Using ABC Narrative Recording to Identify the Function of Problem Behavior:

A Pilot Study

Marc J. Lanovaz[a,*], Malena Argumedes[a], Dona Roy[b],

Julie R. Duquette[c], and Nicholas Watkins[d]

[a] École de psychoeducation, Université de Montréal, C. P. 6128, succ. Centre-Ville, Montréal, QC H3C 3J7 Canada
[b] Département de psychologie, Université du Québec à Montréal, C.P. 8888 succ. Centre-ville, Montréal, QC H3C 3P8 Canada
[c] Centre de réadaptation de l'Ouest de Montréal, 8000 rue Notre-Dame, Lachine, QC H8R 1H2 Canada
[d] Douglas College, P.O. Box 2503, New Westminster, BC V3L 5B2 Canada
* Corresponding author. Tel.: 1 514 343 6111 ext. 81774
E-mail address: marc.lanovaz@umontreal.ca

Abstract

Many professionals report using ABC narrative recording to identify the function of problem behavior in children with developmental disabilities, but research has not established whether their analyses yield valid results. Thus, the purpose of this study was to examine whether the function identified by expert reviewers using ABC narrative recordings would match the one identified by a functional analysis (FA) and the Questions About Behavioral Function (QABF) scale in four children with developmental disabilities. The functions identified by all twelve experts using the ABC narrative recordings matched at least one of the functions identified by the FA for three of four participants. The experts' analyses also agreed with the informant-based assessment at a statistically significant level for two of three participants with a conclusive QABF. Altogether, the results suggest that ABC narrative recording may be useful to generate hypotheses to identify the function of problem behavior, but that more research is needed before recommending its use as a standalone functional behavior assessment.

Keywords: ABC narrative recording, descriptive assessment, functional analysis, indirect assessment, problem behavior, QABF

Using ABC Narrative Recording to Identify the Function of Problem Behavior:

A Pilot Study

1. Introduction

Conducting a functional behavior assessment is generally considered best practice in the assessment and treatment of problem behavior in individuals diagnosed with a developmental disability (Horner & Carr, 1997; Iwata & Dozier, 2008). Clinicians and researchers use functional behavior assessments to identify the contingencies that maintain the occurrence of problem behavior (Matson & Minshawi, 2007). The results of functional behavior assessments subsequently equip the clinician to develop a function-based intervention to replace problem behaviors with more socially appropriate alternatives. For example, a function-based intervention for problem behavior maintained by attention may involve teaching the individual to appropriately request attention using a picture whereas an intervention unrelated to function may involve the delivery of an arbitrary reinforcer (e.g., edibles) when the problem behavior has not occurred for a certain period of time. Researchers have shown that function-based interventions produce more clinically desirable effects than interventions unrelated to function (e.g., Ingram, Lewis-Palmer, & Sugai, 2005; Iwata, Pace, Cowdery, & Miltenberger, 1994; Matson, Bamburg, Cherry, & Paclawskyj, 1999).

The two methods that have the most empirical support for identifying the function of problem behavior are functional analyses (FA) and checklist-based functional assessments (Matson & Mishawi, 2007). The FA involves the systematic manipulation of antecedent and consequent events during analog conditions (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994). A function is identified when a problem behavior occurs more frequently in the presence than in the absence of certain antecedents and consequences. The FA has been shown to

be effective in identifying the contingencies maintaining a wide variety of behavior in individuals with developmental disabilities (Hanley, Iwata, & McCord, 2003). Given that the procedures allow the identification of causal relations, FA methods are typically considered the gold standard by which to identify the contingencies maintaining a problem behavior (Fox & Davis, 2005). However, oft-cited barriers to its implementation include risks associated with evoking possibly dangerous behavior, the potential absence of the variables maintaining the problem behavior in an analog setting, concerns related to strengthening or increasing the problem behavior as well as the required time, resources, and effort (Neef & Peterson, 2007). Although researchers have proposed several solutions to minimize the impact of these barriers (Iwata & Dozier, 2008), designing FA still requires professional expertise and sufficient time to observe the occurrence of the problem behavior.

One alternative to FA methods is the use of checklist-based questionnaires, which rely on responses provided by informants (e.g., caregivers, parents, staff). Examples of questionnaires include the Motivation Assessment Scale (MAS; Durand & Crimmins, 1988), the Functional Analysis Screening Tool (FAST; Iwata & DeLeon, 1995), and the Questions About Behavioral Function (QABF; Matson & Vollmer, 1995). To date, only the use of the QABF has amassed substantial empirical support (Matson, Tureck, & Rieske, 2012). The QABF is a 25-item questionnaire that assesses five functions—attention, escape, nonsocial, physical, and tangible. Researchers have shown that interventions based on the results of the QABF were more effective than those that were not (Matson et al., 1999). Moreover, studies comparing the results of the QABF with the results of FA have shown exact agreement between the two assessments in 43% to 86% of individuals with developmental disabilities (Hall, 2005; Healy, Brett, & Leader, 2013; Paclawskyj, Matson, Rush, Smalls, & Vollmer, 2001; Smith, Smith, Dracobly, & Pace, 2012;

Tarbox et al., 2009; Watkins & Rapp, 2013). One concern with questionnaires (including the

QABF) is that the identified function may differ from one respondent to another (Smith et al.,

2012). However, it may be the case that problem behavior may be under different discriminative

control across raters (i.e., the function may vary from person to person), which could explain the

moderate interrater reliability. This concern not withstanding, the QABF enjoys acceptable to

good interrater reliability (e.g., Paclawskyj, Matson, Rush, Smalls, & Vollmer, 2000; Zaja,

Moore, Van Ingen, & Rojahn, 2010) and is empirically valid (e.g., Freeman, Walker, &

Kaufman, 2007; Matson & Boisjoli, 2007).

Even though FA and the QABF have the most empirical support in the research literature,

descriptive methods continue to be employed (Desrochers, Hile, & Williams-Mosely, 1997;

Ellingson, Miltenberger, & Long, 1999; Love, Carr, Almason, & Petursdottir, 2009). In a survey

on clinical practices in early intervention, Love et al. (2009) reported that 96% of professionals

used descriptive assessments whereas only 56% used experimental FA. Descriptive assessments

involve the direct observation of antecedents and consequences associated with the problem

behavior in the individual's natural environment (Sloman, 2010). Clinicians and researchers may

use one of three methods to collect data during descriptive assessments: scatterplot analysis,

ABC continuous recording, and ABC narrative recording (Neef & Peterson, 2007).

The scatterplot analysis involves collecting data on the temporal distribution of the

problem behavior throughout the day in order to identify potential functional relations

(Touchette, MacDonald, & Langer, 1985). However, in a study with 20 participants diagnosed

with an intellectual disability, Kahng et al. (1998) found that the scatterplot produced poor

agreement between observers for 5 participants and failed to detect temporal patterns in the

remaining 15 participants. Thus, empirical support for the scatterplot is currently insufficient (Matson & Minshawi, 2007).

With ABC continuous recording, clinicians collect real-time continuous data on the problem behavior as well as on the events that precede and follow it. Calculations are then carried out to determine the probability of the behavior given the occurrence of specific antecedents, consequences, or both (McComas, Moore, Dahl, Hartman, Hoch, & Symons, 2009). Researchers have repeatedly shown that ABC continuous recording does not match the results of FA in a large majority of cases (Hall, 2005; Lerman & Iwata, 1993; Tarbox et al., 2009; Thompson & Iwata, 2007). In the previously cited studies, the results of both analyses matched for only 14% to 33% of cases. These results are no better than random guessing, which should identify the same function in 25% of cases (i.e., assuming that each behavior only has one function). One explanation for the weak validity of descriptive assessment is that attention is often correlated with problem behavior regardless of its function (St. Peter, Vollmer, Bourret, Borrero, Sloman, & Rapp, 2005). A second explanation is that the function of problem behavior maintained by thin reinforcement schedules may be difficult to detect given that the contingencies may only be observable for a fraction of occurrences (Marion, Touchette, & Sandman, 2003). As such, ABC continuous recordings are not recommended as standalone methods to identify the function of problem behavior (Thomson & Iwata, 2007).

A third type of descriptive assessment is ABC narrative recording, which involves anecdotally reporting the events correlated with the problem behavior. For example, a parent may be asked to describe in writing the environmental events immediately before and after a problem behavior. Ellingson et al. (1998) reported that approximately 30% of professionals used this type of functional behavior assessment in their practice. In a more recent study, Lerman,

Hovanetz, Strobel, and Tetreault (2009) reported that 12 of 16 schoolteachers had experience with ABC narrative recordings prior to their study. Yet, despite its popularity, we were unable to locate a single study that examined the validity of ABC narrative recordings in identifying the function of problem behavior.

The analysis of narrative recordings by professionals may not necessarily yield the same results as other descriptive assessments. In contrast with quantitative analyses used in the ABC continuous recording, professionals may qualitatively analyze ABC narrative data. One potential advantage of conducting qualitative rather than quantitative analyses is that the clinician may take into consideration the topography of the behavior and the limitations of descriptive assessments, two qualitative variables overlooked by statistical analyses (e.g., conditional probabilities, lag sequence analyses). For example, researchers have shown that stereotypy is maintained by nonsocial reinforcement in more than 90% of cases (Matson et al., 2009; Wilke et al., 2012). Assume that a child with an autism spectrum disorder engages in hand flapping maintained by nonsocial reinforcement, but that his mother consistently provides attention contingent on the behavior (e.g., she asks him to stop). Knowing that false positives (concluding that a consequence produces behavioral change when it fact it does not; a Type-I error) for attention are very common and that stereotypy is nearly always maintained by nonsocial reinforcement, a clinician qualitatively analyzing ABC narrative recordings may conclude that hand flapping was maintained by nonsocial reinforcement despite most instances being followed by attention. In contrast, computing conditional probabilities would have erroneously concluded that hand flapping was maintained by attention.

Moreover, researchers often recommend collecting narrative data to generate a hypothesis before implementing a FA (Hanley, 2012; Neef & Peterson, 2007). Whether the hypotheses

generated by the clinicians are generally accurate remains unknown. If the narrative analysis fails

to yield the correct function in a majority of cases, its clinical utility would be limited. The first

step is to examine whether the analysis of ABC narrative recordings produces results consistent

with other well-established methods to identify the function of problem behavior. Thus, the

purpose of this study was to examine whether the function identified by expert reviewers using

narrative recordings would match the function identified by a FA and a QABF.

2.   Method

2.1. Participants and Target Behaviors

Four children diagnosed with a developmental disability participated in the research

study. Luke was an 8-year-old boy diagnosed with autism. He engaged in hand flapping, defined

as two or more up or down or side-to-side hand movements. Luke's hand flapping interfered

with his academic activities as well as his daily tasks. He used a picture exchange

communication system to make requests. Wendy was a 12-year-old girl diagnosed with a mild

intellectual disability. She engaged in self-pinching, defined as any contact between her skin and

fingers, mouth, or object capable of leaving visible marks (e.g., scratches, small depression in

skin). The child would often take objects (e.g., fork, USB key) and press them on her forearm.

Wendy's self-inflicted wounds sometimes required medical treatment due to their severity. She

communicated via sign language and a picture exchange communication system, both without

assistance. Julie was a 3-year-old girl diagnosed with autism. She engaged in self-injurious

behavior (SIB), defined as any behavior directed towards herself capable of causing harm.

Examples of SIB included biting, head banging, eye poking, and hitting. Her SIB interfered

considerably with her early intensive behavioral intervention program and required the use of a

helmet during high-intensity occurrences. Julie communicated her needs by pointing or grabbing

her communicative partner. Finally, Brian was a 13-year-old boy diagnosed with autism. He engaged in repetitive vocalizations, defined as acontextual sounds or words produced by the nose, mouth, or throat. Brian's repetitive vocalizations occurred in both learning and leisure contexts. He did not have a formal means of communicating with others. With the exception of Julie, all assessments took place in each child's home. For Julie, we conducted the QABF and the descriptive assessment at her home, but due to the limited availability of her caregiver, we carried out the FA in an early intervention center. The informants who responded to the QABF and collected narrative data were the child's mother (Luke and Wendy), father (Julie), or caregiver with whom he lived (Brian).

2.2.Procedures

To examine the validity of analyses conducted using ABC narrative recordings, an interviewer first administered the QABF with each participant's caregiver consistent with the rubrics outlined in Matson and Vollmer (1995). The caregiver was then taught to collect ABC narrative data on the occurrence of the target behavior. Once the caregiver had collected sufficient narrative data (defined below), a trained professional blind to the results of both the QABF and the ABC narrative analyses conducted a FA.

2.2.1. Questions About Behavioral Function (QABF)

At the start of each child's participation in the study, an interviewer administered the QABF to a caregiver either in person or over the telephone. The QABF is a 25-item rating scale (five items per function) that takes less than 15 min to conduct. Prior to each administration, the interviewer and the caregiver together operationally defined the target behavior. The interviewer then read each QABF item aloud and the caregiver responded whether the statement applied never (0), rarely (1), sometimes (2), or often (3). As recommended by Matson and Vollmer

(1995), we concluded that a function was identified when four or five of items were endorsed (i.e., a response of rarely, sometimes, or often) by the caregiver and when no other function received an endorsement score higher than three. Additionally, we conducted a more in-depth analysis of the severity scores when more than one function had an endorsement score of at least four (Watkins & Rapp, 2013). To minimize bias, the caregiver remained blind to the results of the QABF until the end of the child's participation in the study.

2.2.2.   ABC narrative recording

Once the QABF had been completed, we used behavioral skills training to teach each caregiver how to collect ABC narrative data. Behavioral skills training is a teaching strategy that has amassed strong empirical support for teaching skills to caregivers of children with developmental disabilities (e.g., Lafasakis & Sturmey, 2007; Luiselli, 2011; Miles & Wilder, 2009; Seiverling, Williams, Sturmey, & Hart, 2012). First, the trainer provided the caregiver with specific written and oral data collection instructions (detailed instructions available from first author). Then, the trainer modeled the procedures by recording narrative data while viewing pre-recording video vignettes in the presence of the caregiver. Immediately after the trainer role-modeled data collection with three vignettes, the caregiver practiced. During each practice trial, we required the caregiver to record her name, the date, the time, as well as a detailed narrative description of the antecedents, target behavior, and consequences. During rehearsal, the trainer provided positive comments on the data that were correctly collected and corrective feedback on data that were incorrectly recorded. Training continued until the caregiver correctly identified the antecedent, behavior, and consequence for three consecutive video vignettes. When the caregiver met this criterion, the trainer reviewed the operational definition of the child's target behavior with the caregiver and instructed her to collect data for at least 20 occurrences. The trainer spoke

with each caregiver via telephone once every one to two weeks to respond to questions and verify progress.

Once the caregiver had completed collecting narrative data, we transcribed the observations using a word processor while removing all personal identifying information (e.g., name of child, caregiver, sibling). We transcribed two sets of ABC narrative recordings for each participant; one set included the target behavior and the second set omitted the target behavior. We manipulated the data in this manner to test the hypothesis that knowing the topography of a target behavior would influence the subsequent analysis. Twelve expert reviewers qualitatively analyzed the narrative data and developed hypotheses on the function of each problem behavior. The expert reviewers were all certified by the Behavior Analyst Certification Board ® (BACB), held minimally a master's degree, had at least five years of experience ($M = 14$, range: 5-33) applying and supervising behavior analytic assessments and interventions, and reported using functional behavior assessment as part of his or her practice, research, or both.

The expert reviewers were blind to the results of both the QABF and the FA. They did not have access to the diagnosis or the age of each participant. The only information available to them was the events recorded by the caregiver during ABC narrative recording (i.e., antecedents, behavior [in half the cases], consequences, date, and time of occurrence). For each participant, the expert reviewer could propose either a single function or dual functions. Identifying a second function was optional and expert reviewers were instructed to select a second function only if he or she hypothesized that the behavior was multiply controlled. To facilitate the comparison with the results of the FA and QABF, the expert reviewers chose amongst the following four functions: attention, escape, nonsocial, and tangible. For two of the participants, the expert reviewer received ABC narrative data that included the behavior column. For the other two

participants, the behavior column remained blank and any information detailing the target

behavior had been removed from the antecedent and consequence columns. In these cases, the

expert reviewer was blind to the topography of the target behavior. We developed six

combinations of data sets; each set included incomplete recordings for two of the participants

(i.e., without the behavior column) and complete recordings for the remaining two participants

(i.e., with the behavior column). We randomly assigned two expert reviewers to each of the six

data sets.

2.2.3.   Functional analysis

A trained professional blind to the results of the QABF and the descriptive analysis

conducted a FA with each participant. Prior to completing the FA, she conducted a brief

interview with each caregiver and observed the child in order to set up conditions most likely to

evoke the target behavior (Hanley, 2012). For Luke and Brian, she followed a sequence similar

to the model proposed by Vollmer, Marcus, Ringdahl, and Roane (1995). That is, she first

alternated the conditions in a multielement design. If the results remained undifferentiated, the

child participated in a series of no-interaction conditions to examine whether the behavior would

persist in the absence of social consequences. Because Wendy and Julie both engaged in

potentially harmful SIB, we only tested one or two functions in pairwise comparisons to

minimize the risk of injury. The professional only tested functions that were hypothesized to

maintain the behavior based on her preliminary interview and observations (Iwata & Dozier,

2008). We conducted 5 min conditions with Luke, Brian, and Wendy. Conditions lasted 10 min

for Julie because we provided access to a preferred stimulus for 30 s contingent on the target

behavior; as such, conducting 5-min sessions would not have provided sufficient opportunities to

observe differentiation between the conditions.

In the no-interaction condition, the child had no access to either preferred stimuli or social consequences (e.g., attention); all behaviors were ignored. In the attention condition, the child had ongoing access to preferred toys and the caregiver was present but engaged in other tasks (e.g., reading). When the child engaged in the target behavior, the caregiver provided immediate attention (e.g., saying "stop doing that" while placing her hand on the child's shoulder). In the demand condition, an adult presented requests every 30 s. If the child did not comply within 5 s, the adult initiated a three-step, least-to-most prompting sequence. When the child engaged in the target behavior, the demand or prompt sequence was interrupted for 30 s. In the control condition, the child had noncontingent access to his or her preferred stimuli and the adult provided attention on a 30-s fixed-time schedule. In the tangible condition, the child had access to her preferred stimulus prior to the condition. When the condition began, the adult removed access to the preferred stimulus. Contingent on the occurrence of the target behavior, the child regained access to her preferred stimulus for 30 s. We only conducted the tangible condition when we hypothesized that the target behavior may be maintained by access to a preferred item (Julie only).

During all conditions, the professional measured the duration of hand flapping for Luke, repetitive vocalizations for Brian, and self-pinching for Wendy. For Julie, she measured the frequency of SIB. A second observer measured the target behavior for at least 25% of conditions for each participant for a total 29% of sessions. We calculated interobserver agreement (IOA) using the block-by-block method with10-s intervals (Mudford, Taylor, & Martin, 2009). The mean IOAs were 96% (range: 90%-100%) for Luke's hand flapping, 85% (range: 57%-97%) for Brian's repetitive vocalizations, 96% (range: 89%-99%) for Julie's self-injurious behaviors, and 90% (range: 60%-100%) for Wendy's self-pinching.

2.3. Data Analysis

First, we analyzed the results of the FA and of the QABF to identify the function

maintaining the problem behavior. Second, we compared the results of the FA with the function

proposed by each expert reviewer to examine whether the functions matched or not. An exact

match occurred when the function(s) identified by the expert reviewer and the FA were the same.

A partial match occurred when (a) the expert reviewer identified two functions, one of which

matched the FA, (b) the FA identified two functions, whereas the expert reviewer identified only

one of these two functions, or (c) the two methods being compared yielded two functions, but

only one of the two matched. A disagreement occurred when none of the functions identified

were the same. Third, we applied the same rules when comparing the responses of the expert

reviewers to the results of the QABF. Then, we added the number of exact and partial matches

for each participant and used a binomial distribution to determine the probability of these results

occurring as a result of chance. Specifically, we calculated the probability that random guessing

would produce a number of partial and exact matches equal to or higher than what we observed.

Finally, we collapsed the data with and without the behavior column and compared the results

using Fisher's exact test.

3. Results

Figure 1 shows the results of the FA (left panels) and QABF (right panels) for each

participant. Levels of Luke's hand flapping (top tier) were higher in the no-interaction condition

relative to the other conditions, suggesting that the target behavior was maintained by nonsocial

reinforcement. On the QABF scale, the caregiver endorsed four or five items for all functions.

Given that the severity was highest for nonsocial reinforcement and that behavior maintained by

nonsocial reinforcement may occur under many conditions (Hagopian et al., 1997), we

concluded that this pattern suggested a nonsocial function. Based on her hypotheses, the

professional conducted two pairwise comparisons with Wendy (second tier), which indicated that

self-pinching was maintained by both nonsocial reinforcement and attention. The results of the

QABF unambiguously indicated that Wendy's target behavior was maintained by nonsocial

reinforcement. The pairwise comparison for Julie (third tier) suggested that she engaged in SIB

to access preferred items. Julie's caregiver endorsed four items on the QABF—all social

functions. Thus, the QABF scale suggested that Julie's behavior was socially reinforced (i.e., all

social functions were endorsed four times, but not the nonsocial function); however, the exact

function remained unclear (inconclusive results). For Brian (bottom tier), levels of repetitive

vocalizations remained initially low and undifferentiated across conditions. When we introduced

the series of no-interaction conditions, the target behavior increased and persisted in the absence

of social consequences. The only function endorsed four or five times on the QABF was

nonsocial, consistent with the nonsocial function identified by the FA.

Table 1 presents the results of the qualitative analyses conducted by each expert reviewer

for each of the four participants. For Luke, the analyses of ten of twelve expert reviewers exactly

matched the results of the FA and of the QABF and the responses for the remaining two expert

reviewers were partial matches (binomial test, $p < .0001$). These results strongly suggested that

the expert reviewers accurately identified the function based solely on the ABC narrative

recordings. The twelve expert reviewers also identified at least one of the two functions of

Wendy's target behavior as identified by the FA (binomial test, $p < .05$; note that the probability

differs from the others because it is easier to guess one of the functions when the target behavior

has two functions), indicating that these results were unlikely to be the result of chance.

Similarly for Wendy, nine of twelve experts at least partly agreed with the results of the QABF, which was also statistically significant (binomial test, $p < .05$).

All expert reviewers identified that Julie's behavior was at least partly maintained by tangible reinforcement, which indicated that expert reviewers were performing significantly better than random guessing (binomial test, $p < .0001$). The most common false positive for Julie was the hypothesis of an escape function, which was selected by more than half expert reviewers (i.e., 7). We could not compare the results of the experts with the QABF scale as the analyses of the latter were inconclusive. Finally, only three expert reviewers identified the same function as the FA and QABF for Brian's repetitive vocalizations, which was not statistically significant (binomial test, $p = .92$). Notably, ten expert reviewers identified false positives for attention and five false positives for tangibles. We subsequently compared the number of exact matches, partial matches, and incorrect matches across types of recordings (i.e., with and without the behavior column) using Fisher's exact test, but the results were not statistically significant ($p = .37$). Our analysis suggests that having access to the topography of the behavior did not significantly alter the accuracy of the expert reviewers' qualitative analyses.

4. Discussion

Overall, the results of the study suggest that expert reviewers may accurately identify the function of problem behavior based solely on information contained in ABC narrative recordings. Specifically, the functions identified by the twelve expert reviewers always partially or exactly matched the ones identified by the FA for three of four participants. The experts' analyses also matched the informant-based results at a statistically significant level for two of three participants with a clear QABF. These results were unexpected because prior research shows that the quantitative analysis of ABC continuous data generates hypotheses with no

greater accuracy than random guessing (e.g., Hall, 2005; Tarbox et al., 2009; Thompson & Iwata, 2007). Even though our results indicate that knowing the topography of a target behavior did not statistically improve the accuracy of expert reviewers, one hypothesis that may explain the discrepancy is that the experts considered other variables that are overlooked by statistical analyses. For example, the expert reviewers may have taken into consideration the limitations of descriptive assessments (e.g., the high rate of false positives for attention, the effects of intermittent reinforcement schedules) when analyzing the ABC narrative recordings. Alternatively, the expert reviewers may have attributed more weight to certain recorded events than others, which would yield results that differ from a statistical analysis that assigns the same weight to all recorded events.

For one of the participants (i.e., Brian), the function identified using the ABC narrative recordings partially or exactly matched the ones identified by both the FA and QABF for only three expert reviewers. One potential explanation is that the target behavior may have been correlated with events (e.g., attention) that did not maintain engagement in the target behavior. An alternative explanation is that the caregiver only recorded the behavior in specific situations (e.g., when she was playing with the child), which could have biased the analysis. Another noteworthy pattern is that the false positives identified for each participant were often similar across expert reviewers. For example, ten of twelve expert reviewers indicated that Brian's target behavior was at least partly maintained by attention. Similarly, seven of twelve expert reviewers suggested that Julie's target behavior was also maintained by escape. These patterns suggest that the expert reviewers may have employed similar strategies to identify the function of the target behavior.

To our knowledge, our study is the first to examine whether professionals can accurately identify the function of problem behavior using ABC narrative recordings. Given that the use of ABC narrative recordings appears to be widespread amongst professionals, this study provides support for its use. The analysis in which we removed the behavior column further extends research by suggesting that topography did not strongly influence the decisions taken by professionals. These results were surprising because studies have suggested that some functions may be more likely for certain topographies (e.g., Matson et al., 1999; Medeiros, Rojahn, Moore, & Ingen, 2013). As such, partly relying on the topography for some types of behavior may improve the validity of the analysis conducted by the professional. Finally, our comparisons of the functions identified by the FA and the QABF are consistent with other studies (e.g., Paclawskyj et al., 2001; Tarbox et al., 2009; Watkins & Rapp, 2013) and provide further support for the use of checklist-based assessments; that is, the function identified by the QABF exactly matched the one identified by the FA for two of the participants and partially matched for another. In particular, our results support those of Watkins and Rapp (2013) with respect to problem behaviors maintained by nonsocial reinforcement.

The main clinical implication of our study is that the results may support the use of ABC narrative recordings to generate hypotheses about the function of problem behavior, which can be later confirmed using another type of assessment. For three of four participants, the analysis conducted by all expert reviewers identified at least one of the functions of the target behavior. Furthermore, the expert reviewers were as accurate as the QABF in identifying the function of problem behavior. Because our study was the first to examine the accuracy of the analysis of expert reviewers using ABC narrative recordings, our results should be viewed as preliminary. Until the results are replicated by other researchers, we do not recommend using ABC narrative

recordings as a standalone assessment method. Nonetheless, ABC narrative recordings may be used to (a) generate hypotheses, (b) provide support for a function identified via another assessment method, and (c) help inform the choice of other methods of observation.

The results of the study are limited in at least three ways. First, we did not measure the fidelity of the narrative data collected by caregivers. As such, caregivers may have failed to report all occurrences of the target behavior or may have described some occurrences incorrectly. Even though we trained the caregivers using well-established practices in the field (i.e., behavioral skills training), caregivers may have interrupted their data collecting when they were busy with other tasks. This limitation notwithstanding, the data collection procedures employed in this study closely approximate how parents collect data in applied settings—a finding that potentially increases the ecological validity of our results. Second, we only tested one function for Julie during the FA. This limitation may have produced false negatives for other functions. From an ethical standpoint, we could not test multiple functions because her self-injury was extremely severe and thus posed a danger to her physical health. Nevertheless, the function with the highest severity score in the QABF was tangible reinforcement, consistent with the results of the FA. Finally, the target behaviors of three of four participants were at least partly maintained by nonsocial reinforcement. Thus, it is unclear whether the analysis of ABC narrative recordings would be as accurate in the identifying the function of socially reinforced behavior.

Future research should systematically replicate our procedures with a larger number of participants with a wider variety of target behaviors and behavioral functions. To examine the influence of correlated events on the professional's analysis, researchers should also collect and analyze ABC continuous data as part of the study. If the ABC continuous and narrative data analyses produced different results, we could conclude that the professionals were relying on

additional variables to make decisions. In the same vein, researchers should identify the variables

that influence the professional's decision making in order to improve accuracy. Finally, future

research should examine whether using ABC narrative recordings alone or in combination with

other assessment methods is a cost effective means by which to identify the function of problem

behavior. In the end, whether or not collecting ABC narrative data reduces the time and

resources involved in identifying the function of problem behavior will largely dictate its clinical

utility.

Acknowledgments

References

Desrochers, M. N., Hile, M. G., & Williams-Mosely, T. L. (1997). Survey of functional

assessment procedures used with individuals who display mental retardation and severe

problem behaviors. *American Journal on Mental Retardation, 101*, 535-546.

Durand, V.M. & Crimmins, D.B. (1988). Identifying the variables maintaining self –injurious

behavior. *Journal of Autism and Developmental Disorders*, 18, 99-117.

Ellingson, S. A., Miltenberger, R. G., & Long, E. S. (1999). A survey of the use of functional

assessment procedures in agencies serving individuals with developmental disabilities.

Behavioral Interventions, 14, 187-198.

Fox, J., & Davis, C. (2005). Functional behavior assessment in schools: Current research

findings and future directions. *Journal of Behavioral Education, 14,* 1-4.

Freeman, K. A., Walker, M., & Kaufman, J. (2007). Psychometric properties of the Questions

About Behavioral Function Scale in a child sample. *American Journal on Mental

Retardation, 112,* 122-129.

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,* 313-326.

Hall, S.S. (2005). Comparing descriptive, experimental and informant-based assessments of

problem behaviors. *Research in Developmental Disabilities, 26*, 514-526.

Hanley, G. P. (2012). Functional assessment of problem behavior: Dispelling myths, overcoming

implementation obstacles, and developing new lore. *Behavior Analysis in Practice, 5,* 54-

72.

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: a

    review. *Journal of Applied Behavior Analysis, 36,* 147-185.

Healy, O., Brett, D., & Leader, G. (2013). A comparison of experimental functional analysis and

    the Questions About Behavioral Function (QABF) in the assessment of challenging

    behavior of individuals with autism. *Research in Autism Spectrum Disorders*, *7*, 66-81.

Horner, R. H., & Carr, E. G. (1997). Behavioral support for students with severe disabilities

    functional assessment and comprehensive intervention. *The Journal of Special Education,*

    31, 84-104.

Ingram, K., Lewis-Palmer, T., & Sugai, G. (2005). Function-based intervention planning:

    Comparing the effectiveness of FBA function-based and non-function-based intervention

    plans. *Journal of Positive Behavior Interventions, 7*, 224-236.

Iwata B. A. & DeLeon I. G. (1995). *The Functional Analysis Screening Tool (FAST).* University

    of Florida, FL.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a

    functional analysis of self-injury. *Journal of Applied Behavior Analysis*, *27,* 197-209.

    (Reprinted from: *Analysis and Intervention in Developmental Disabilities*, 2, 3-20, 1982).

Iwata, B. A., & Dozier, C. L. (2008). Clinical application of functional analysis methodology.

    Behavior Analysis in Practice, 1, 3-9.

Iwata, B. A., Pace, G. M., Cowdery, G. E., & Miltenberger, R. G. (1994). What makes extinction

    work: an analysis of procedural form and function. *Journal of Applied Behavior Analysis*,

    27, 131-144.

Kahng, S., Iwata, B. A., Fischer, S. M., Page, T. J., Treadwell, K. R., Williams, D. E., & Smith,

    R. G. (1998). Temporal distributions of problem behavior based on scatter plot analysis.

    Journal of Applied Behavior Analysis, 31, 593-604.

Lafasakis, M., & Sturmey, P. (2007). Training parent implementation of discrete-trial teaching:

    Effects on generalization of parent teaching and child correct responding. *Journal of*

    Applied Behavior Analysis, 40, 685-689.

Lerman, D. C., Hovanetz, A., Strobel, M., & Tetreault, A. (2009). Accuracy of teacher-collected

    descriptive analysis data: A comparison of narrative and structured recording formats.

    Journal of Behavioral Education, 18, 157-172.

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.

Love, J. R., Carr, J. E., Almason, S. M., & Petursdottir, A. I. (2009). Early and intensive

    behavioral intervention for autism: A survey of clinical practices. *Research in Autism*

    Spectrum Disorders, 3, 421-428.

Luiselli, J. K. (2011). Training parents and other care providers. In J. K. Luiselli (Ed.), *Teaching*

    and behavior support for children and adults with autism spectrum disorder: A

    practitioner's guide (pp. 212-216). New York, NY: Oxford University Press.

Marion, S. D., Touchette, P. E., & Sandman, C. A. (2003). Sequential analysis reveals a unique

    structure for self-injurious behavior. *American Journal on Mental Retardation, 108,* 301-

    313.

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. *Research in Developmental Disabilities, 20,*

163-175.

Matson, J. L., & Boisjoli, J. A. (2007). Multiple versus single maintaining factors of challenging

behaviours as assessed by the QABF for adults with intellectual disabilities. *Journal of*

Intellectual and Developmental Disability, 32, 39-44.

Matson, J. L., & Minshawi, N. F. (2007). Functional assessment of challenging behavior:

Toward a strategy for applied settings. *Research in Developmental Disabilities*, *28*, 353-

361.

Matson, J. L., Tureck, K., & Rieske, R. (2012). The Questions About Behavioral Function

(QABF): Current status as a method of functional assessment. *Research in*

Developmental Disabilities, 33, 630-634.

Matson, J. L., & Vollmer, T. R. (1995). *User's guide – Questions About Behavioral Function*

(QABF). Baton Rouge, LA: Scientific Publishers.

McComas, J. J., Moore, T., Dahl, N., Hartman, E., Hoch, J., & Symons, F. (2009). Calculating

contingencies in natural environments: Issues in the application of sequential analysis.

Journal of Applied Behavior Analysis, 42, 413-423.

Medeiros, K., Rojahn, J., Moore, L. L., & Ingen, D. J. (2013). Functional properties of behaviour

problems depending on level of intellectual disability. *Journal of Intellectual Disability*

Research. Advanced online publication. doi:10.1111/jir.12025

Miles, N. I., & Wilder, D. A. (2009) The effects of behavioral skills training on caregiver

implementation of guided compliance. *Journal of Applied Behavior Analysis, 42,* 405-

410.

Mudford, O. C., Taylor, S. A., & Martin, N. T. (2009). Continuous recording and interobserver

      agreement algorithms in the Journal of Applied Behavior Analysis (1995-2005). *Journal*

      of Applied Behavior Analysis, 42, 165-169.

Neef, N. A., & Peterson, S. M. (2007). Functional behavior assessment. In J. O. Cooper, T. E.

      Heron, & W. L. Heward, *Applied behavior analysis* (pp. 500-524). Columbus, OH:

      Merrill Publishing Co.

Paclawskyj, T. R., Matson, J. L., Rush, K. S., Smalls, Y., & Vollmer, T. R. (2000). Questions

      about behavioral function (QABF): A behavioral checklist for functional assessment of

      aberrant behavior. *Research in Developmental Disabilities, 21,* 223-229.

Paclawskyj, T. R., Matson, J. L., Rush, K. S., Smalls, Y. & Vollmer, T. R. (2001). Assessment of

      convergent validity of the Questions About Behavioral Function scale with analogue

      functional analysis and the Motivation Assessment Scale. *Journal of Intellectual*

      Disability Research, 45, 484-494.

Seiverling, L., Williams, K., Sturmey, P., & Hart, S. (2012). Effects of behavioral skills training

      on parental treatment of children's food selectivity. *Journal of Applied Behavior Analysis,*

      45, 197-203.

Sloman, K. N. (2010). Research trends in descriptive analysis. *The Behavior Analyst Today, 11,*

      20-35.

Smith, C. M., Smith, R. G., Dracobly, J. D., & Pace, A. P. (2012). Multiple-respondent anecdotal

      assessments: An analysis of interrater agreement and correspondence with analogue

      assessment outcomes. *Journal of Applied Behavior Analysis*, *45*, 779-795.

St. Peter, C. C., Vollmer, T. R., Bourret, J. C., Borrero, C. S., Sloman, K. N., & Rapp, J. T.

(2005). On the role of attention in naturally occurring matching relations. *Journal of*

Applied Behavior Analysis, 38, 429-443.

Tarbox, J., Wilke, A. E., Najdowski, A. C., Findel-Pyles, R. S., Balasanyan, S., Caveney, A. C.,

... & Tia, B. (2009). Comparing indirect, descriptive, and experimental functional

assessments of challenging behavior in children with autism. *Journal of Developmental*

and Physical Disabilities, *21*, 493-514.

Thompson, R.H. & Iwata, B.A. (2007) A comparison of outcomes from descriptive and

functional analyses of problem behavior. *Journal of Applied Behavior Analysis, 40,* 333-

338.

Touchette, P. E., MacDonald, R. F., & Langer, S. N. (1985). A scatter plot for identifying

stimulus control of problem behavior. *Journal of Applied Behavior Analysis, 18,* 343-351.

Vollmer, T. R., Marcus, B. 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.

Watkins, N., & Rapp, J. T. (2013). The convergent validity of the Questions About Behavioral

Function scale and functional analysis for problem behavior displayed by individuals

with autism spectrum disorder. *Research in Developmental Disabilities*, *34*, 11-16.

Wilke, A. E., Tarbox, J., Dixon, D. R., Kenzer, A. L., Bishop, M. R., & Kakavand, H. (2012).

Indirect functional assessment of stereotypy in children with autism spectrum disorders.

Research in Autism Spectrum Disorders, *6*, 824-828.

Zaja, R. H., Moore, L., Van Ingen, D. J., & Rojahn, J. (2011). Psychometric comparison of the

functional assessment instruments QABF, FACT and FAST for self-injurious, stereotypic

and aggressive/destructive behaviour. *Journal of Applied Research in Intellectual Disabilities*, *24*, 18-28.
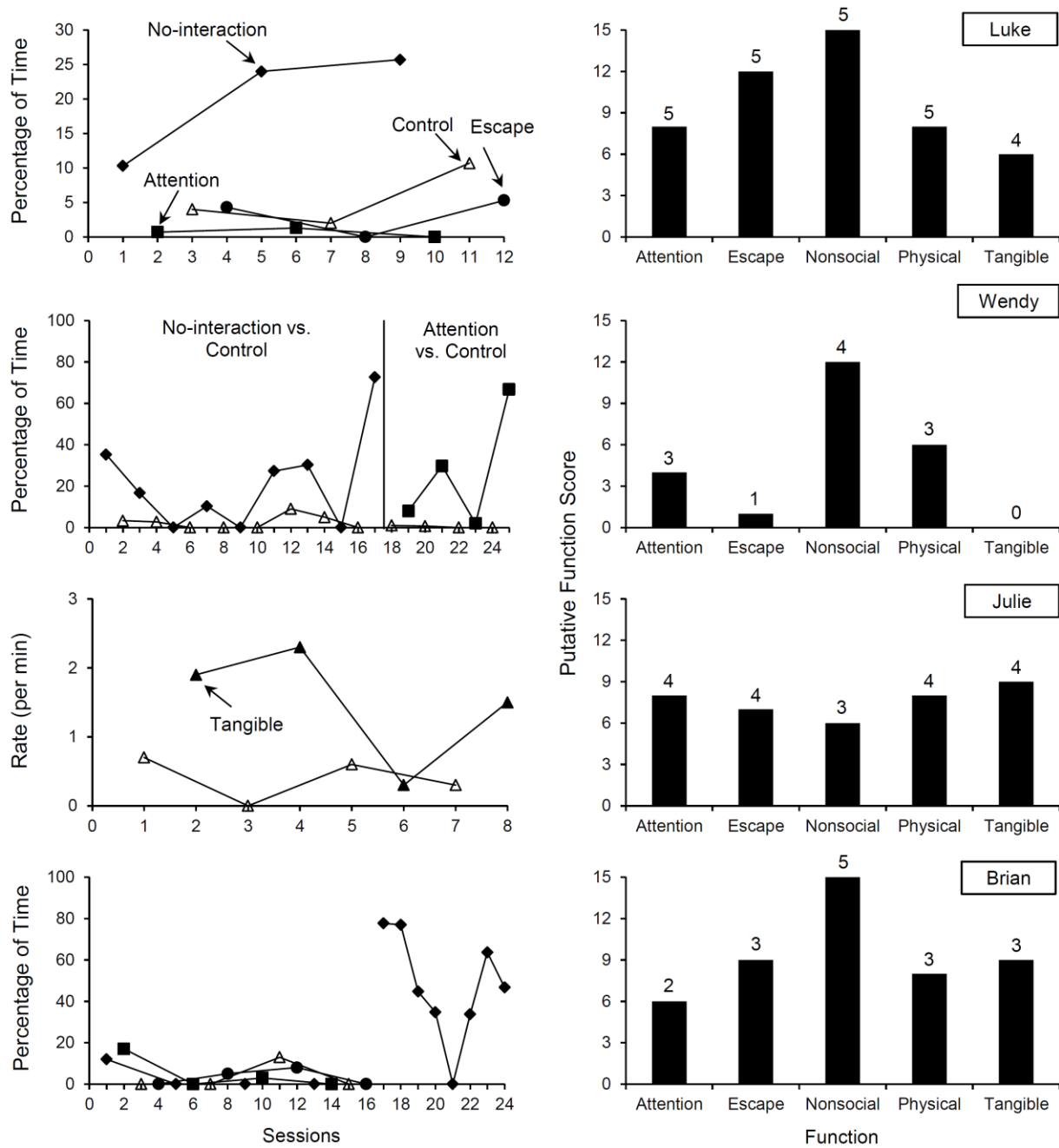
Figure 1. Results of the functional analyses presented in line graphs for Luke, Wendy, Julie, and

Brian (left panels). The corresponding QABF severity scores are presented as bars (right panels).

Numbers above each bar indicate the number of items endorsed by the caregiver.