

*FALSE-POSITIVE TANGIBLE OUTCOMES OF FUNCTIONAL ANALYSES*GRIFFIN W. ROOKER, BRIAN A. IWATA, JILL M. HARPER,  
TARA A. FAHMIE, AND ERIN M. CAMP

UNIVERSITY OF FLORIDA

Functional analysis (FA) methodology is the most precise method for identifying variables that maintain problem behavior. Occasionally, however, results of an FA may be influenced by idiosyncratic sensitivity to aspects of the assessment conditions. For example, data from several studies suggest that inclusion of a tangible condition during an FA may be prone to a false-positive outcome, although the extent to which tangible reinforcement routinely produces such outcomes is unknown. We examined susceptibility to tangible reinforcement by determining whether a new response was acquired more readily when exposed to a tangible contingency relative to others commonly used in an FA (Study 1), and whether problem behavior known *not* to have a social function nevertheless emerged when exposed to tangible reinforcement (Study 2). Results indicated that inclusion of items in the tangible condition should be done with care and that selection should be based on those items typically found in the individual's environment.

*Key words:* functional analysis, false-positive assessment, tangible reinforcement, motivating operations

Functional analysis (FA) methodology has become a standard approach for the assessment of severe problem behavior (Hanley, Iwata, & McCord, 2003). In a typical FA (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994), an individual is exposed to a series of conditions in which antecedent and consequent events are systematically varied while their effects on behavior are observed. In spite of its high degree of precision in identifying sources of reinforcement for problem behavior, results of an FA may be influenced occasionally by idiosyncratic sensitivity to various aspects of assessment conditions (therapists, objects, or tasks). One potential source of influence is the presence of tangible reinforcers. McCord and Neef (2005) discussed potential disadvantages of including tangible items as *antecedent events* in the attention condition of an FA, which may mask the reinforcing effects of

attention and produce a false-negative outcome (missing an attention function). The present study focuses on the extent to which delivery of tangible items as consequent events for problem behavior may produce a false-positive outcome (incorrectly identifying a tangible function).

Several sources of data suggest a high degree of sensitivity to tangible reinforcement during the course of an FA. First, although most studies in which an FA was conducted did not routinely contain a tangible condition, those studies that did also reported that a large proportion of problem behavior was maintained by tangible reinforcement (Hagopian, Fisher, Sullivan, Acquisto, & LeBlanc, 1998; Marcus, Vollmer, Swanson, Roane, & Ringdahl, 2001; Vollmer, Marcus, Ringdahl, & Roane, 1995). It is possible that the subjects in these studies were drawn from unusual samples, but another possibility is that the problem behavior of at least some subjects acquired a tangible function during assessment.

Another finding suggestive of influence by tangible reinforcement comes from FA studies in which problem behavior was found to be maintained by multiple control, that is, by more than one source of reinforcement. Beavers and Iwata (2011) examined FA studies published in

---

This research was supported in part by a grant from the Florida Agency for Persons with Disabilities.

Address correspondence to Griffin W. Rooker, who is now at the Department of Behavioral Psychology, Kennedy Krieger Institute, 707 North Broadway, Baltimore, Maryland 21205 (e-mail: Rooker@kennedykrieger.org) or Brian A. Iwata, Psychology Department, University of Florida, Gainesville, Florida 32611 (e-mail: iwata@ufl.edu).

doi: 10.1901/jaba.2011.44-737

the *Journal of Applied Behavior Analysis* and found 88 data sets indicative of multiple control. Of these, tangible reinforcement was reported as one source of maintenance in 41 cases, although a tangible condition was included in only 49 of the FAs. This high percentage may reflect the fact that a tangible condition typically was included based on preexperimental information suggestive of a tangible function; however, exposure to highly preferred stimuli in the tangible condition may have produced a false-positive outcome in at least some cases.

More direct evidence for false-positive tangible outcomes from an FA was reported in two studies that used FA and descriptive analysis (DA) methodologies in a complementary fashion. Shirley, Iwata, and Kahng (1999) found that hand mouthing exhibited by one individual occurred during all conditions of an FA (alone, attention, play, demand, and tangible) but that much higher levels of mouthing occurred during the tangible condition when any of three consequences (rings, food, Connect-4 game) was delivered. They then conducted a DA in the subject's home over the course of 10 15-min sessions to identify consequences that typically followed hand mouthing. The only tangible consequence that was observed was periodic delivery of a towel for drying the hands and mouth. The authors subsequently repeated a portion of the FA with two test conditions, the alone condition and a tangible condition that consisted of delivery of the towel, and observed nearly identical levels of mouthing in both conditions. Thus, tangible items selected from a preference assessment produced high rates of hand mouthing, whereas the tangible consequence identified through naturalistic observation had no effect. In a related study, Galiatsatos and Graff (2003) conducted an FA of one boy's screaming behavior and observed that screaming occurred most often in two tangible conditions: access to a toy and access to an edible item. They then conducted a DA in the student's classroom but never observed the delivery of any tangible

item as a consequence for screaming during 13 10-min observations.

Because tangible items, especially edible items, are highly effective reinforcers that typically are selected based on the results of preference assessments to enhance their value, they may be highly likely to produce new behavior as well as to increase the occurrence of behavior already in one's repertoire. The present studies explored both possibilities. We first (Study 1) determined whether a new response would be acquired more readily when exposed to tangible reinforcement than when exposed to other contingencies commonly used in an FA. If so, results would suggest that behavior in general may be more susceptible to tangible reinforcement. We subsequently (Study 2) determined whether problem behavior that did not have a social function emerged quickly when highly preferred tangible items were delivered as consequences, thereby showing a false-positive outcome.

## STUDY 1: ACQUISITION OF BEHAVIOR UNDER TANGIBLE REINFORCEMENT

### METHOD

#### *Subjects and Setting*

Six individuals who attended either a special education school or an adult day-treatment program for persons with developmental disabilities participated. Subjects were selected based on the criterion that they were neither reported nor observed to engage in problem behavior. Bobby was a 5-year-old boy who had been diagnosed with autism, Justin was a 49-year-old man who had been diagnosed with cerebellar agenesis, Rusty was a 24-year-old man who had been diagnosed with intellectual disability, Evan was a 5-year-old boy who had been diagnosed with speech and language delays, Jane was a 5-year-old girl who was typically developing, and Milton was a 41-year-old man who had been diagnosed with intellectual disability. Bobby, Justin, Rusty, Evan, and Jane could converse in simple sentences, whereas Milton communicated via gestures. None of the subjects had a history of training with the target responses, and at the time of the study none

participated in any program involving the use of food reinforcers. We did not limit access to any particular stimuli outside the sessions. However, the edible and leisure items used in this study typically were unavailable at the school or day program and never were observed to be delivered to subjects outside the sessions.

Sessions were conducted in an empty room at the school or day program that contained a table, chairs, and materials needed to conduct sessions (see below). All sessions were 10 min in duration, and five sessions were conducted daily, usually 4 to 5 days per week.

#### *Preference Assessments*

Two preference assessments were conducted to identify reinforcers to be used during subsequent conditions. A paired-stimulus assessment (Fisher et al., 1992) was conducted with nine edible items. On each trial, two items were presented to a subject, who was allowed to select and consume one of the items. Trials continued until all possible pairs had been presented, and the item selected most often (and on at least 80% of trials) was used as the consequence for responding during the tangible condition of the FA. A multiple-stimulus without replacement assessment (DeLeon & Iwata, 1996) was conducted with eight leisure items. All items were simultaneously presented in an array to a subject, who was allowed to select one item and interact with it for 30 s. On subsequent trials, previously selected items were removed from the array, and this continued until all items had been selected or until the subject made no further selections. The item selected first (highly preferred) was used as the leisure item available during the play condition of the FA; the item selected fourth (moderately preferred) was used as the leisure item available during the attention condition of the FA.

Trained observers used paper and pencil to collect data on subjects' selections, and an independent observer collected data on at least 30% of all trials. Reliability was calculated by comparing observers' records for each trial, which yielded 100% agreement for all subjects.

#### *Experimental Task*

An arbitrary task (switch pressing or card touching) was selected as the dependent variable for each subject. Switch pressing (Evan, Justin, and Rusty) was defined as pressing a button mounted on a plastic panel with sufficient force to produce an audible click. Card touching (Bobby, Jane, and Milton) was defined as placing a finger on a laminated piece of construction paper. These responses were selected because they were simple, discrete responses that subjects had not performed previously in any training situation.

Trained observers used handheld computers to collect data on the frequency of responses, and an independent observer collected data during a mean of 48% of sessions across subjects (range, 32% to 75%). Reliability was calculated by dividing session time into continuous 10-s intervals and comparing observer's records on an interval-by-interval basis. The smaller number of responses in each interval was divided by the larger; these fractions were then summed across the session and multiplied by 100%. Mean agreement across subjects was 98% (range, 97% to 99%).

#### *Functional Analysis*

Subjects were exposed to FA conditions similar to those described by Iwata et al. (1982/1994) with the addition of a tangible condition, and contingencies were placed on each subject's target behavior (button pressing or card touching). Immediately prior to the first session but not during any subsequent sessions, each subject was verbally prompted to engage in the target response once (no consequences were delivered). Alone, attention, demand, play, and tangible conditions were alternated in a multielement design. In the alone condition, the subject was alone in the therapy room with no materials present other than those required to perform the target response. In the attention condition, the therapist and subject were present in the room with a moderately preferred leisure item. If the subject engaged in the target response, the therapist delivered a brief statement of encouragement and physical attention (for approximately 5 s). In the demand

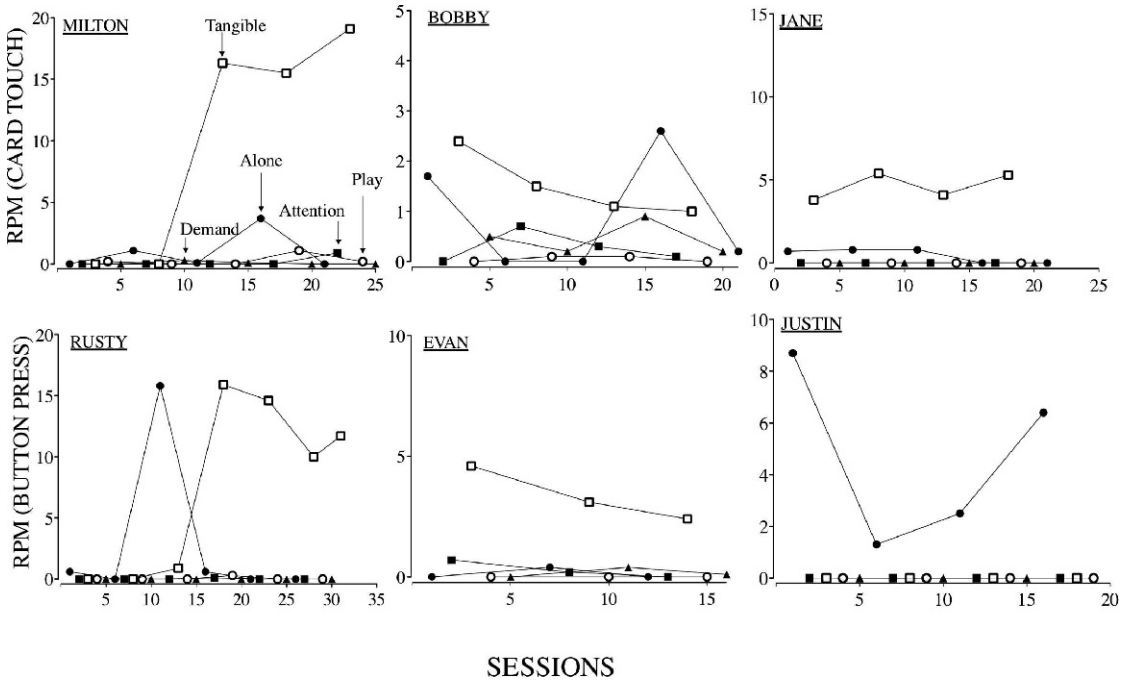


Figure 1. Responses per minute (RPM) of card touches or button presses during functional analysis conditions in Study 1.

condition, the therapist and subject were present in the room along with various tasks. The therapist presented instructional trials continuously throughout the session. If, however, the subject engaged in the target response, instruction was terminated, and tasks were removed for 30 s. In the play condition, the therapist and subject were present in the room with a highly preferred leisure item. The therapist delivered brief attention (as in the attention condition) every 30 s or any time the subject initiated interaction. If the subject engaged in the target response, no consequences were delivered. In the tangible condition, the therapist and subject were present in the room with a highly preferred edible item. If the subject engaged in the target response, the therapist delivered a bite-size portion of the edible item. In each condition, the switch or card was located near the source of reinforcement (e.g., next to the therapist in the attention condition, next to the task materials in the demand condition, next to the edible item in the tangible condition).

## RESULTS

Figure 1 shows results for all subjects. Milton's and Rusty's responding emerged over several exposures to the tangible condition, whereas Bobby, Jane, and Evan immediately began responding in the tangible condition. All five subjects exhibited very little responding in any other condition. By contrast, Justin engaged in high rates of responding only in the alone condition. Because he never exhibited responses in any other condition, he never contacted any of the programmed contingencies of the FA.

Data for five of the six subjects indicate that, when a response having no particular history of reinforcement is exposed to common contingencies in an FA, including those of the tangible condition, responding is more susceptible to tangible reinforcement. When acquisition of a tangible function occurred during the course of an FA, two patterns of responding were observed. In the first pattern (Milton's and Rusty's data), no responding occurred in the initial series of

sessions; thereafter, responding increased dramatically in the tangible condition. In the second pattern (Bobby's, Jane's, and Evan's data), responding occurred at a high rate during the first exposure to the tangible condition and was maintained almost exclusively in that condition. Justin was the only subject who did not acquire the target response in the tangible condition. His button pressing emerged and was maintained in the alone condition. Thus, some aspect of switch pressing per se (e.g., audible noise or tactile stimulation) was sufficient to maintain responding, whereas responses available in the other conditions effectively competed with switch pressing. Although it is possible that additional sessions with Justin would have shown acquisition in other test conditions, the pattern of responding evident in his data was sufficient to conclude that his behavior was maintained by automatic reinforcement and was relatively insensitive to other contingencies.

## STUDY 2: FALSE-POSITIVE TANGIBLE OUTCOMES DURING FUNCTIONAL ANALYSES OF PROBLEM BEHAVIOR

Results of Study 1 showed that a novel response was highly susceptible to a tangible reinforcement contingency relative to other contingencies typically included in an FA for five of the six subjects. However, it is unclear whether responses already in one's repertoire and maintained by contingencies other than tangible reinforcement would be similarly affected. The purpose of Study 2 was to explore this possibility by exposing problem behavior, which already had a history of maintenance by automatic reinforcement, to tangible contingencies.

### METHOD

#### *Subjects and Setting*

Three individuals who attended either a special education school or an adult day-treatment program for persons with developmental disabilities participated. All individuals engaged in noninjurious stereotypy. Amy was a 24-year-old woman

who had been diagnosed with autism and who engaged in body rocking (repetitive, back-and-forth movements from the waist up). Her primary mode of communication was one-word utterances. Oliver was a 42-year-old man who had been diagnosed with intellectual disability and who repeatedly tapped his finger on his body. His primary mode of communication was gesturing. Jimmy was an 8-year-old boy who had been diagnosed with autism and who engaged in finger-play (making a distinctive motion that involved his thumb, index, and middle fingers). His primary mode of communication was sign language (as a note, the distinctive finger motion was not similar to any of Jimmy's signs). None of the subjects were reported to engage in the target behavior to gain access to tangible reinforcement, and at the time of the study none participated in any training programs that involved the use of food or leisure-activity reinforcers. As in Study 1, although we did not limit access to any particular stimuli outside the sessions, the edible and leisure items used in this study typically were unavailable and never were observed to be delivered to subjects outside the sessions.

Sessions were conducted in an empty therapy room that contained a table, chairs, and the materials necessary to conduct the FA conditions (see below). All sessions were 10 min in duration, and four sessions were conducted daily, usually 4 to 5 days per week.

#### *Response Measurement and Reliability*

The dependent variable in all phases of this study was the subject's stereotypy (see previous definitions). Trained observers used handheld computers to record stereotypy on a 10-s partial-interval basis. Data were collected by an independent observer on an average of 54% of sessions, and reliability was calculated as described in the previous study. The mean agreement percentage across subjects was 95% (range, 92% to 98%).

#### *Preference Assessments*

Prior to the FA, stimuli to be used in the play and attention conditions were identified in the



same manner as the previous study (DeLeon & Iwata, 1996). A paired-stimulus assessment (Fisher *et al.*, 1992) was used to identify a highly preferred edible item. The item selected most often (and on at least 80% of trials) was selected for inclusion in the tangible preference assessment condition in the comparison of tangible conditions. For both preference assessments, trained observers used paper and pencil to collect data on selection responses. An independent observer collected data on at least 30% of all trials. Reliability was calculated by comparing observers' records for each trial, which yielded 100% agreement for all subjects.

#### *Descriptive Assessment*

A DA was conducted to identify consequences frequently delivered following occurrences of stereotypy in the natural environment. Subjects were observed during a series of 10-min sessions in their classroom or day program. An observer recorded the delivery of any stimulus within 30 s after every instance of stereotypy. Although the 30-s window for scoring consequent events was larger than that use in typical DAs, it increased the likelihood that relevant stimuli would be identified rather than missed. Sessions continued until at least 12 10-min sessions had been conducted in which stereotypy was observed. The tangible item (or items) observed most frequently was selected as the tangible DA item and was used in the comparison of tangible conditions. If no tangible stimulus was ever delivered, a paired-stimulus preference assessment was conducted using items commonly found in the classroom or day program, and the most preferred item was used in the comparison of tangible conditions.

#### *Functional Analysis*

*Initial FA.* A multielement FA (Iwata *et al.*, 1982/1994) including alone, attention, demand, and play conditions, was conducted for each subject's stereotypy. All sessions were 10 min in duration.

*Comparison of tangible conditions.* Three conditions were conducted in a multielement design.

Table 1  
Results of Preference Assessment (PA) and Descriptive Analysis (DA)

Subject	PA (% trials selected)	DA (frequency)
Amy	chocolate (100%)	juice (1) puzzle (1) video (4)
Oliver	peppermint-chocolate (87.5%)	magazine (87.5%) <sup>a</sup>
Jimmy	yogurt (87.5%)	marker (1)

<sup>a</sup> Identified via secondary preference assessment.

The first was identical to the alone condition of the FA. The second was the tangible preference assessment condition, in which the most highly preferred edible item from the paired-stimulus preference assessment was delivered following each instance of stereotypy. The third condition was the tangible DA condition, in which the stimulus most frequently delivered during the DA was delivered after each instance of stereotypy. All sessions were 10 min in duration.

## RESULTS

Table 1 shows results of the preference assessment and the DA for each subject. Highly preferred edible items were readily identified in the paired-stimulus assessment. The DA identified naturally occurring consequences for Amy's and Jimmy's stereotypy. Three items were identified in Amy's DA: a videotape (presented four times), juice (presented once), and a puzzle (presented once). Jimmy's DA yielded a single delivery of a tangible item (a marker) following stereotypy. No items were ever identified in Oliver's DA, despite the fact that 4 hr (24 10-min sessions) with stereotypy were observed. Because it seemed unlikely that DA would identify any stimulus that was highly correlated with the occurrence of problem behavior, a nine-item paired-stimulus preference assessment was conducted using items commonly found in Oliver's immediate environment. A magazine was determined to be highly preferred and was selected as the item to be included in the tangible DA condition.

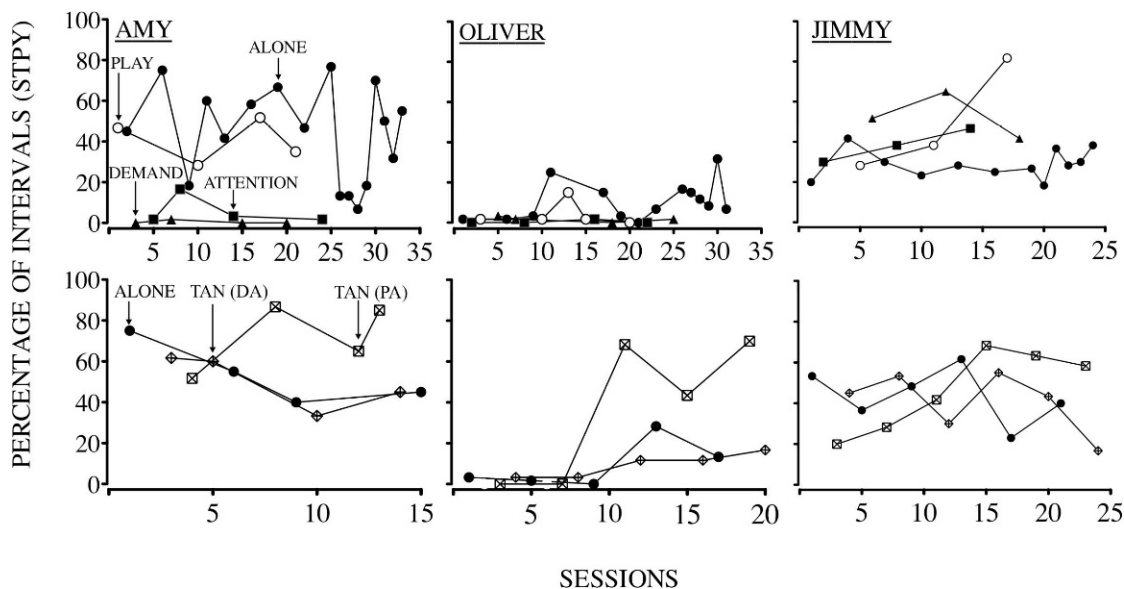


Figure 2. Percentage of intervals of stereotypy (STPY) across functional analysis conditions in Study 2.

Figure 2 (top) shows results of the FA for all subjects. High levels of body rocking occurred in two conditions of Amy's FA (alone and play) and persisted during a subsequent series of alone sessions. Low levels of tapping were observed in all of Oliver's FA conditions. However, responding occurred most consistently in the alone condition and continued to occur during subsequent alone sessions. Moderate to high levels of finger play were observed during Jimmy's FA conditions. Although higher levels of stereotypy occurred in conditions other than the alone condition, stereotypy was maintained during a series of alone sessions. Thus, all subjects' FA results indicated that stereotypy was maintained at least in part by automatic reinforcement.

Figure 2 (bottom) shows results of the comparison of tangible conditions. Amy's and Oliver's stereotypy occurred more often and more consistently in the tangible preference assessment condition. By contrast, stereotypy in their tangible DA condition was similar to that observed when no consequences whatsoever were delivered (alone condition). Jimmy initially engaged in somewhat less stereotypy in the tangible preference assessment

condition than in the alone and tangible DA conditions, but after several sessions he engaged in stereotypy most often in the tangible PA condition.

## DISCUSSION

Results from these studies indicate that a response with either no history of reinforcement (Study 1) or a different history of reinforcement (Study 2) is highly susceptible to the effects of tangible reinforcement. Results of Study 1 were not surprising given that edible items are used often as reinforcers in the acquisition of new responses and were selected based on outcomes of a preference assessment. As a result, the highly preferred items used in Study 1 were more effective in establishing the target response than were consequences available in the other conditions (attention or escape from task demands). Results of Study 2 showed that stereotypy maintained by automatic reinforcement was more likely to occur under a condition of tangible reinforcement (tangible preference assessment) than during either the alone condition or one in which stimuli selected from subjects' natural environments (tangible DA) were delivered as

consequences. These results demonstrated a false-positive FA outcome for tangible reinforcement, replicating the findings of Shirley *et al.* (1999).

Results from both studies suggest that tangible reinforcers, particularly in the form of edible items, are highly likely to produce a positive outcome when included as a condition in an FA. Thus, one recommendation is not to include a tangible FA condition unless there is a strong suspicion that problem behavior is, in fact, maintained by tangible reinforcement. Even then, it is possible that the use of highly preferred tangible stimuli might produce false-positive outcomes. Therefore, a related recommendation is that, when a tangible condition is included, it should reflect tangible items that actually are delivered as a consequence for problem behavior. If none are identified, as was the case for Oliver (Study 2), perhaps the tangible condition is unnecessary in spite of anecdotal reports to the contrary.

Although the present results were highly consistent, some qualifications are in order. First, it is possible that additional functions for the arbitrary response may have emerged in Study 1 had conditions been conducted for a longer period of time. Still, the fact that responding emerged more quickly under tangible reinforcement for five of six subjects illustrates the superior effects of tangible reinforcers relative to others. Second, although the tangible DA items delivered as consequences in Study 2 had no reinforcing effect on stereotypy, it is possible that other items encountered in subjects' daily environments may have maintained their stereotypy. However, the delivery of such items would have been so intermittent that it never occurred during repeated observations over an extended period of time. Third, the inclusion of edible items in both studies makes it unclear if the effects observed also would have been obtained with leisure items. Finally, data from Study 2 may have limited generality to problem behavior other than that maintained by automatic reinforcement. However, if behavior maintained by automatic reinforcement is highly sensitive to tangible reinforcers, one

would suspect that behavior already maintained by social reinforcement may be even more likely to come under the control of other (tangible) reinforcers.

Data from several studies have indicated that specific antecedent and consequent events may influence the occurrence of problem behavior during an FA, producing both false-positive (Galiatsatos & Graff, 2003; Shirley *et al.*, 1999) and false-negative (O'Reilly, Lancioni, King, Lally, & Dhomhnaill, 2000; Roscoe, Carreau, MacDonald, & Pence, 2008; Roscoe, Rooker, Pence, & Longworth, 2009) outcomes. In each case, clarification of results required further analysis. The fact that such findings and the subsequent need to make procedural adjustments have been reported rarely underscores the general robustness of the conditions contained in a typical FA. Nevertheless, these exceptions may be informative to researchers and clinicians who observe findings during treatment contrary to what was expected based on assessment results.

## REFERENCES

- Beavers, G. A., & Iwata, B. A. (2011). Prevalence of multiply controlled problem behavior. *Journal of Applied Behavior Analysis, 44*, 593–597.
- DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*, 519–533.
- Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*, 491–498.
- Galiatsatos, G. T., & Graff, R. B. (2003). Combining descriptive and functional analyses to assess and treat screaming. *Behavioral Interventions, 18*, 123–138.
- Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquistio, J., & LeBlanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis, 31*, 211–235.
- Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147–186.
- 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)
- Marcus, B. A., Vollmer, T. R., Swanson, V., Roane, H. R., & Ringdahl, J. E. (2001). An experimental analysis of aggression. *Behavior Modification*, 25, 189–213.
- McCord, B. E., & Neef, N. A. (2005). Leisure items as controls in the attention condition of functional analyses. *Journal of Applied Behavior Analysis*, 38, 417–426.
- O'Reilly, M. F., Lancioni, G. E., King, L., Lally, G., & Dhomhnaill, O. N. (2000). Using brief assessments to evaluate aberrant behavior maintained by attention. *Journal of Applied Behavior Analysis*, 33, 109–112.
- Roscoe, E. M., Carreau, A., MacDonald, J., & Pence, S. T. (2008). Further evaluation of leisure items in the attention condition of functional analyses. *Journal of Applied Behavior Analysis*, 41, 351–364.
- Roscoe, E. M., Rooker, G. W., Pence, S. T., & Longworth, L. J. (2009). Assessing the utility of a demand assessment for functional analysis. *Journal of Applied Behavior Analysis*, 42, 819–825.
- Shirley, M. J., Iwata, B. A., & Kahng, S. (1999). False-positive maintenance of self-injurious behavior by access to tangible reinforcers. *Journal of Applied Behavior Analysis*, 32, 201–204.
- 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.

Received October 1, 2010

Final acceptance May 24, 2011

Action Editor, Joel Ringdahl