

Do Functional Analyses Probe Learning Histories, Or Create Them? An Exploratory  
Investigation

Joseph M. Lambert & Nealetta J. Houchins-Juárez

Vanderbilt University

Author Note

Joseph M. Lambert, Department of Special Education, Vanderbilt University; Nealetta J. Houchins-Juárez, Department of Special Education, Vanderbilt University.

We thank Rachel Mottern, Andrea Perkins, Casey Chauvin, Gillian Cattey, Kate Chazin, Kristen Stankiewicz, Lilly Stiff, Sara Sheffler, Sarah Reynolds, Sarah Shaw, and Naomi Parikh for their assistance collecting and summarizing data for this study and for coordinating and delivering high-quality therapeutic services for participants following this study's conclusion. We also thank the Treatment and Research Institute for Autism Spectrum Disorders for providing clinic space and Jessica Torelli for her feedback on an earlier draft of this manuscript.

Correspondence concerning this article should be addressed to Joseph M. Lambert, Department of Special Education, Vanderbilt University, Nashville, TN 37203. Fax (615) 343-1570. Email: joseph.m.lambert@vanderbilt.edu

Do Functional Analyses Probe Learning Histories, Or Create Them? An Exploratory  
Investigation

Joseph M. Lambert & Nealetta J. Houchins-Juárez

Vanderbilt University

Author Note

Joseph M. Lambert, Department of Special Education, Vanderbilt University; Nealetta J. Houchins-Juárez, Department of Special Education, Vanderbilt University.

We thank Rachel Mottern, Andrea Perkins, Casey Chauvin, Gillian Cattey, Kate Chazin, Kristen Stankiewicz, Lilly Stiff, Sara Sheffler, Sarah Reynolds, Sarah Shaw, and Naomi Parikh for their assistance collecting and summarizing data for this study and for coordinating and delivering high-quality therapeutic services for participants following this study's conclusion. We also thank the Treatment and Research Institute for Autism Spectrum Disorders for providing clinic space and Jessica Torelli for her feedback on an earlier draft of this manuscript.

Correspondence concerning this article should be addressed to Joseph M. Lambert, Department of Special Education, Vanderbilt University, Nashville, TN 37203. Fax (615) 343-1570. Email: joseph.m.lambert@vanderbilt.edu

### Abstract

During functional analysis (FA), therapists arrange contingencies between potential reinforcers and problem behavior. It is unclear whether this fact, in and of itself, facilitates problem behavior's acquisition of new (false-positive) functions. If problem behavior can come under the control of contingencies contrived between it and known reinforcers for which there is no direct history, then outcomes of reinforcer analysis (RA) should perfectly predict FA outcomes. This study evaluated the degree to which RA outcomes corresponded with FA outcomes for eight participants referred to a university-based outpatient clinic for problem behavior. For 75% (6 of 8) of participants, correspondence was imperfect. These findings appear to support the construct validity of contemporary interpretations of FA data.

*Keywords:* divergent construct validity, false positive, functional analysis, reinforcer analysis, problem behavior

## Do Functional Analyses Probe Learning Histories, Or Create Them? An Exploratory Investigation

Current publication practices (Hanley, Iwata, & McCord, 2003; Beavers, Iwata, & Lerman, 2013), conventional wisdom (Iwata & Dozier, 2008; Hanley, 2012; Mace, 1994), and a growing research base (e.g., Hurl, Wightman, Haynes, & Virues-Ortega, 2016) suggest optimal intervention outcomes for problem behavior are most likely when pre-intervention assessments include functional analysis (FA; Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994). During FAs, experimenters systematically establish and abolish the reinforcing value of programmed consequences (often isolated by stimulus class [e.g., attention, tangibles] or process [e.g., escape]) to probe the evocative and reinforcing effects these manipulations have on problem behavior. Conditions that consistently produce elevated rates of responding, relative to no-consequence control conditions, are purported to shed light on problem behavior's reinforcement history and are used to implicate its function(s).

Despite majority consensus about the superiority of FA to alternative approaches to function identification, few practitioners actually employ this technology in their day-to-day practice (Oliver, Pratt, & Normand, 2015; Roscoe, Phillips, Kelly, Farber, & Dube, 2015). One reason for this may be that some stakeholders raise concerns about various aspects of FA methodology (e.g., Hastings & Noone, 2005; Martin, Gaffan, & Williams, 1999; Solnick & Ardoin, 2010; Sturmey, 1995).

### **Potential Concerns with FA Methodology and Interpretation**

During FAs, practitioners deliberately evoke and reinforce problem behavior. Because reinforcement, by definition, increases the future probability of behavior (Catania, 2013, Cooper, Heron, & Heward, 2007), a valid concern may be that FAs are counter-therapeutic and might

intensify the nature of existing problem behavior in naturalistic settings (by increasing its rate or magnitude), or might shape or establish new functions for problem behavior (by delivering reinforcers not commonly provided in naturalistic settings). To evaluate the former, Shabani et al. (2013) collected data on problem behavior in each of four participants' classrooms during the 10 min immediately prior to, and following, FA sessions and compared pre/post rates. Although results were inconsistent across participants (with evidence suggestive of generalization in some cases and contrast in others), differences were generally negligible. Call, Findley, and Reavis (2012) and Call et al. (2017) conducted similar experiments and produced similar results.

These studies shed light on the ways FAs may, or may not, influence baseline rates of problem behavior in naturalistic settings. However, they do not evaluate whether FAs are instructional by nature. That is, results from studies like Call et al. (2012), Call et al. (2017), and Shabani et al. (2013) cannot prove or disprove functions identified through FAs are the product of artificially contrived contingencies arranged during the assessment process. In fact, the results of other research (e.g., Galiatsatos & Graff, 2003; Rooker, Iwata, Harper, Fahmie, & Camp, 2011; Shirley, Iwata, & Kahng, 1999) lead us to question the validity of functions identified in certain conditions of the analysis.

For example, Galiatsatos and Graff (2003) and Shirley et al. (1999) found tangibles could maintain problem behavior during FAs, even though tangibles were never delivered for problem behavior during an extended series of descriptive assessments conducted in participants' typical settings. Similarly, Shirley et al. and Rooker et al. (2011) both found highly preferred tangible deliveries not observed in typical settings could increase rates of automatically-maintained problem behavior, relative to conditions in which no tangibles (or tangibles that *did* follow problem behavior in typical settings) were delivered. Finally, Rooker et al. demonstrated

contingent tangibles were more effective at establishing and maintaining high rates of an arbitrary task than either contingent attention or contingent escape, suggesting tangibles were more likely to establish new functions for a target response than other reinforcers commonly programmed into FA conditions. The conglomerate of findings across these studies led Rooker et al. to conclude that tangible conditions should be omitted from most FAs of problem behavior and, when included, results from this condition should be interpreted with caution.

Importantly, across these studies, highly preferred tangibles were identified via preference assessments. By contrast, the specific qualities of attention and escape were never systematically evaluated. Although this aligns with how consequences are often selected for FAs (cf., Slocum, Grauerholz-Fisher, Peters, & Vollmer, 2018), it represents a procedural inconsistency that allows for alternative interpretations of data published on this topic. For example, it may be that the contingent delivery of any highly preferred consequence, irrespective of stimulus class (e.g., tangible or attention) or process (e.g., positive or negative) would have been equally likely to establish functional relations that did not exist prior to the analysis. If functions identified via FAs are artifacts of learning that has occurred during the analysis (i.e., if FAs generate “false positives”, in the sense that functional relations did not exist prior to the analysis), it would be hard to refute claims that FA procedures are flawed and outcomes invalid (Martin et al., 1999). By extension, similar criticisms could be levied against the validity of intervention outcomes in studies that used FA conditions as baseline for treatment evaluations (e.g., Sturmey, 1995).

The prospect of FA creating the histories it is merely meant to detect has far-reaching implications. For example, if FAs do not reliably examine pre-assessment histories, then practitioners should be cautious about attributing historical significance to FA outcomes.

Related, if FA methodology only evaluates the reinforcing effect of programmed consequences on target responses, then FAs are no different than reinforcer analyses (RA) and should not carry a special distinction. By extension, the practice of targeting and reinforcing problem behavior during FAs would be difficult to justify because the strategy would not add value to assessment outcomes and would be unnecessarily taxing on the social validity of behavior-change technology.

Similar to FAs, reinforcer analyses (RA) are experiments designed to evaluate whether programmed contingencies can maintain elevated rates, magnitudes, or allocations of responding across time (e.g., Buchmeier, Baker, Reuter-Yuill, & MacNeill, 2018; Call, Trosclair-Lasserre, Findley, Reavis, & Shillingsburg, 2012; Cote, Thompson, Hanley, & McKerchar, 2007; DeLeon & Iwata, 1996; Hagopian, LeBlanc, & Maglieri, 2000; Hodos, 1961; Quick, Baker, & Ringdahl, 2018; Tarbox, Tarbox, Ghezzi, Wallace, & Yoo, 2007). Unlike FAs, RAs do not purport to probe pre-experimental learning histories. In fact, experimenters will often provide instructions, forced-choice exposures, or both before initiating the analysis to ensure that negative outcomes implicate ineffective contingencies rather than insufficient experience (cf. Buchmeier et al., 2018; Call et al., 2012; Cote et al., 2007; DeLeon & Iwata, 1996; Hagopian et al., 2000; Hodos, 1961; Quick et al., 2018; Tarbox et al., 2007).

### **Pre-Experimental History with Reinforcers**

Certainly, frequent contact with powerful reinforcers can facilitate a response's acquisition of new functions (cf., Rooker et al., 2011). However, during FAs, participants are not instructed to engage in problem behavior and forced-choice exposures to programmed contingencies do not occur. Rather, therapists arrange contingencies between potentially reinforcing consequences and problem behavior and probe to see whether manipulated

discriminative stimuli ( $S^D$ ) and establishing operations (EO) evoke problem behavior. If a participant is not already inclined to engage in problem behavior following EO onset, then problem behavior should never contact programmed consequences and new learning should not occur. As the goal of FA is not to teach, but rather to examine what participants have already learned (by arranging the conditions most likely to evoke demonstrations of that learning), FAs should only produce functional relations when those relations existed prior to analysis.

Rooker et al. (2011) and Shirley et al. (1999) both identified circumstances under which FAs don't live up to this ideal by demonstrating it is possible to teach new relations *during* the analysis. However, Rooker et al. and Shirley et al. targeted automatically maintained problem behavior. That is, they evaluated the effect superimposed contingencies of arbitrary reinforcement would have on the rates of problem behavior demonstrated to occur frequently in the absence of social consequences; thus guaranteeing frequent contact between it and superimposed consequences and therefore creating the very history the analysis was only meant to detect.

It is unclear whether arbitrary contingencies of reinforcement superimposed onto problem behavior with socially mediated functions would be equally as likely to implicate false-positive functions during FAs. This is because experimenters can more precisely control socially-mediated problem behavior (cf. responding in control conditions in Jessel, Hanley, & Ghaemmaghami, 2016; Jessel, Ingvarsson, Metras, Kirk, & Whipple, 2018). Increased control likely decreases the potential for within-analysis experiences with historically irrelevant contingencies and may decrease the probability of producing false-positive outcomes during FAs of socially-mediated problem behavior (although see Jessel, Hausman, Schmidt, Darnell, & Kahng, 2014). Thus, it remains unclear whether FAs of socially-mediated problem behavior can

or do implicate false-positive functions.

To date, little empirical work has been done on this topic. Perhaps one reason for this stems from the difficulties associated with confirming or negating extra-experimental contact with programmed consequences. Although Rooker et al. (2011) and Shirley et al. (1999) used descriptive assessment (DA) results as comparisons against which to evaluate the ecological validity of tangible functions identified during FAs, tangible delivery is a fairly discriminable event and is relatively easy to quantify. By contrast, various qualities and magnitudes of attention and escape can be sufficient to maintain behavior (cf. Kelly, Roscoe, Hanley, & Schlichenmeyer, 2014; Lerman, Kelley, Vorndran, Kuhn, & LaRue, 2002) and may be harder to detect during uncontrolled observations. Furthermore, individuals' histories are often complex (Beavers & Iwata, 2011) and schedules of reinforcement in the natural environment can be lean and unpredictable (Call et al., 2017). When behavior occurs with any degree of variability, even DAs that include extended-duration observations (e.g., 60 min) are unlikely to generate representative samples of real-world circumstances (Tiger et al., 2013) or reliably isolate the antecedent and consequent events responsible for maintaining problem behavior (Pence, Roscoe, Bourret, & Ahearn, 2009).

Thus, we concluded it is problematic to suggest functions implicated during analogue FAs are "false" because target consequences are not observed (or correctly quantified) during formal DAs. This conclusion presents a problem when evaluating the validity of contemporary interpretations of FA data because there does not appear to be an acceptable standard for confirming or negating pre-experimental reinforcement histories, outside of the analysis itself.

### **Construct Validation**

Construct validation is a term which describes the process by which investigators defend

inferences about data when attributed meaning extends beyond obtained results (Cronbach & Meehl, 1955). To engage in this process, investigators propose nomologicals that theoretically connect constructs to observed outcomes and inform falsifiable hypotheses that can later be tested. When correspondence between predicted and obtained outcomes is high, there is evidence that interpretations of data are correct and that assessments measure what they are intended to measure (Yoder, Lloyd, & Symons, 2018). Although construct validation is typically done through statistical analysis of data obtained through group design (cf. Yoder et al.), the concept of divergent construct validity may hold relevance to the current issue. *Divergent validity* is demonstrated when a relevant variable is not associated with outcomes with which it should not be associated.

For example, we can state that persistent patterns of behavior are a product of experience with reinforcement. From this, we can hypothesize that contingencies of reinforcement that have been learned will control behavior whereas contingencies of reinforcement that have not been learned will not (even though they could, given sufficient experience). We might also say that both FAs and RAs are designed to empirically demonstrate the control that reinforcement can have over behavior but only FAs are designed to detect pre-experimental experiences with reinforcement. Said differently, in order to establish a functional relation during an RA, the programmed contingency need only be reinforcing. However, to establish a functional relation during FA, the programmed contingency needs to be reinforcing *and* the individual would theoretically need to have experienced the contingency prior to the FA.

As problem behavior is unlikely to be maintained by all known reinforcers (for indirect evidence of this, see discussion of the prevalence of multiply controlled problem behavior in Beavers & Iwata, 2011), every functional relation established during RAs should not be

replicated during FAs. Therefore, if we conduct both analyses and compare outcomes across a number of participants, we can evaluate (indirectly) the validity of contemporary interpretations of FA data in ways that are analogous to how others have assessed divergent construct validity.

Specifically, if RAs and FAs are functionally equivalent, then RA results should consistently and perfectly predict FA results across participants. By contrast, if RAs and FAs each identify what they purport to identify (i.e., reinforcing contingencies in the case of the former and experience with said contingencies in the case of the latter), then RA results should serve as unreliable predictors of FA results. Each time the outcomes of RAs and FAs do not perfectly correspond, we accumulate evidence that they measure different things and demonstrate that something more than a simple contingency of reinforcement is needed to establish functional relations between reinforcers and target responses during FAs (presumably pre-experimental experience).

Thus, the purpose of this study was to compare results of RAs with results of FAs when implementers and systematically identified consequences were held constant across assessments. We propose that imperfect correspondence between assessments across participants would contribute divergent construct validity to contemporary interpretations of FA outcomes.

## **Method**

### **Participants and Setting**

We included the first eight individuals referred for services at a university-based outpatient clinic who (a) engaged in problem behavior hypothesized to be socially mediated during pre-assessment interviews, (b) for whom informed consent could be obtained, (c) who completed RAs and FAs with matched tangible, attention, and escape conditions, and (d) for whom acceptable interobserver agreement (IOA) and procedural fidelity scores could be

obtained. Participant ages, parent-reported diagnoses, and researcher-generated ratings of communication skills are displayed in Table 1. All preference assessments, RAs, and FAs were conducted in a clinical room equipped with table, chairs, condition-specific materials, and a one-way mirror.

### **Response measurement**

The primary dependent variable during RAs was the rate of occurrence of an arbitrary behavior. The primary dependent variable during FAs was the rate of occurrence of problem behavior. Arbitrary behaviors targeted during RAs were individualized for each participant and ranged from wiping tables to clapping hands. To be targeted, arbitrary responses needed to be: (a) free operant, (b) easily performed by the participant, and (c) not automatically maintained. Similarly, problem behavior was individualized for each participant, based on caregiver report. Arbitrary and problem behaviors selected for each participant are listed in Table 1. Definitions are available in this study's *supplemental materials* section.

Using handheld computers with data-collection software, observers trained with the video-based protocol outlined by Dempsey, Iwata, Fritz, & Rolider (2012) collected frequency data on arbitrary responses during RAs and frequency data on problem behavior during FAs.

### **Interobserver Agreement**

A second observer simultaneously but independently collected data on all dependent variables for no less than 25% of total sessions conducted during each analysis (i.e., RA and FA) for each participant. We compared primary and reliability data and calculated interobserver agreement (IOA) by dividing RA and FA sessions into 10-s bins, dividing agreements about the occurrence of each dependent variable within each bin by the sum of agreements and

disagreements, and multiplying by 100. Bin scores were then averaged to generate session IOA scores. Average IOA for each participant is reported in Table 2.

### **Procedural Fidelity**

Observers used yes/no checklists to evaluate therapist fidelity to programmed procedures during both RAs and FAs. We calculated fidelity scores by dividing the number of “yes” (marked when a procedure was implemented as programmed) by the sum of “yes” and “no” and multiplying by 100. Average fidelity scores for each participant are shown in Table 2.

### **General Procedures**

All data for this study were obtained during the initial weeks of service provision for each participant, prior to the onset of intervention. All participants attended two appointments per week and each appointment lasted two hours. All participants progressed through the same general sequence of events. First, we conducted open-ended functional-assessment interviews (Iwata, DeLeon, & Roscoe, 2013; O’Neill et al., 1997) with primary caregivers to operationalize problem behavior and identify an array of consequences that might plausibly follow problem behavior in naturalistic settings. We then organized these consequences into the classes of social interaction commonly arranged during FAs (i.e., attention, tangible, & escape).

Next, we incorporated variables from the above-mentioned categories into a series of preference and demand assessments to identify the consequence most likely to function as reinforcement for each class of social interaction. Then, we conducted RAs to test whether contingent access to, or escape from, identified events would function as reinforcement. Finally, we conducted FAs of problem behavior. With the exclusion of programmed breaks, the entirety of each appointment with each participant was dedicated to obtaining data from these assessments, which were conducted in the order and fashion described above and below.

To facilitate discrimination between conditions during RAs and FAs (Conners et al., 2000), we assigned a specific therapist (i.e., a graduate student in Special Education and Applied Behavior Analysis), who wore a specific colored shirt (e.g., red, blue, white), to each condition tested. We always conducted RAs prior to FAs, manipulated identical qualities and durations of antecedent and consequent events within and across the conditions of both analyses, and used the same schedule-correlated stimuli (e.g., therapists and colors) to signal the same active contingencies during both analyses. We did this to ensure that negative functions implicated during FAs weren't artifacts of insufficient exposure to  $S^D$ s, EOs, or experience with consequences. That is, we arranged our experiment in this way to capitalize on learning that occurred during RAs in a way that would increase the probability of achieving perfect correspondence across analyses. Notwithstanding, we used different control techniques (i.e., reversal [RA] and multielement [FA]) to minimize the possibility of response induction attributable to similarities in experimental design.

Because we presumed that any response would quickly come under the control of contingent access to a reinforcer if explicit instructions were provided, or if experience with relevant contingencies was ensured, we provided both contingency reviews and forced-choice exposures during RAs of arbitrary responses but provided neither during FAs of problem behavior. Likewise, we conducted RAs using 2-min sessions and FAs using 5-min sessions. We did this to demonstrate it was possible to establish the reinforcing effect of programmed consequences in less time than what was allotted during FAs. Assuming non-correspondences between analyses would be due to positive functional relations established during RAs but not FAs, we made these modifications to address potential concerns about whether better correspondences would have occurred if FAs had been carried out for longer periods of time.

**Preference assessments.** To identify each participant's highest preferred tangible item from arrays identified during interviews, therapists either conducted a multiple stimulus without replacement (MSWO; DeLeon & Iwata, 1996) or a paired stimulus (Fisher et al., 1992) preference assessment. To identify participants' highest preferred form of attention, therapists identified four potentially preferred social interactions using information from parent reports and informal observations and pitted them against each other in discrete-trial paired-stimulus social-interactions preference assessments adapted from procedures described by Clay, Samaha, Bloom, Bogoev, and Boyle (2013). Likewise, to identify a low-probability demand for each participant (from which escape might function as a reinforcer), we identified five to ten potential low-probability demands using information from parent reports and informal observations and conducted discrete-trial demand assessments adapted from procedures outlined by Roscoe, Rooker, Pence, & Longworth (2009). Highest-preferred tangibles, highest-preferred attention, and lowest-probability demands for each participant are listed in Table 1.

**Reinforcer analysis.** Procedures for this analysis were adapted from procedures described by DeLeon and Iwata (1996). We conducted tangible, attention, and escape RAs (in that order) with each participant. Session duration was 2 min. During sessions, clinical rooms contained participants and condition-specific therapists. Prior to initiating RAs, therapists selected a single arbitrary response to reinforce across all RAs for each participant. Arbitrary responses needed to fall within participants' current repertoires (e.g., elbow touch, cup flip) and could not be automatically maintained. If a target response persisted across sessions during the first baseline condition of the tangible RA (suggesting automatic control), data from those sessions were discarded, a new response was identified, and a new RA was initiated. Only

materials required to engage in target responses and those required to deliver programmed consequences were present during RA sessions.

Each RA entailed A (baseline) conditions and B (Fixed Ratio [FR] 1) conditions. During baseline conditions, therapists did not interact with participants nor did they react to targeted responses. Prior to the first session of each FR1 condition, participants were given 30 s access to programmed consequences (i.e., highly preferred tangibles, highly preferred attention, or no demands). At session onset, therapists denied participants access to highly preferred tangible items (tangible RA), removed their attention (attention RA) or began presenting low-probability demands using a two-step prompting hierarchy (vocal, manual guidance; 5-s inter-prompt interval) (escape RA). Following each target response, therapists produced 15 s access to programmed consequences (i.e., tangible, attention, or escape). During FR1 conditions of escape RAs, compliance produced brief praise and a new demand. If problem behavior occurred, it did not produce programmed consequences, nor did it delay reinforcement for target responses (anecdotally, problem behavior rarely emerged during RAs and did not occur contiguously with target responding).

Prior to the first session of every condition (baseline or FR1), therapists conducted contingency reviews concurrently with forced-choice exposures to programmed consequences (i.e., no consequence [baseline] or 15-s access to reinforcement [FR1]). Contingency reviews entailed verbal descriptions of programmed consequences for target responses (i.e., “when you do this, nothing happens” [baseline]. Or, “when you do this, this happens” [FR1]). Because therapists and caregivers both questioned whether a single contingency review would be sufficient to teach Becky and Diana about programmed consequences, therapists also provided

within-session vocal reminders of relevant contingencies (i.e., “remember, when you do x, you get y”) on a fixed-time 30-s schedule during both baseline and FR1 conditions.

Therapists conducted a minimum of three sessions of each condition and implemented an ABAB reversal design for any RA in which rates of target responding increased in FR1 conditions relative to rates in baseline conditions. Therapists conducted a minimum of five sessions in FR1 conditions before concluding that a demonstration of effect had not occurred and discontinuing an RA.

**Functional analysis.** Sessions were 5 min in duration. During FAs, every participant was exposed to traditional test (i.e., tangible, attention, escape) and control (play) conditions. A select subset of participants were also exposed to idiosyncratic test conditions (e.g., social avoidance, divided attention in a mixed-gender group context) implicated during interviews or informal interactions (Hanley, Jin, Vanselow, & Hanratty, 2014; Schlichenmeyer, Roscoe, Rooker, Wheeler, & Dube, 2013). During play conditions, no demands were presented, participants were given continuous access to highly preferred tangibles, and attention was delivered at least once every 30 s. Prior to every test session, participants were provided with 30-s access to putative reinforcers. At session onset, reinforcers were removed, relevant EOs presented, and problem behavior produced 15 s access to programmed consequences. If responses targeted during RAs occurred during FAs they did not produce programmed consequences, nor did they delay reinforcement for problem behavior (anecdotally, target responses rarely emerged during FAs and did not occur contiguously with target responding).

With the exception of responses targeted, session duration, and experimental design, tangible, attention, and escape FA sessions were identical to tangible, attention, and escape RA sessions. Likewise, ignore sessions were identical to baseline RA sessions. Although we did not

compare idiosyncratic FA conditions to RA conditions, we included them in the analysis to demonstrate that functions could be identified in cases in which problem behavior was insensitive to consequences programmed into tangible, attention, or escape conditions.

### **Results**

We interpreted preference assessment, RA, and FA results for each participant using visual analysis and summarized our findings in Table 1. We identified a case of correspondence when consequences that controlled target responding during RAs (i.e., produced three demonstrations of effect [as evidenced by deviations from level, trend or variability in responding following condition onset]) also controlled problem behavior during FAs (as determined by criteria described in Roane, Fisher, Kelley, Mevers, & Bouxsein, 2013), and vice versa. We identified a case of non-correspondence if a consequence controlled responding during one analysis but not the other. Across participants, RA outcomes corresponded with FA outcomes for 50% (4 of 8) of tangible conditions, 50% (4 of 8) of attention conditions, and 62.5% (5 of 8) of escape conditions.

Importantly, when evaluating correspondence, we only compared conditions that manipulated identical parameters and qualities of consequences. Thus, if escape from demands (identified from previous demand assessments) did not function as a reinforcer during either analysis but escape from social interactions (identified during parent interviews and included as an idiosyncratic test condition during Becky's FA) functioned as a reinforcer during the FA (see Becky's data in Figure 1), we counted this as a case of correspondence (even though "escape" was implicated during the FA but not during the RA).

To inform discussions of construct validity, we considered cases of correspondence and non-correspondence FA against the full functional profile of each participant's problem behavior.

Specifically, we proposed that cases of “perfect correspondence” (i.e., participants for whom RAs and FAs never produced instances of non-correspondence) contributed evidence to support assumptions of functional equivalence between RAs and FAs. By contrast, we proposed that cases of “imperfect correspondence” (i.e., participants for whom RAs and FAs produced at least one instance of non-correspondence) contributed evidence to support assumptions of functional asymmetry. Thus, using logic inspired by work on divergent construct validity, we proposed that a high percentage of perfect correspondences would challenge the construct validity of contemporary interpretations of FA data and a high percentage of imperfect correspondences would support it.

When considered in this way, obtained results generally produced outcomes which supported the construct validity of contemporary interpretations of FA data. For 25% (2 of 8) of participants (i.e., Becky and Ian), RA results predicted FA results (Figure 1); either by indicating that tangible, attention, and escape would function as reinforcers for problem behavior (Ian), or that they would not (Becky). For 75% (6 of 8) of participants (i.e., Adam, Carrie, Diana, Earl, George, & Heather) RA results did not predict FA results (Figure 2); either because they erroneously predicted tangible, attention, or escape would function as a reinforcer for problem behavior (Adam, Carrie, Earl, George, and Heather), or because they erroneously predicted attention and escape would not (Diana).

Before initiating either analysis, we hypothesized that Heather’s self-injurious behavior (SIB) served a different function than her aggression and we conducted separate FAs of each (shown in the bottom right panels of Figure 2). During the FA of SIB, responding did not demonstrate sensitivity to programmed consequences. Conversely, Heather’s aggression quickly came under the control of contingent attention.

## Discussion

The purpose of this study was to evaluate the likelihood of analogue FAs shaping new functions of problem behavior during the assessment process. Using logic akin to that used in studies of divergent construct validity, we proposed that imperfect correspondence between RAs and FAs would demonstrate that the practice of contriving contingencies between problem behavior and known reinforcers, in and of itself, is likely insufficient to establish new functional relations during FAs (i.e., to produce “false positives”). After completing matched RA and FA conditions for eight participants, we found perfect correspondence between RA and FA outcomes for only 25% of cases, which represented the extremes of what was possible (i.e., all consequences functioned as reinforcers [Ian], or no consequences functioned as reinforcers [Becky]). We interpret this as low and propose that our findings indirectly validate conclusions drawn about behavior function derived from traditional FAs of problem behavior conducted in analogue settings.

This study’s limitations should be noted. First, we did not track rates of problem behavior during RAs or rates of arbitrary responding during FAs; so, it is impossible to determine whether non-targeted behavior occurred during either analysis. Notwithstanding, we do not see this possibility as a major threat to the study’s findings. Had RA targets occurred during FAs, we would have treated them like precursors to problem behavior, allowed them to contact extinction, and expected an escalation to culminate in the occurrence of problem behavior (e.g., Smith & Churchill, 2002). Conversely, if problem behavior occurred during RAs, it is possible that extinction suppressed its occurrence to a level from which it never recovered during the matched FA condition(s), potentially providing an alternative explanation for false positive functions identified by RAs. However, clear functional relations between problem behavior and

maintaining consequences were obtained in the first iterations of non-synthesized FA-test conditions during 66.7% (6 of 9) of our FAs; a percentage that falls well above what might be expected (i.e., 40-50%), given previous work on the topic (Hagopian et al., 2013; Lambert et al., 2017; Slaton, Hanley, & Raftery, 2017). This suggests that problem behavior was at strength at the onset of FAs.

Another limitation of our study was that RAs and FAs differed in several ways. For example, we did not attempt to match problem-behavior-response effort or duration when identifying arbitrary responses during RAs. Furthermore, we used different control techniques across assessments (i.e., withdrawal [RA] and multielement [FA]). Finally, our population was fairly heterogeneous and it is unknown whether a more homogenous sample would have produced greater correspondence. Future researchers may wish to explore whether specific participant characteristics (e.g., age, diagnosis, language ability, sophistication of problem-solving repertoires, etc.) moderate the probability of assessment correspondence.

Notwithstanding these limitations, our results provide evidence suggesting that FA-of-problem-behavior outcomes are a function of more complex variables than the simple arrangement of contingencies between problem behavior and known reinforcers. Although our data do not allow us to examine why, one of two explanations seems plausible. First, participants simply never learned that problem behavior would produce programmed consequences during the FA. Alternatively, it is possible that certain response topographies are less sensitive to some forms for reinforcement than others. In both cases, it does not appear as though problem behavior is likely to come under the control of contingencies of reinforcement that have been contrived but not experienced. We interpret these results as evidence which supports the construct validity of FA methodology. That is, we suggest that low correspondence between RA

results and FA results lends validity to the assumption that functional relations identified during FAs implicate important aspects of individuals' idiosyncratic learning histories.

As a point of discussion, our results appear to contribute to a growing body of evidence that cautions practitioners from using preference- or reinforcer-assessment results, or FAs of alternative behavior, as replacements for FAs of problem behavior. For example, Schieltz et al. (2010), LaRue et al. (2011), and Lambert, Bloom, Kunnavatana, Boyle, & Samaha (2018) all compared the results of FAs of socially appropriate behavior under motivational control (i.e., "mands") with FAs of problem behavior and found varying degrees of correspondence (75% in LaRue et al., 20% in Schieltz et al., and 50% and 0% in Lambert et al.). Similarly, Berg et al. (2007) compared preference assessment results with FAs of problem behavior and obtained 75% correspondence. Our own findings (i.e., 25% perfect correspondence) tell a similar story. The conglomerate of these findings suggests that FAs of alternative behavior (e.g., arbitrary responses, mands, choice responses) provide unreliable predictions of the functions of problem behavior.

Perhaps the most obvious reason for this stems from the fact that reinforcers responsible for maintaining problem behavior represent only a small subset of the total reinforcers available to a given context. Because different responses produce different consequences and because access to problem behavior's functional reinforcers would not necessarily abolish alternative reinforcers (for example, see the behavioral-economics concepts of commodity *substitutes* and *compliments* [Madden, 2000]), it follows that alternative responses (e.g., choice responses, mands, etc.) could be evoked by EOs that do not evoke problem behavior and that problem behavior's abolishing operations might not abolish those alternative responses.

From a practical standpoint, our findings may be useful because they underscore the related but distinct purposes of FAs and RAs. Specifically, RAs are designed to evaluate the potential reinforcing impact of contingent consequences. Although FAs are *capable of* identifying powerful reinforcers, they are more likely than RAs to “miss” them because analysis is constrained to the subset of consequences hypothesized to be functionally related to problem behavior. Thus, RAs are the more appropriate analysis when practitioners need to identify consistently effective instructional consequences (e.g., when working in skill-acquisition paradigms). By contrast, RAs and preference assessments will, on occasion, accurately predict functions of problem behavior (cf. Ian’s results), and can be useful for identifying consequences for inclusion in the experimental conditions of FAs (e.g., Slocum et al., 2018). However, FAs are the more appropriate assessment when awareness of pre-experimental histories is necessary to inform treatment components whose efficacy depends on function.

The potential adverse repercussions of assuming equivalence between RA results and FA results may be best typified by Heather’s case. Not only did RA results incorrectly predict the function of either topography of problem behavior (i.e., SIB and aggression), we also did not have a clear method for identifying which topography to assign the implicated function. Had we not thought to conduct separate FAs for SIB and aggression, we may have assumed that both shared the same function. Likewise, had we equated RA results with FA results, we may have assumed that either or both had a tangible function. All of these assumptions would have been wrong and function-dependent treatment components (e.g., extinction) informed by them would be unlikely to generate favorable therapeutic outcomes (Iwata, Pace, Cowdery, & Miltenberger, 1994).

A few additional points may be worth discussing. First, we anticipated that the most likely disagreement between analyses would occur when RAs identified a reinforcer that did not maintain problem behavior during FAs (i.e., RAs would implicate a false-positive function). Thus, we arranged our experiment in such a way to decrease the possibility of this occurring. That is, we always conducted RAs before FAs and FA sessions (5 min) were 2.5 times longer than RA sessions (2 min). Thus, participants had always experienced programmed consequences multiple times before any FA began. Furthermore, in cases in which RAs identified reinforcers that did not maintain problem behavior during FAs, there was clear evidence that exposure to programmed EOs could evoke high rates of independent responding across a time frame (i.e., 2 min) that was much briefer than was allotted during FA sessions (i.e., 5 min). As a result, it would be hard to argue that lack of responding during FA sessions was due to insufficient exposure to EOs (as is sometimes the case; cf. Kahng, Abt, & Schonbachler, 2001).

Because we anticipated that RAs would only implicate false positive functions for problem behavior, we were surprised to see that they implicated a false-negative for Diana. That is to say, Diana would not clap to escape from low-probability demands but she would aggress. Although we cannot draw strong conclusions from her data because RA sessions were considerably shorter than FA sessions (and Diana may have eventually clapped with more time), it is worth noting that RA sessions were conducted in quick succession (thus, Diana was exposed to 10 min of near-continuous demands). It is also worth noting that Diana was reminded she could clap to escape every 30 s (and had experienced this contingency during forced-choice exposures) and that these same demands evoked aggression during the first escape session of her FA (unfortunately, we did not track aggression during her escape RA and cannot say whether aggression occurred, and perhaps competed with clapping, during that assessment). Thus,

although any conclusions about her results would only be speculative, the possibility that reinforcers might maintain some topographies of behavior but not others (in spite of instructions and direct experience with relevant contingencies) suggests a certain inflexibility in operant membership that might further protect against invalid conclusions of problem-behavior function drawn from FA data. This could be worth exploring in future research.

It is likely important to note that tangible functions of problem behavior were only identified during 37.5% (3 of 8) of cases and were identified for less than half of participants for whom sensitivity to tangible reinforcement was demonstrated during RAs (i.e., 7 participants). Tangible functions were less prevalent than escape functions, (4 of 8 cases) but were more prevalent than attention functions (2 of 8 cases). In a review on multiply controlled problem behavior conducted by Beavers and Iwata (2011), the authors noted a relatively high percentage of confirmed tangible functions (83.7%) amongst FAs analyzed, when they considered the number of confirmed functions (41 cases) against the number of opportunities to evaluate function (49 FAs which included a tangible condition). Beavers and Iwata, and Rooker et al. (2011), both presented two hypotheses to account for this. One was that researchers only included tangible conditions in published FAs when there was good reason to believe that problem behavior served a tangible function. The other was that tangible conditions are more likely than other FA conditions to produce false-positive functions.

Although the findings of Rooker et al. and Shirley et al. (1999) suggest the latter may be the case for automatically-maintained problem behavior (but see Dozier, Iwata, Wilson, Thomason-Sassi, & Roscoe, 2013 for contradictory findings), our own results (both in terms of general non-correspondence between tangible RA and FA outcomes and in terms of the moderate percentage of cases in which tangible functions were identified) suggest a more

plausible account is that published FAs have only included tangible conditions when tangible functions have been suspected and have not included them otherwise. To the extent to which this is true, including tangible conditions in FAs of problem behavior may not be as detrimental to the validity of FA results as previously indicated. However, as we did not explicitly design our study to test either hypothesis, additional research on the topic appears to be warranted.

### References

- Beavers, G. A., & Iwata, B. A. (2011). Prevalence of multiply controlled challenging behavior. *Journal of Applied Behavior Analysis, 44*, 593-597. doi:10.1901/jaba.2011.44-593
- Beavers, G. A., Iwata, B. A., & Lerman, D. C. (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis, 46*, 1-21. doi:10.1002/jaba.30
- Berg, W. K., Wacker, D. P., Cigrand, K., Merkle, S., Wade, J., Henry, K., & Wang, Y. C. (2007). Comparing functional analysis and paired-choice assessment results in classroom settings. *Journal of Applied Behavior Analysis, 40*, 545-552.
- Buchmeier, A. L., Baker, J. C., Reuter-Yuill, L. M., & MacNeill, B. R. (2018). Considerations for preference and reinforcer assessments with older adults with developmental disabilities. *Behavior Analysis: Research and Practice, 18*, 103-116. doi: [10.1037/bar0000085](https://doi.org/10.1037/bar0000085)
- Clay, C. J., Samaha, A. L., Bloom, S. E., Bogoev, B. K., & Boyle, M. A. (2013). Assessing preference for social interactions. *Research in Developmental Disabilities, 34*, 362-371. doi: 10.1016/j.ridd.2012.07.028
- Call, N. A., Findley, A. J., & Reavis, A. R. (2012). The effects of conducting a functional analysis on problem behavior in other settings. *Research in Developmental Disabilities, 33*, 1990-1995. doi: 10.1016/j.ridd.2012.06.001
- Call, N. A., Reavis, A. R., Clark, S. B., Parks, N. A., Cariveau, T., & Muething, C. S. (2017). The effects of conducting a functional analysis on problem behavior in other settings: A descriptive study on potential interaction effects. *Behavior Modification, 41*, 609-625. doi: 10.1177/0145445517696052

- Call, N. A., Trosclair-Lasserre, N. M., Findley, A. J., Reavis, A. R., & Shillingsburg, M. A. (2012). Correspondence between single versus daily preference assessment outcomes and reinforcer efficacy under progressive-ratio schedules. *Journal of Applied Behavior Analysis, 45*, 763-777. doi: [10.1901/jaba.2012.45-763](https://doi.org/10.1901/jaba.2012.45-763)
- Conners, J., Iwata, B. A., Kahng, S., Hanley, G. P., Worsdell, A. S., & Thompson, R. H. (2000). Differential responding in the presence and absence of discriminative stimuli during multielement functional analyses. *Journal Of Applied Behavior Analysis, 33*, 299-308. doi: [10.1901/jaba.2000.33-299](https://doi.org/10.1901/jaba.2000.33-299)
- Cote, C. A., Thompson, R. H., Hanley, G. P., & McKerchar, P. M. (2007). Teacher report and direct assessment of preferences for identifying reinforcers for young children. *Journal of Applied Behavior Analysis, 40*, 157-166. doi: [10.1901/jaba.2007.177-05](https://doi.org/10.1901/jaba.2007.177-05)
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis (2nd ed.)*. Upper Saddle River, NJ: Pearson.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin, 52*, 281. doi: [10.1037/h0040957](https://doi.org/10.1037/h0040957)
- Catania, A. C. (2013). *Learning (5<sup>th</sup> ed.)*. Cornwall-on-Hudson, NY: Sloan publishing, LLC.
- 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. doi:[10.1901/jaba.1996.29-519](https://doi.org/10.1901/jaba.1996.29-519)
- Dempsey, C. M., Iwata, B. A., Fritz, J. N. & Rolider, N. U. (2012). Observer training revisited: a comparison of in vivo and video instruction. *Journal of Applied Behavior Analysis, 45*, 827-832. doi: [10.1901/jaba.2012.45-827](https://doi.org/10.1901/jaba.2012.45-827)

- Dozier, C. L., Iwata, B. A., Wilson, D. M., Thomason-Sassi, J. L., & Roscoe, E. M. (2013). Does supplementary reinforcement of stereotypy facilitate extinction?. *Journal of applied behavior analysis*, 46, 242-255.
- 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. doi: 10.1901/jaba.1992.25-491
- Galiatsatos, T. G., & Graff, R. (2003). Combining descriptive and functional analyses to assess and treat screaming. *Behavioral Interventions*, 18, 123-138. doi: 10.1002/bin.133
- Hagopian, L. P., LeBlanc, L. A., & Maglieri, K. A. (2000). Noncontingent attention for the treatment of excessive medical complaints in a medically fragile man with mental retardation. *Research in Developmental Disabilities*, 21, 215–221. doi: 10.1016/S0891-4222(00)00035-4
- Hagopian, L. P., Rooker, G. W., Jessel, J., & DeLeon, I. G. (2013). Initial functional analysis outcomes and modifications in pursuit of differentiation: A summary of 176 inpatient cases. *Journal of Applied Behavior Analysis*, 46, 88-100. doi:10.1002/jaba.25
- 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. doi: 10.1007/BF03391818
- 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. doi:10.1901/jaba.2003.36-147

- Hanley, G. P., Jin, C. S., Vanselow, N. R., & Hanratty, L. A. (2014). Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments. *Journal of Applied Behavior Analysis, 47*, 16-36. doi: 10.1002/jaba.106
- Hastings, R. P., & Noone, S. J. (2005). Self-injurious behavior and functional analysis: ethics and evidence. *Education and Training in Developmental Disabilities, 40*, 335-342.
- Hodos, W. (1961). Progressive ratio as a measure of reward strength. *Science, 134*, 943-944. doi:10.1126/science.134.3483.943
- Hurl, K., Wightman, J., Haynes, S. N., & Virues-Ortega, J. (2016). Does a pre-intervention functional assessment increase intervention effectiveness? A meta-analysis of within-subject interrupted time-series studies. *Clinical Psychology Review, 47*, 71-84.
- Iwata, B.A., DeLeon, I.G. & Roscoe, E.M. (2013). Reliability and validity of the functional analysis screening tool. *Journal of Applied Behavior Analysis, 46*, 271-284. doi:10.1002/jaba.31
- 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. doi:10.1901/jaba.1994.27-197.
- Iwata, B.A., & Dozier, C.L. (2008). Clinical application of functional analysis methodology. *Behavior Analysis in Practice, 1*, 1-9.
- 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. doi: 10.1901/jaba.1994.27-131

- Jessel, J., Hanley, G. P., and Ghaemmaghmi, M. (2016). Interview-informed synthesized contingency analyses: Thirty replications and reanalysis, *Journal of Applied Behavior Analysis*, 49, 576-595. doi: 10.1002/jaba.316
- Jessel, J., Hausman, N. L., Schmidt, J. D., Darnell, L. C., & Kahng, S. (2014) The development of false-positive outcomes during functional analyses of problem behavior. *Behavioral Interventions*, 29, 50-61. doi: 10.1002/bin.1375
- Jessel, J., Ingvarsson, E. T., Metras, R., Kirk, H., & Whipple, R. (2018). Achieving socially signification reductions in problem behavior following interview-informed synthesized contingency analysis: a summary of 25 outpatient applications. *Journal of Applied Behavior Analysis*, 51, 130-157. doi: 10.1002/jaba.436
- Kahng, S. W., Abt, K. A., & Schonbachler, H. E. (2001). Assessment and treatment of low-rate high-intensity problem behavior. *Journal of Applied Behavior Analysis*, 34, 225-228. doi: 10.1901/jaba.2001.34-225
- Kelly, M. A., Roscoe, E. M., Hanley, G. P., and Schlichenmeyer, K (2014). Evaluation of assessment methods for identifying social reinforcers, *Journal of Applied Behavior Analysis*, 47, 113-135. doi: 10.1002/jaba.107
- Lambert, J.M., Bloom, S.E., Kunnavatana, S.S., Boyle, M., & Samaha, A.L. (2018). Evaluating correspondence between functional analyses of mands and functional analyses of problem behavior. *European Journal of Behavior Analysis*. doi: [10.1080/15021149.2018.1535162](https://doi.org/10.1080/15021149.2018.1535162)
- Lambert, J. M., Staubitz, J. E., Torelli Roane, J., Houchins-Juárez, N. J., Juárez, A. P., Sanders, K. B. and Warren, Z. E. (2017). Outcome summaries of latency-based functional analyses

- conducted in hospital inpatient units. *Journal of Applied Behavior Analysis*, 50, 487–494.  
doi:10.1002/jaba.399
- LaRue, R. H., Sloman, K. N., Weiss, M., Delmolino, L., Hansford, A., Szalony, J., & ...  
Lambright, N. M. (2011). Correspondence between traditional models of functional  
analysis and a functional analysis of manding behavior. *Research In Developmental  
Disabilities*, 32, 2449-2457.
- Lerman, D. C., Kelley, M. E., Vorndran, C. M., Kuhn, S. A. C., & LaRue Jr., R. H. (2002).  
Reinforcement magnitude and responding during treatment with differential  
reinforcement, *Journal of Applied Behavior Analysis*, 35, 29-48. doi:  
10.1901/jaba.2002.35-29
- Mace, F. C. (1994). The significance and future of functional analysis methodologies. *Journal of  
Applied Behavior Analysis*, 27, 385–392. doi:10.1901/jaba.1994.27-385
- Madden, G. (2000). A behavioral economics primer. In Bickel, W. K., & Vuchinich, R. E.  
(Eds.). *Reframing health behavior change with behavioral economics* (pp. 3-26).  
Mahwah, NJ: Taylor & Francis
- Martin, N. T., Gaffan, E. A., & Williams, T. (1999). Experimental functional analyses for  
challenging behavior: a study of validity and reliability. *Research in Developmental  
Disabilities*, 20, 125-146. doi: 10.1016/S0891-4222(98)00037-7
- Oliver, A. C., Pratt, L. A., & Normand, M. P (2015). A survey of functional behavior assessment  
methods used by behavior analysts in practice. *Journal of Applied Behavior Analysis*, 48,  
817–829. doi: 10.1002/jaba.256

- O'Neill, R.E., Horner, R. H., Albin, R. W., Sprague, J. R., Storey, K., & Newton, J. S. (1997). *Functional Assessment and Program Development for Problem Behavior*. Pacific Grove, CA: Brooks/Cole Publishing.
- Pence, S. T., Roscoe, E. M., Bourret, J. C., & Ahearn, W. H. (2009). Relative contributions of three descriptive methods: implications for behavioral assessment. *Journal of Applied Behavior Analysis, 42*, 425-446. doi: [10.1901/jaba.2009.42-425](https://doi.org/10.1901/jaba.2009.42-425)
- Quick, M. J., Baker, J. C., & Ringdahl, J. E. (2018). Assessing the validity of engagement-based and selection-based preference assessments in elderly individuals with neurocognitive disorder. *Behavior Analysis: Research and Practice, 18*, 92-102. doi: [10.1037/bar0000070](https://doi.org/10.1037/bar0000070)
- Rooker, G. W., Iwata, B. A., Harper, J. M., Fahmie, T. A., & Camp, E. M. (2011). False-positive tangible outcomes of functional analyses. *Journal of Applied Behavior Analysis, 44*, 737-745. doi: [10.1901/jaba.2011.44-737](https://doi.org/10.1901/jaba.2011.44-737)
- Roane, H. S., Vollmer, T. R., Ringdahl, J. E. and Marcus, B. A. (1998). Evaluation of a brief stimulus preference assessment. *Journal of Applied Behavior Analysis, 31*, 605–620. doi: [10.1901/jaba.1998.31-605](https://doi.org/10.1901/jaba.1998.31-605)
- Roane, H. S., Fisher, W. W., Kelley, M. E., Mevers, J. L., & Bouxsein, K. J. (2013). Using modified visual-inspection criteria to interpret functional analysis outcomes. *Journal of Applied Behavior Analysis, 46*, 130-146. doi: [10.1002/jaba.13](https://doi.org/10.1002/jaba.13)
- Roscoe, E.M., Phillips, K.M., Kelly, M.A., Farber, R., & Dube, W.V. (2015). A statewide survey assessing practitioners' use and perceived utility of functional assessment. *Journal of Applied Behavior Analysis, 48*, 830-844. doi: [10.1002/jaba.259](https://doi.org/10.1002/jaba.259)

- Roscoe, E.M., Rooker, G.W., Pence, S.T., & Longworth, L.K. (2009). Assessing the utility of a demand assessment for functional assessment for functional analysis. *Journal of Applied Behavioral Analysis, 42*, 819-825. doi: 10.1901/jaba.2009.42-819
- Schieltz, K., Wacker, D., Harding, J., Berg, W., Lee, J., & Padilla Dalmau, Y. (2010). An evaluation of manding across functions prior to functional communication training. *Journal of Developmental & Physical Disabilities, 22*, 131-147. doi:10.1007/s10882-009-9181-5
- Schlichenmeyer, K. J., Roscoe, E. M., Rooker, G. W., Wheeler, E. E., & Dube, W. V. (2013). Idiosyncratic variables that affect functional analysis outcomes: a review (2001-2010). *Journal of Applied Behavioral Analysis, 46*, 339-348. doi: 10.1002/jaba.12
- Shabani, D., Carr, J., Pabico, R., Sala, A., Lam, W., & Oberg, T. (2013). The effects of functional analysis test sessions on subsequent rates of problem behavior in the natural environment. *Behavioral Interventions, 28*, 40-47. doi: 10.1002/bin.1352
- Shirley, M. J., Iwata, B. A., & Kahng, S. W. (1999). False-positive maintenance of self-injurious behavior by access to tangible reinforcers. *Journal of Applied Behavior Analysis, 32*, 201-04. doi: 10.1901/jaba.1999.32-201
- Slaton, J. D., Hanley, G. P., and Raftery, K. J. (2017). Interview-informed functional analyses: A comparison of synthesized and isolated components, *Journal of Applied Behavior Analysis, 50*, 252-277. doi: 10.1002/jaba.384
- Slocum, S. K., Grauerholz-Fisher, E., Peters, K. P., & Vollmer, T. R. (2018). A multicomponent approach to thinning reinforcer delivery during noncontingent reinforcement schedules. *Journal of Applied Behavior Analysis, 51*, 61-69. doi: 10.1002/jaba.427

- Smith, R. G., & Churchill, R. M. (2002). Identification of environmental determinants of behavior disorders through functional analysis of precursor behaviors. *Journal Of Applied Behavior Analysis, 35*, 125-136. doi: 10.1901/jaba.2002.35-125
- Solnick, M., & Ardoin, S. (2010). A quantitative review of functional analysis procedures in public school settings. *Education and Treatment of Children, 33*, 153-175
- Sturme, P. (1995). Analogue baselines: A critical review of the methodology. *Research in Developmental Disabilities, 16*, 269-284. doi: 10.1016/0891-4222(95)00014-E
- Sigafoos, J., & Sagers, E. (1995). A discrete-trial approach to the functional analysis of aggressive behaviour in two boys with autism. *Journal of Intellectual and Developmental Disability, 20*, 287-297. doi:10.1080/07263869500035621
- Tarbox, R. S., Tarbox, J., Ghezzi, P. M., Wallace, M. D., & Yoo, J. H. (2007). The effects of blocking mouthing of leisure items on their effectiveness as reinforcers. *Journal of Applied Behavior Analysis, 40*, 761-765. doi: [10.1901/jaba.2007.761-765](https://doi.org/10.1901/jaba.2007.761-765)
- Tiger, J. H., Miller, S. J., Mevers, J. L., Mintz, J. C., Scheithauer, M. C., & Alvarez, J. (2013). On the representativeness of behavior observation samples in classrooms. *Journal of Applied Behavior Analysis, 46*, 424-435. doi: [10.1002/jaba.39](https://doi.org/10.1002/jaba.39)
- Yoder, P. J., Symons, F. J., & Lloyd, B. P. (2018). *Observational measurement of behavior (2<sup>nd</sup> Ed.)*. Brookes Publishing.

Table 1

*Participant and Outcome Summaries*

Name	Age	Diagnosis	Language Ability	Arbitrary Bx	Problem Bx	Condition	Preference Assessment	Reinforcer Analysis	Functional Analysis	Correspondence
Adam	23	ASD	4	Table Wipe	Agg	Tangible: Video clips		X		No
						Attention: Handshake				Yes
						Escape: Stand/sit		X	X	Yes
Becky	12	ASD	3	Elbow Touch	Agg	Tangible: Music				Yes
						Attention: Reprimand		--		Yes
						Escape: 2-step directives			--	Yes
Carrie	19	Schizencephaly, CP, ADHD, OCD, SPD, Anxiety, & DD	4	Table Wipe	Agg	Tangible: Music		X		No
						Attention: Fist bump		X		No
						Escape: Stack books		X	X	Yes
Diana	19	Down Syndrome	2	Clap	Agg	Tangible: Music		X	X	Yes
						Attention: Reprimand		--	X	No
						Escape: Stand/sit			X	No
Earl	16	ASD	4	Cup Flip	Persev. Speech	Tangible: Music		X	X	Yes
						Attention: Talk		X		No
						Escape: Stand/sit		X		No
George	18	ASD, Down Syndrome	3	Elbow Touch	PD	Tangible: iPod		X		No
						Attention: Hug		X		No
						Escape: Open/close book				Yes
Heather	7	Prader-Willi Syndrome	2	Clap	SIB	Tangible: Blanket		X		No
						Attention: High five				Yes
						Escape: Puzzle		X		No
Ian	6	ASD	4	Nose Touch	Agg	Tangible: iPad & play doh		X	X	Yes
						Attention: "Quiet" tickles		X	X	Yes
						Escape: Stand/sit		X	X	Yes

Note: X = functional relation confirmed (e.g., 3 or more demonstrations of effect); -- = functional relation suggested but not confirmed (e.g., only two demonstrations of effect); Bx = behavior; ASD = Autism Spectrum Disorder; CP = Cerebral Palsy, ADHD = Attention Deficit Hyperactivity Disorder; OCD = Obsessive-Compulsive Disorder; SPD = Sensory Processing Disorder; DD = Developmental Delay; Language ability = 1= non-vocal, 2 = 1-word utterances, 3 = short diffluent sentences, 4 = full fluency; Agg = aggression; SIB = Self-Injurious Behavior; PD = Property Destruction; Persev = Perseverative

Table 2

*Interobserver Agreement and Procedural Fidelity*

		Pref. Assessments		Reinforcer Analysis		Functional Analysis	
		IOA	PF	IOA	PF	IOA	PF
Adam	Tangible	90.3%	100%	91.7% (100%)	100% (83.3%)		
	Attention	100%	100%	100% (100%)	100% (100%)	97.5% (94.7%)	99% (100%)
	Escape	100%	100%	93.7% (100%)	98.9% (91.7%)		
Becky	Tangible	95%	100%	100% (75%)	100% (100%)		
	Attention	83.3%	100%	97.2% (100%)	100% (100%)	99.6% (83.3%)	100% (83.3%)
	Escape	92%	100%	99% (100%)	100% (100%)		
Carrie	Tangible	99%	86%	96.3% (100%)	96.3% (100%)		
	Attention	100%	100%	100% (100%)	98.3% (100%)	99.6% (95%)	100% (75%)
	Escape	100%	100%	97.1% (100%)	100% (100%)		
Diana	Tangible	100%	95.2%	100% (64.3%)	100% (57.1%)		
	Attention	100%	100%	99.4% (85.7%)	100% (85.7%)	98.3% (62.5%)	98.4% (50%)
	Escape	96%	97%	100% (100%)	100% (25%)		
Earl	Tangible	100%	100%	95% (35.7%)	100% (64.3%)		
	Attention	98%	100%	92.5% (57.1%)	100% (85.7%)	99.5% (70.5%)	97.7% (63.6%)
	Escape	100%	100%	99.1% (63.6%)	100% (72.7%)		
George	Tangible	100%	100%	91.1% (86.7%)	100% (46.7%)		
	Attention	100%	100%	95.5% (83.3%)	100% (22.2%)	99.1% (86.1%)	98.2% (38.9%)
	Escape	100%	100%	100% (92.3%)	100% (61.5%)		
Heather	Tangible	100%	100%	97.6% (80%)	100% (80%)		
	Attention	100%	100%	100% (100%)	100% (50%)	99% (92.1%)	99.4% (80.4%)
	Escape	96%	100%	95.3% (75%)	100% (56.3%)		
Ian	Tangible	100%	95%	100% (50%)	100% (58.3%)		
	Attention	100%	-	100% (25%)	100% (66.7%)	96% (75%)	97.8% (100%)
	Escape	100%	-	100% (50%)	100% (41.7%)		

Note: Non-parenthetical percentages represent mean interobserver agreement and procedural fidelity scores for each assessment. Parenthetical percentages represent the percentage of total sessions for which interobserver agreement and procedural fidelity was evaluated. Pref = preference; IOA = interobserver agreement; PF = procedural fidelity.

### Figure Captions

*Figure 1:* Reinforcer and functional analysis results for two correspondence cases (Ian & Becky). Target responses are listed below graphs in lower left-hand corner of each relevant column. BL = baseline and FR1 = fixed-ratio 1. Note that y-axis scales differ for RA graphs (left three columns) and FA graphs (far right column).

*Figure 2:* Reinforcer and functional analysis results for six non-correspondence cases (Adam, Carrie, Diana, Earl, George, & Heather). Target responses are listed below graphs in lower left-hand corner of each relevant column. BL = baseline and FR1 = fixed-ratio 1. Note that y-axis scales differ for RA graphs (left three columns) and FA graphs (far right column).

# Reinforcer Analysis

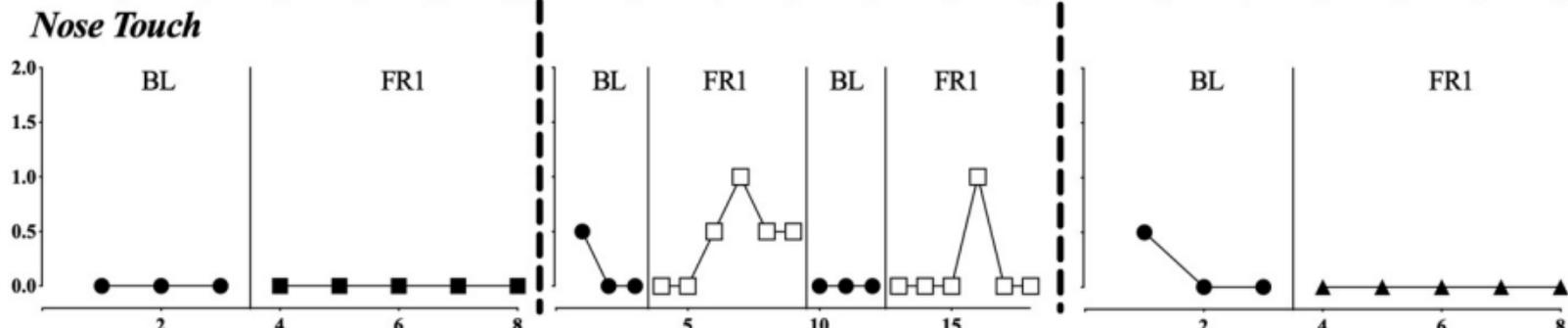
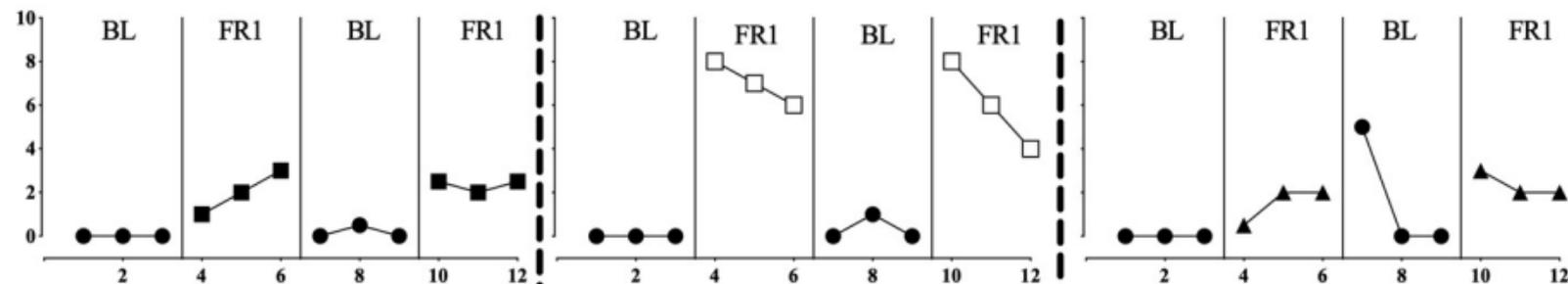
## Tangible

## Attention

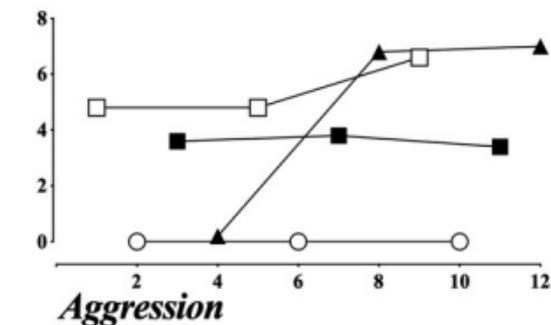
## Escape

# Functional Analysis

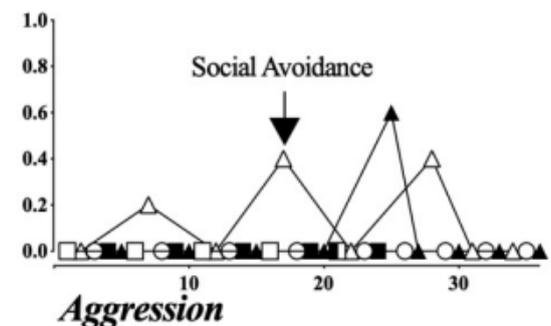
RESPONSES PER MINUTE



SESSIONS



Ian



Becky

# Reinforcer Analysis

## Tangible

## Attention

## Escape

# Functional Analysis

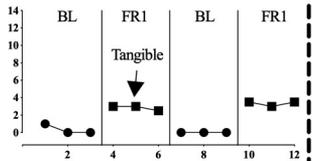


Table Wipe

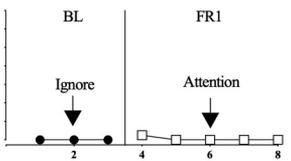
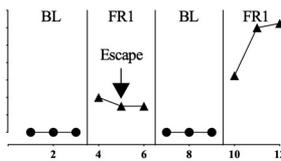
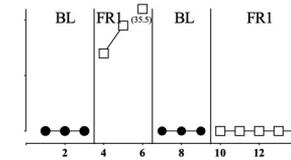


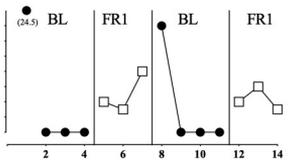
Table Wipe



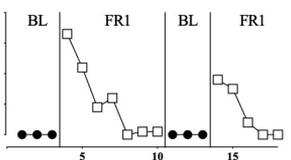
Clap



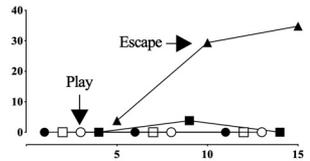
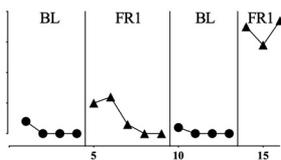
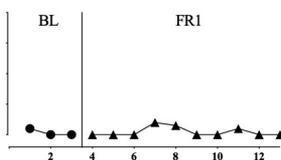
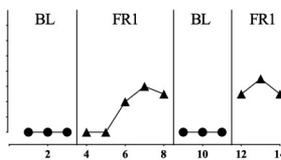
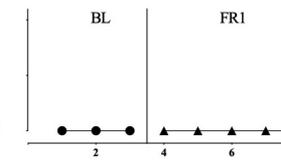
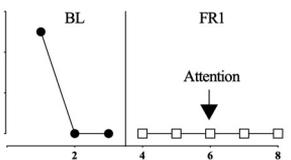
Cup Flip



Elbow Touch

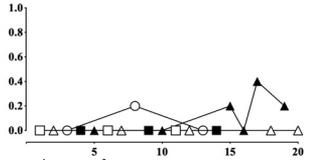


Clap



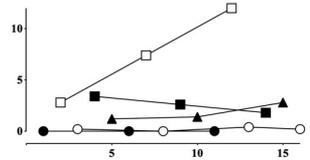
Aggression

Adam



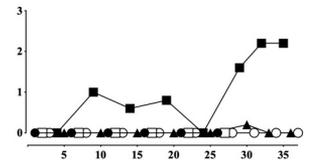
Aggression

Carrie



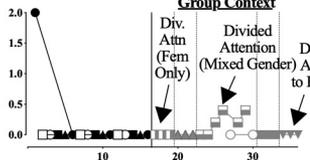
Aggression

Diana



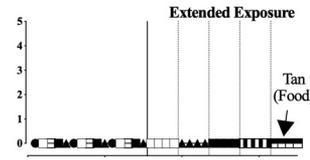
Perseverative Speech

Earl



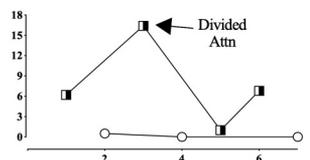
Property Destruction

George



SIB (FA1)

Heather



Aggression (FA2)

RESPONSES PER MINUTE

SESSIONS