CORRESPONDENCE BETWEEN OUTCOMES OF BRIEF AND EXTENDED FUNCTIONAL ANALYSES

SUNGWOO KAHNG AND BRIAN A. IWATA

UNIVERSITY OF FLORIDA

We compared results obtained from 50 sets of functional analysis data from assessments of self-injurious behavior (SIB), 35 of which showed clear response patterns and 15 of which were undifferentiated, with those obtained from two abbreviated methods of assessment: (a) a brief functional analysis, consisting of the first session of each condition from the full functional analysis, and (b) a within-session analysis, in which data from the brief analysis were regraphed to show minute-by-minute changes in response rates during a session. Results indicated that outcomes of the brief and within-session analyses corresponded with those of the full functional analyses in 66.0% and 68.0% of the cases, respectively. Further examination of results indicated a tendency for the brief analysis to identify a large proportion of positive cases (both true and false positives) and for the within-session analysis to identify a large proportion of negative cases (true and false negatives).

DESCRIPTORS: functional analysis, self-injurious behavior, aggression

---

Functional analysis methodology has emerged as a powerful tool in the assessment and treatment of severe behavior disorders over the past 15 years. Results from a number of studies have demonstrated that interventions based on the identification of behavioral function are more likely to be effective than those that are arbitrarily chosen. This finding has been demonstrated in individual cases (Carr & Durand, 1985; Day, Rea, Schussler, Larsen, & Johnson, 1988; Iwata, Pace, Cowdery, & Miltenberger, 1994; Repp, Felce, & Barton, 1988) as well as across large client populations (Iwata, Pace, Dorsey, et al., 1994).

Although the advantages of functional analysis methodology have been clearly established, its implementation may not always be practical. For example, thorough functional analyses may be infeasible under time constraints imposed in certain settings (e.g., outpatient clinics). Under such conditions, a procedure described as the "brief functional analysis" has been proposed as an alternative method of assessment (Northup et al., 1991). In the first phase of this assessment, an individual typically is exposed once (sometimes twice) to a subset of test and control conditions during 5- or 10-min sessions, resulting in a multielement design with a single data point in each condition. A "contingency reversal phase," in which contingencies are placed on an alternative response, is usually conducted in a subsequent phase of the analysis.

In a large-scale evaluation ($N = 79$) of the brief functional analysis, Derby et al. (1992) reported that instances of the target behavior were observed in only two thirds of the cases and that the behavior's maintaining variable was identified in approximately 75% of those cases (for an overall "hit rate" of about 50%). The most likely explanation for their inability to identify a behavioral function in a higher proportion of cases is that limited exposure to the assessment conditions did not allow participants to either

contact or discriminate differences among contingencies. Another possibility is that discrimination occurred within a session but was not reflected in the data; that is, the overall session average may have obscured a within-session trend produced by either acquisition or extinction.

To the extent that behavior change occurs during the course of a single session, examination of response rates during brief intervals within a session may allow detection of such trends. Vollmer, Iwata, Zarcone, Smith, and Mazaleski (1993) compared the results of these within-session analyses with those obtained from full functional analyses for 4 individuals who engaged in self-injurious behavior (SIB) and obtained correspondence between the two sets of data in three of the cases. Based on these results, Vollmer, Marcus, Ringdahl, and Roane (1995) proposed a hierarchical series of assessments progressing from brief to full functional analyses. The starting point in their sequence was a within-session analysis of data from single exposures to each of several assessment conditions. Six of their 20 participants showed differential responding (i.e., higher rates of behavior under one condition) during these brief within-session analyses. However, the overall degree of correspondence across assessment (or data analysis) procedures was undetermined because the progression of analyses was terminated once differential responding was observed.

The purpose of this study was to examine correspondence among results from assessments based on full functional analyses, brief functional analyses, and within-session analyses. To accomplish this goal, we first conducted full functional analyses of SIB and aggression. Subsequently, we examined rates of responding during the first session of each analysis, graphed as overall session averages as well as within-session values, and compared these data to those obtained from the full functional analyses.

METHOD

Participants and Setting

Participants were 50 individuals (23 females, 27 males) who lived in a state residential facility and who had been referred for assessment and treatment of SIB or aggression. All participants were diagnosed with mental retardation (moderate, 4%; severe, 6%; profound, 90%). Their ages ranged from 20 to 68 years ($M = 37.5$).

All sessions were conducted in individual therapy rooms at a specialized program for the assessment and treatment of SIB located on the grounds of the facility. The rooms were equipped with furniture and other materials necessary to conduct the various experimental conditions.

Response Measurement and Reliability

Topographies of SIB and aggression, including specific response definitions, are listed in Table 1. All data were collected on handheld computers (Assistant Model AST102) and were summarized as either number of responses per minute or as the percentage of 10-s intervals during which responding occurred. Interobserver agreement was assessed by having a second observer collect data simultaneously with, but independent of, the primary observer. Agreement percentages were calculated in one of two ways, based on interval-by-interval comparison of observers' records. For frequency measures, the smaller number of responses was divided by the larger number in each interval; these fractions were then summed, divided by the total number of intervals in the session, and multiplied by 100%. For interval measures, the number of intervals containing scoring agreements was divided by the total number of intervals and multiplied by 100%. The percentage of sessions during which reliability was assessed ranged from 10% to 75% ($M = 32.2\%$) across participants. Mean interobserver agreement for

Table 1
Response Topographies, Definitions, and Distributions

Topography	Definition	% of participants[a]
SIB		
Biting	Closure of upper and lower teeth on any part of the body	18.0 ($n = 9$)
Body hitting	Audible or forceful contact of one body part against another	24.0 ($n = 12$)
Eye poking	Forceful contact of a finger within the ocular area	4.0 ($n = 2$)
Foot stomping	Audible or forceful contact of the foot to the floor	2.0 ($n = 1$)
Head hitting	Audible or forceful contact of a body part against the head or face	44.0 ($n = 22$)
Head banging	Audible or forceful contact of the head against a stationary object	22.0 ($n = 11$)
Hand mouthing	Contact of fingers or hand against lips or tongue	24.0 ($n = 12$)
Scratching	Raking the skin with fingernails or rubbing against objects	14.0 ($n = 7$)
Aggression		
Hitting	Audible or forceful contact of a body part against another person	4.0 ($n = 2$)
Biting	Closure of upper and lower teeth on any part of another person's body	2.0 ($n = 1$)
Scratching	Raking the skin with fingernails or rubbing against another person	4.0 ($n = 2$)

[a] Numbers and percentages do not correspond to the total sample size because a number of participants exhibited more than one topography.

individuals ranged from 82.1% to 100% ($M = 92.5\%$).

Functional Analysis

Functional analyses based on procedures described by Iwata, Dorsey, Slifer, Bauman, and Richman (1994/1982) were conducted using a combination of the following conditions: attention, tangible, demand, alone, and play. One individual was also exposed to a variation of the demand condition involving escape from loud noise (music). Sessions were sequenced in a multielement design and were 10 to 15 min in duration. Two to eight sessions were conducted per day, usually 5 days per week.

Attention. This condition assessed behavioral sensitivity to positive reinforcement in the form of attention. The experimenter and participant were in a room with leisure items, to which the participant had free access. The experimenter delivered approximately 5 s of attention to the participant (e.g., "Don't do that, you'll hurt yourself")

contingent on the occurrence of a target behavior, but ignored the participant otherwise.

Tangible. This was a variation of the attention condition. Prior to the session, the experimenter allowed the participant free access to a leisure item or food for a brief time. When the session began, the experimenter removed the item but returned it for 30 s (or delivered a small piece of food) contingent on the occurrence of a target behavior.

Demand. This condition assessed sensitivity to negative reinforcement in the form of escape from task demands. The experimenter presented instructional tasks to the participant, usually on a fixed-time (FT) 30-s schedule using a three-prompt sequence (an initial instruction, followed by a gestural prompt and physical guidance as needed). If the participant complied with an instruction, the experimenter delivered praise. If the individual exhibited a target behavior at any time during the sequence, the experimenter removed the task materials and ig-

nored the participant until the next scheduled trial. We observed that 1 individual's behavior during this condition seemed to escalate in the presence of noise, and the possibility that SIB was maintained by escape from noise was assessed during a variation of the demand condition. The experimenter played country music on a cassette player throughout the session but terminated it for 30 s when the participant engaged in SIB.

Alone. This condition was implemented only with individuals who engaged in SIB to determine whether SIB persisted in the absence of social reinforcement. The participant was alone in a room without any leisure items, and no social consequences were delivered following occurrences of SIB.

Play. This condition served as the control. Leisure items were available throughout the session, and the experimenter delivered noncontingent attention to the participant on an FT 30-s schedule. No instructions were presented, and no social consequences were delivered following occurrences of target behaviors.

Data Analysis and Interpretation

Graphs showing session-by-session values were prepared for each complete set of functional analysis data. From these, brief functional analyses were created by plotting data from only the first session of each condition. Finally, data from the first session of each condition were replotted on a minute-by-minute basis to form the within-session analyses.

All data (with identifying information removed) were examined by seven to eight behavior analysts experienced in conducting functional analyses. The evaluators first reviewed each set of full functional analysis data and reached a consensus about the variable maintaining a given participant's aberrant behavior: positive reinforcement (higher responding during the attention or access to tangible items condition), negative reinforce-

Table 2

Overall Percentage and Number of Cases for Which Results of the Brief and Within-Session Analyses Matched or Did Not Match Results of the Full Functional Analyses ($N = 50$)

Type of analysis	Correspondence with full analysis	
	Percentage	Number of cases
Brief functional analysis	66.0	33
Within-session analysis	68.0	34
Either analysis	80.0	40
Both analyses	54.0	27

ment (higher responding during the escape from demand or loud noise condition), and automatic reinforcement (higher responding during the alone condition). Alternatively, the evaluators agreed that the results were uninterpretable. These decisions were based on examination of the data with respect to changes in level, trend, or both within and across assessment conditions. The evaluators then examined each of the 50 brief and within-session functional analyses and reached a consensus, using the same criteria, about the variable maintaining the aberrant behavior, or agreed that the data were uninterpretable.

Interpretations based on the brief and within-session functional analyses were compared to those based on the full functional analyses, and the number of participants for whom interpretations based on the brief or within-session functional analysis matched those of the full analyses was determined. A correspondence percentage for the brief and within-session analyses was calculated by dividing the number of agreements (with the full analyses) by the total number of agreements plus disagreements and multiplying by 100%.

RESULTS

As shown in Table 2, there was little difference in overall correspondence when re-

Table 3
Percentage (and Number) of Hits and Misses for the Brief and Within-Session Analyses Based on Outcomes
of the Full Functional Analyses

| | Results of full functional analysis | | | |
| | Clear (*n* = 35) | | Unclear (*n* = 15) | |
Type of analysis	True positive	False negative	True negative	False positive
Brief functional analysis	77.1 (27)	22.9 (8)	40.0 (6)	60.0 (9)
Within-session analysis	62.9 (22)	37.1 (13)	80.0 (12)	20.0 (3)
Both analyses	60.0 (21)	40.0 (14)	40.0 (6)	60.0 (9)
Either analysis	80.0 (28)	20.0 (7)	80.0 (12)	20.0 (3)

sults from brief and within-session analyses were compared with those from full analyses. Interpretations of the brief and within-session analyses corresponded with those of the full analyses in 66% and 68% of the cases, respectively. When *either* of the abbreviated analyses was used as the basis for comparison, correspondence with the full analyses was very high (80%), although *both* of the abbreviated analyses showed correspondence only about half the time (54%). In addition, there were six cases (12%) in which only the results of the brief analysis corresponded with those of the full analysis and seven cases (14%) in which only the results of the within-session analysis corresponded with those of the full analysis.

The data in Table 2 provide rough estimates of the overall predictive power of the brief and within-session analyses but they do not reflect the sensitivity of either analysis to positive or negative assessment outcomes. Both "hit" (true positive, true negative) and "miss" (false positive, false negative) rates can be determined by examining correspondence for the brief and within-session analyses when the full functional analyses are divided based on outcome (i.e., clear vs. unclear results). These data are presented in Table 3. When results of the full functional analysis were clearly differentiated (positive), correspondence (true positive rate) for the brief analysis (77.1%) was noticeably higher than for the within-session analysis (62.9%).

However, when results of the full functional analysis were undifferentiated (negative), correspondence (true negative rate) for the within-session analysis (80.0%) was much higher than for the brief analysis (40.0%). Thus, it appeared that the brief analysis was more likely to identify a function if one were present, whereas the within-session analysis was more likely to identify the absence of a clear function.

Figure 1 shows representative cases in which both the brief and within-session analyses did and did not show correspondence with their clearly differentiated full functional analyses. Hans' full functional analysis showed clear results from the outset, in that his SIB occurred almost exclusively in the tangible condition (positive reinforcement). That being the case, both of the abbreviated analyses showed similarly clear results. Timmy's full functional analysis indicated that his SIB was maintained by escape from demands (negative reinforcement). However, similar rates of SIB occurred during the first session of both the demand and alone conditions, which were reflected in the brief and within-session analyses.

The top panel of Figure 2 shows a case in which only the brief analysis corresponded with a clearly differentiated full analysis. Both of these analyses indicated that Britt's SIB was maintained by attention (positive reinforcement), although results were clearer in the full analysis. However, examination of
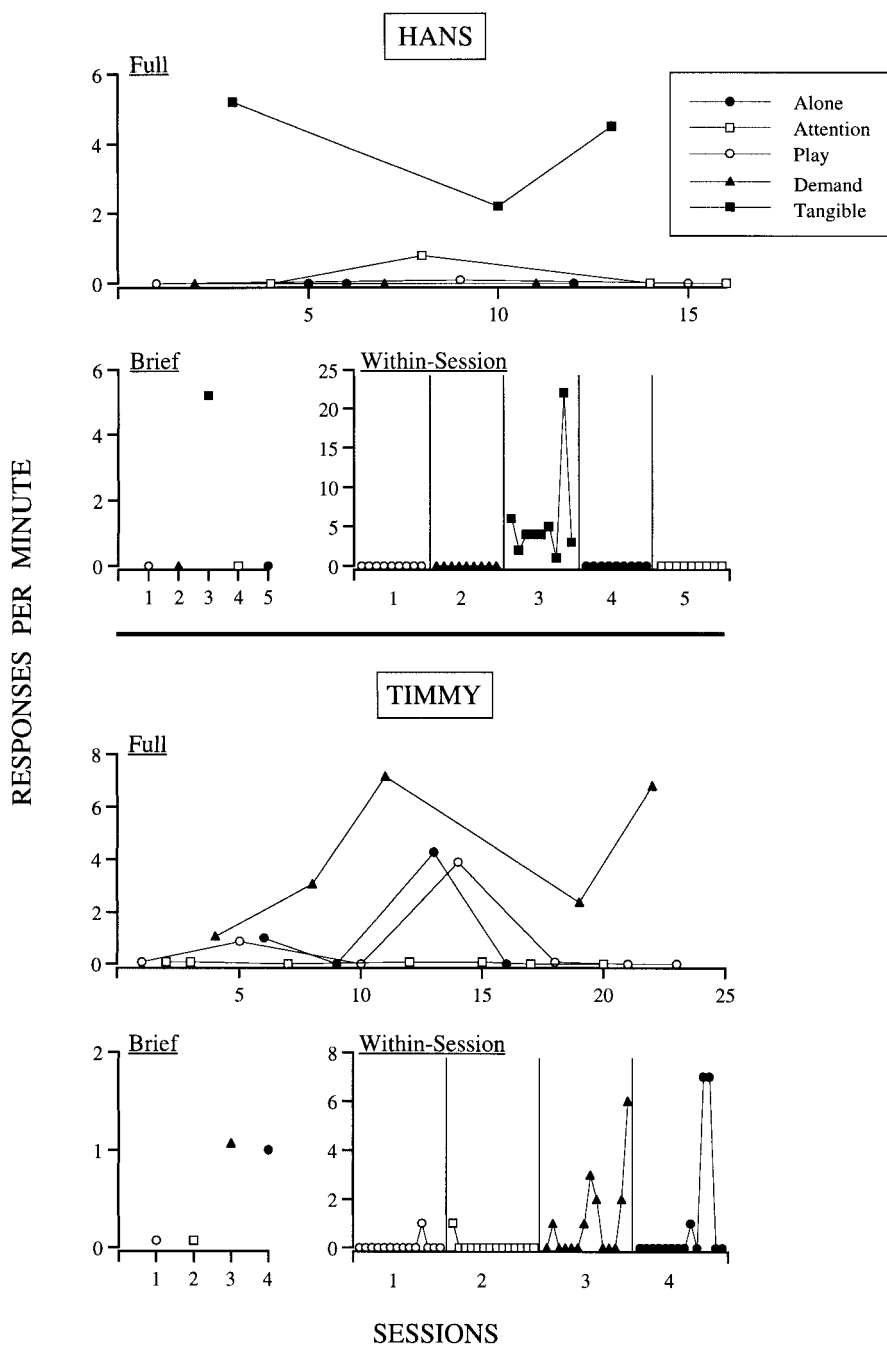
Figure 1.   Representative cases in which both the brief and within-session analyses showed correspondence (top panel, Hans) and did not show correspondence (bottom panel, Timmy) with their respective full analyses.

Britt's within-session patterns of responding failed to identify a clear behavioral function because moderate levels of SIB throughout the attention session were somewhat ob-scured by periodic spikes in SIB at various points during other sessions.

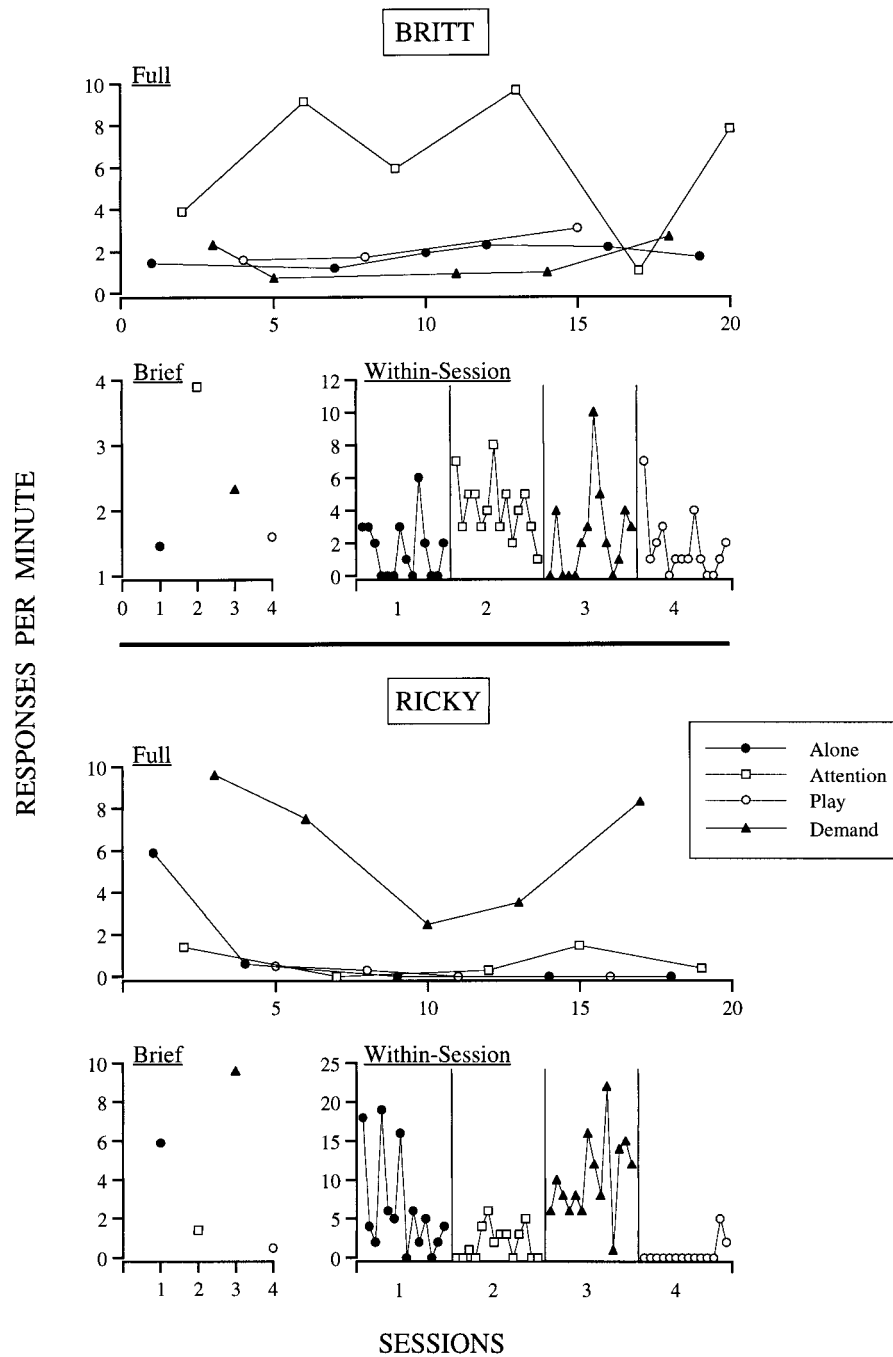The bottom panel of Figure 2 shows the sole case in which only the within-session

Figure 2.   Cases in which only the brief analysis corresponded with the full analysis (top panel, Britt) and in which only the within-session analysis corresponded with the full analysis (bottom panel, Ricky).

analysis matched a clear full analysis. Ricky consistently engaged in higher rates of SIB during the demand condition (full analysis). However, a high rate of SIB was also ob-

served during the first session of the alone condition. During Ricky's first demand and alone sessions, opposite trends were observed (within-session analysis): SIB decreased
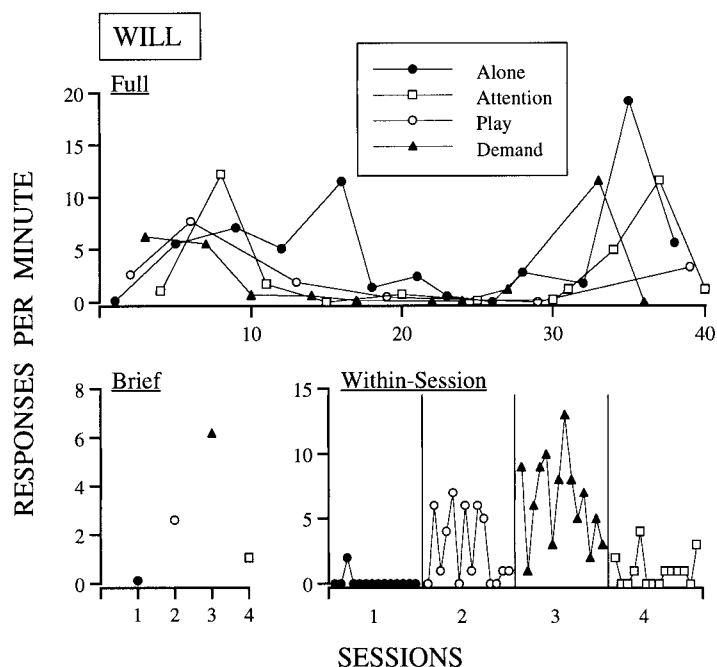
Figure 3. A case in which neither of the abbreviated analyses corresponded with an undifferentiated full analysis.

throughout the alone session but increased throughout the demand session. These trends were evident in the within-session analysis but not in the brief analysis.

Finally, Figure 3 shows a case in which results of the brief and within-session functional analyses differed from that of an undifferentiated full functional analysis. Data from the full analysis showed high response rates in every condition periodically, but responding was highest in the demand session during the first set of conditions. This unrepresentative pattern was reflected in both the brief and within-session analyses.

DISCUSSION

Results of the present study indicated that outcomes obtained from brief functional analyses, and from within-session response patterns contained in such analyses, showed a similar and moderately high degree of overall correspondence (about two thirds) with outcomes based on full functional analyses.

Correspondence between the brief and full functional analyses was even higher (77.1%) when results of the full functional analyses were clear, but was much lower (40%) when results of the full functional analyses were unclear. These data suggest that the brief analysis produced a high proportion of true positives but also a high proportion of false positives, reflecting a general tendency for the brief analysis to identify a behavioral function, regardless of whether one was apparent in the full functional analysis. Opposite results were obtained for the within-session analyses, which yielded a much higher rate of true negatives and a somewhat higher rate of false negatives than did the brief analysis, reflecting a tendency for the within-session analysis to produce unclear (negative) results.

This apparent difference in the types of errors produced by the two abbreviated analyses is probably a direct function of the method of data display. The brief analysis shows the total amount of responding dur-

ing a session. When behavior occurs in more than one assessment condition, as is often the case, an initially higher rate in one condition will be interpreted as a potential true positive, although this difference may or may not be sustained across repeated exposures to assessment conditions (cf. Britt's data in Figure 2 with Will's data in Figure 3); hence, the high proportion of both true and false positives. By contrast, within-session data analysis provides greater detail with respect to response variability and trends. But a high degree of unsystematic variability in several conditions, which may be reflected as periodic spiking, might obscure differences in overall response rates between conditions (see Figure 2, Britt). Thus, in addition to the fact that within-session analyses are based on a single exposure to assessment conditions, the within-session analysis may not yield clear results unless response rates differ markedly throughout a session (see Figure 1, Hans) or unless variable response rates show different trends (see Figure 2, Ricky); hence, the high proportion of both true and false negatives.

In actual practice, it is highly unlikely that a within-session analysis would be conducted without first examining overall response rates obtained in the brief functional analysis because the former are derived from the latter. At issue, then, is the extent to which a within-session analysis improves predictive power once a brief analysis has been conducted. Given the high true positive rate of the brief analysis (77.1%), further examination of within-session response patterns might not be helpful in identifying many more cases. In fact, there was only one case in which the within-session analysis identified a clear function when the brief functional analysis did not (see Figure 2, Ricky). However, that single case was significant because the within-session analysis revealed opposite trends in two data sets that could not be detected by examining the session totals,

and the extent to which such trends exist is unknown unless data are examined using both types of analysis.

A less obvious but major benefit of within-session data analysis may be to reduce the high false positive rate of the brief analysis. When results of the full functional analyses were unclear (indicating the absence of an apparent function), the brief analysis incorrectly identified a function in 60% of the cases but the within-session analysis did so in only 20% of the cases. In many of these cases, high overall response rates in one condition were primarily a function of isolated responding (spikes) during a few of the intervals, which was evident only when examining within-session response patterns.

The above conclusions must be tempered somewhat due to several limitations in the present study. First, all of the analyses were derived from data based on session lengths of 15 min. In a number of studies in which brief functional analyses have been used, session length was limited to 5 min (e.g., Derby et al., 1997). Had the brief analyses in this study been based on 5-min sessions, it is likely that correspondence with the full functional analyses would have been lower, as can be seen through an examination of Ricky's data (Figure 2, within-session analysis). His SIB was much higher during the first half of the alone condition than it was during the first half of the demand condition, which would have affected interpretation of both of the brief analyses had session duration been limited to 5 min. To the extent that discrimination (or extinction) occurs during a session, the use of very brief sessions may not allow sufficient contact with the contingencies whose effects (or lack thereof) are being evaluated.

A second limitation was that comparisons made in this study were based on brief functional analyses containing no reversal probes because data for the brief analyses were taken from existing data sets (i.e., full functional

analyses). Such probes, which consist of a reversal of the contingency (via differential reinforcement, extinction, etc.) suspected of maintaining the problem behavior, are often conducted as an adjunct to the brief functional analyses (e.g., Derby et al., 1992; Harding, Wacker, Cooper, Millard, & Jensen-Kovalan, 1994). It is possible that the false positive rate of the brief analysis would have been lower had we included reversal probes.

Third, it is unclear what effect previous exposure to assessment contingencies may have had on the results of this study. We were unable to accurately determine the history of exposure to previous functional analyses for each of our participants, although any such analyses would have been conducted at least several months prior to their inclusion in the present study. Nevertheless, a few of the participants may have shown immediate response differentiation in the first session of each assessment condition due to prior exposure to the same contingencies. If so, the proportion of true positives for both abbreviated forms of assessment would have been lower than that reported here.

Finally, all of the interpretations presented in the results were based on group consensus following examination of graphed data. Although visual analysis is the most widely used method for interpreting data from single-subject analyses, it is possible that slightly different outcomes would have been obtained had we adopted a different strategy. For example, Hagopian et al. (1997) proposed a number of formal criteria for interpreting the results of functional analyses. We did not use these rules in the present study because none of data for the brief analyses, and only some of the data for the full analyses, contained the requisite number of data points (10 per condition). In addition, it is important to note that the decision-making

rules proposed by Hagopian et al. were themselves based on consensus opinion.

In spite of the above limitations, our results suggest that the brief functional analysis, based on single exposures to contingencies that may maintain problem behavior, may be adequate when circumstances do not permit repeated observation of behavior under multiple assessment conditions. Although the hit rate (proportion of true positives and true negatives) obtained with the brief analysis was not exceedingly high, it was sufficient to provide justification for its use when time constraints are paramount, especially when the alternative consists of verbal report measures, whose reliability has been shown to be generally poor. In addition, because within-session analysis of several data points takes little time, its use as a standard adjunct to the brief functional analysis also seems justified if, as was the case in the present study, it enhances the interpretive power of the assessment (albeit marginally).

REFERENCES

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18,* 111–126.

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.

Derby, K. M., Wacker, D. P., Berg, W., DeRaad, A., Ulrich, S., Asmus, J., Harding, J., Prouty, A., Laffey, P., & Stoner, E. A. (1997). The long-term effects of functional communication training in home settings. *Journal of Applied Behavior Analysis, 30,* 507–531.

Derby, K. M., Wacker, D. P., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis, 25,* 713–721.

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.

Harding, J., Wacker, D. P., Cooper, L. J., Millard, T., & Jensen-Kovalan, P. (1994). Brief hierarchical assessment of potential treatment components with children in an outpatient clinic. *Journal of Applied Behavior Analysis, 27,* 291–300.

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., 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.

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., 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.

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.

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.

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.

STUDY QUESTIONS

1. What, according to the authors, are some factors that may hinder response differentiation during brief functional analyses?

2. What were the apparent establishing operations and reinforcement contingencies for SIB in each of the functional analysis conditions?

3. How were data for the brief and within-session analyses generated and interpreted?

4. Summarize the results obtained with respect to overall correspondence between outcomes of the brief and within-session analyses when compared to those of the full functional analyses.

5. What results were obtained from the analysis of "hit" and "miss" rates, and what are the clinical implications of these data?

6. According to the authors, what characteristics of the brief and within-session analyses probably accounted for observed differences between their respective hit and miss rates?

7. Under what set of circumstances is the within-session analysis most valuable, and to what extent was this observed in the present study?

8. How might the results have been affected by methodological limitations or procedural differences between this and previous studies?

Questions prepared by Jana Lindberg and April Worsdell, The University of Florida