assessment Scale: failure to replicate with aggressive behavior. *Res Dev Disabil*, 15, 333–342.
- Spreat, S., & Connelly, L. (1996). Reliability analysis of the Motivation Assessment Scale. *Am J Ment Retard*, 100, 528–532.
- Sturmey, P. (1996). *Functional Analysis in Clinical Psychology*. New York: John Wiley & Sons.
- Taylor, J. C., Ekadahl, M. M., Romanczyk, R. G., & Miller, M. L. (1994). Escape behavior in task situations: task versus social antecedents. *J Autism Develop Dis*, 24, 331–344.
- 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.
- 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. *Res Dev Disabil*, 12, 349–360.