

AN EVALUATION OF TRIAL-BASED FUNCTIONAL ANALYSES  
IN CLASSROOM SETTINGS

By

Blair P. Lloyd

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

in

Special Education

August, 2013

Nashville, Tennessee

Approved:

Dr. Joseph H. Wehby

Dr. Erik W. Carter

Dr. Craig H. Kennedy

Dr. Paul J. Yoder

## ACKNOWLEDGEMENTS

I extend my deepest gratitude to the following faculty members who have served on my doctoral committee and offered their thoughtful feedback and guidance throughout this and other research projects: Dr. Erik Carter, Dr. Craig Kennedy, Dr. Jeremy Veenstra-Vanderweele, Dr. Joseph Wehby, Dr. Mark Wolery, and Dr. Paul Yoder. I also thank Samantha Goldman, Michelle Harvey, Daniel Sherlock, and Emily Weaver for their assistance with data collection and commitment to delivering high-quality services to the students and teachers who participated in this project. I am grateful to the school administrators, behavior specialists, teaching staff, students, and parents who made this study possible. Finally, I thank the Semmel Family for their generous financial contributions, which have supported many doctoral dissertations including my own.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	ii
LIST OF TABLES .....	v
LIST OF FIGURES .....	vi
Chapter	
I. INTRODUCTION.....	1
II. METHOD .....	11
Participants and Settings.....	11
Response Definitions.....	14
Student behaviors .....	14
Teacher behaviors.....	16
Data Collection.....	17
Descriptive assessments .....	17
Experimental trials.....	19
Procedural fidelity .....	19
Observer Training and Reliability .....	21
Experimental Design .....	24
Study Procedures .....	25
TA interview.....	25
Descriptive assessments .....	25
TA training on assessment procedures .....	26
Trial-based FA.....	27
TA training on intervention procedures .....	33
Intervention.....	33
TA questionnaire .....	39
Data Analysis.....	39
Descriptive assessments .....	39
Experimental trials.....	41
Trial durations .....	41

III. RESULTS.....	43
Descriptive Assessments .....	43
Abhi .....	43
Sid .....	45
Davis .....	47
Gretchen .....	49
Trial-based FA .....	51
Abhi .....	51
Sid .....	52
Davis .....	54
Gretchen .....	55
Intervention.....	58
Abhi .....	59
Sid .....	60
Davis .....	61
Gretchen .....	63
Social Validity Measures.....	67
TA questionnaire .....	67
Trial durations .....	69
IV. DISCUSSION .....	71
Challenges and Limitations .....	72
Future Research .....	77
Conclusion.....	81
Appendix	
A. DESCRIPTIVE ASSESSMENT DATA COLLECTION FORM.....	82
B. TBFA FIDELITY DATA COLLECTION FORM.....	83
C. INTERVENTION DATA COLLECTION FORM .....	84
REFERENCES.....	85

## LIST OF TABLES

Table	Page
1. Student (top) and TA (bottom) participant demographics .....	12
2. Procedural fidelity and inter-observer agreement on fidelity of experimental trials .....	21
3. Inter-observer agreement on descriptive assessment data .....	23
4. TA responses on social validity questionnaire .....	68
5. Duration of trial-based FAs and interventions by participant .....	70

## LIST OF FIGURES

Figure	Page
1. Results of descriptive assessments for Abhi .....	44
2. Results of descriptive assessments for Sid .....	46
3. Results of descriptive assessments for Davis .....	48
4. Results of descriptive assessments for Gretchen.....	50
5. Results of trial-based FA for Abhi .....	52
6. Results of trial-based FA for Sid .....	53
7. Results of trial-based FA for Davis .....	55
8. Results of trial-based FA for Gretchen including all problem behaviors.....	56
9. Results of trial-based FA for Gretchen excluding out-of-seat behavior.....	58
10. Results of intervention trials for Abhi .....	59
11. Results of intervention trials for Sid.....	61
12. Results of intervention trials for Davis.....	62
13. Results of intervention trials for Gretchen including all problem behaviors .....	64
14. Results of intervention trials for Gretchen excluding out-of-seat behavior .....	65

## CHAPTER I

### INTRODUCTION

The functional analysis (FA) is an experimental assessment approach used to identify environmental variables evoking and maintaining problem behavior. FA procedures based on those developed by Iwata, Dorsey, Slifer, Bauman, and Richman (1982/1994) have become the standard experimental assessment method to identify the function(s) of problem behavior and develop function-based interventions. These FA procedures include the systematic manipulation of environmental antecedents and consequences of problem behavior across multiple test conditions and one control condition, all within the context of highly controlled settings. Each test condition is designed to evaluate a hypothesized reinforcer. A condition testing for positive reinforcement in the form of adult attention may involve restricted adult attention as an antecedent and attention delivery as a consequence for problem behavior. A condition testing for negative reinforcement in the form of escape from instructions may involve the delivery of task instructions as an antecedent and removal of task instructions as a consequence for problem behavior. A control condition typically involves the delivery of all hypothesized reinforcers as antecedents with no programmed consequences to minimize the likelihood of problem behavior. Increased levels of problem behavior in one or more test conditions relative to the control condition leads to confirmation of the corresponding reinforcement hypotheses.

FA procedures were originally designed for individuals with significant disabilities who engaged in severe topographies of problem behavior (e.g., self-injury, aggression) in the context of controlled settings (e.g., hospitals, institutions, large clinical facilities; Iwata et al., 1982/1994;

Iwata, Pace, Kalsher, Cowdery, & Cataldo, 1990; Mace, Page, Ivancic, & O'Brien, 1986). Over time, however, more youth with significant disabilities began transitioning to less restrictive educational settings as a result of special education reform (IDEA, 1990; U.S. Department of Education, 1995). In addition, a broader collection of assessment procedures (both non-experimental and experimental) used to evaluate environmental variables related to problem behavior (i.e., *functional assessment* methods) were extended to a wider range of behaviors and made accessible to service providers, including educators and practitioners (e.g., O'Neill, Horner, Albin, Storey, & Sprague, 1990). In 1997, the reauthorization of the Individuals with Disabilities Education Act (IDEA, 1990) mandated that IEP teams use functional assessment methods to develop and/or revise behavioral interventions before a student may be transferred to a more restrictive educational placement. These amendments, however, did not elaborate on how functional assessments and function-based interventions should be conducted. These legislative changes added a sense of urgency to establish practical and effective methods for conducting functional assessments in the context of school settings.

Attempts to implement the standard FA procedures in school settings have met a variety of challenges (e.g., Lalli, Browder, Mace, & Brown, 1993). Systematic manipulations in general are more difficult to implement in classrooms as compared to highly controlled laboratory and clinical settings (Ellis & Magee, 1999). School administrators may be unwilling to allow environmental manipulations in which differential levels of problem behavior are predicted, as some of these manipulations necessarily involve setting the occasion for increased rates of problem behavior (Desrochers, Hile, & Williams-Moseley, 1997; Repp, 1994). In addition, a number of procedural elements associated with traditional FA procedures present challenges when applied in a classroom context. First, FAs take considerable time to complete. Multiple



exposures to each of several experimental conditions with session durations of 10-15 min typically require several hours of total assessment time (Iwata et al., 1994; Lloyd, in preparation; Lydon, Healy, O'Reilly, & Lang, 2012; Tincani, Castrogiovanni, & Axelrod, 1999). Second, FA procedures are complex and labor-intensive (Axelrod, 1987; Doss & Reichle, 1989). Distinct combinations of antecedents and response-dependent consequences must be implemented continuously throughout sessions and with consistency across conditions. Data collection methods also require continuous observation throughout all sessions. Third, FAs may pose increased risk to the participants involved, as procedures involve conditions that repeatedly evoke and reinforce problem behaviors (Najdowski, Wallace, Ellsworth, MacAleese, & Cleveland, 2008).

For many of the reasons described above, the majority of FAs in school settings have been conducted in analogue environments (i.e., areas separate from the natural classroom setting, including empty or auxiliary classrooms) rather than students' typical classrooms (Ervin et al., 2001; Solnick & Ardoin, 2010). Analogue settings allow increased control of antecedent and consequent variables while minimizing the influence of extraneous variables. In addition, analogue settings eliminate risks to classmates in cases of high-intensity aggressive or disruptive behavior. The use of analogue settings, however, often is not feasible due to limited resources (e.g., physical space, additional staff) and may not be desirable when students are removed from their instructional setting for extended periods of time.

In addition, it is unclear whether results of assessments conducted in analogue settings by experimenters would be relevant to the natural settings in which behavioral support is needed. In fact, discrepancies in the outcomes of FAs conducted with different contextual variables have been identified in the literature. FA outcomes have been shown to vary based on setting (e.g.,

Lang et al., 2008, 2009), the individuals delivering consequences (English & Anderson, 2006; Ringdahl & Sellers, 2000), and idiosyncratic tangible stimuli (Carr, Yarbrough, & Langdon, 1997). Analogue FAs also may produce false positive outcomes when programmed consequences are included that do not reflect those typical of the natural setting (e.g., Anderson & Long, 2002). For these reasons, conducting assessments in the same context in which intervention is needed preserves the ecological validity of assessment results (Conroy, Fox, Crain, Jenkins, & Belcher, 1996). In fact, preliminary evidence suggests interventions based on results of assessments conducted in natural classroom settings may be more effective than those based on assessments conducted in separate settings (English & Anderson, 2006), though additional research is needed to address this question.

Despite the challenges associated with conducting experimental analyses in natural settings, the identification of practical methods to test hypotheses of behavior function in classrooms remains critical. The trial-based FA (Sigafoos & Sagers, 1995) has been identified as a variation on traditional FA methodology (Lloyd, in preparation; Lydon et al., 2012). The trial-based FA is a method of testing response-reinforcer relations via brief trials that are distributed over time and embedded into naturally-occurring routines. The first trial-based FAs (Sigafoos & Meikle, 1996; Sigafoos & Sagers, 1995) were designed as a variation of the traditional FA developed by Iwata et al. (1982/1994) and were conducted in classrooms of a school and therapy center for children with autism. This approach incorporated the assessment procedures into natural classroom routines. Each trial lasted up to 2 min and consisted of a test interval and a control interval. During the test interval, a hypothesized establishing operation

(EO)<sup>1</sup> was present until the first instance of problem behavior or until 60 s elapsed (whichever occurred first). During the control interval, the hypothesized reinforcer was delivered such that the EO was absent for 60 s. Thus, if problem behavior occurred at any point during the test interval, reinforcement was delivered for the remainder of the trial. This aspect of the procedure preserved the reinforcement contingency necessary to identify behavior function yet minimized the number of reinforced problem behaviors per trial to one. Trial types were conducted in a mixed or random order until 20 trials per condition were completed. Data from each condition were summarized across trials to identify conditions in which the occurrence of problem behavior reliably occurred during the test (EO-present) interval and did not occur during the control (EO-absent) interval.

Wallace and Knights (2003) replicated the use of 2-min test-control trials to assess correspondence between what they referred to as brief (i.e., trial-based) FAs and extended (i.e., traditional) FAs. This study was conducted with three adults with developmental disabilities in a vocational training setting. A procedural distinction between this study and the previous studies by Sigafoos and Sagers (1995) and Sigafoos and Meikle (1996) was that test conditions were ordered according to a modified pairwise design. That is, rather than being rapidly alternated, a series of test-control trials for one condition (e.g., attention) was conducted, followed by a series of test-control trials for another condition (e.g., escape). In addition, all test intervals lasted 60 s such that multiple instances of problem behavior could occur prior to the control interval. The number of problem behaviors occurring during this interval, rather than the percentage of trials in which problem behavior occurred, was the dependent measure. Wallace and Knights (2003)

---

<sup>1</sup> An EO is an antecedent variable that increases the effectiveness of a reinforcer (e.g., deprivation of attention increases the effectiveness of attention as a reinforcer; Keller & Shoenfield, 1950; Michael, 1982).

identified correspondence between the trial-based FA and extended FA for two of three participants, with partial correspondence identified for the third participant. In addition, authors found the brief FAs required approximately 88% less time than the extended FAs (36 min and 5.2 hr, respectively).

LaRue et al. (2010) also evaluated correspondence between trial-based and traditional FAs for five participants with developmental disabilities in classrooms and vocational rooms within a clinical setting. During the trial-based FA, data were collected on the latency to problem behavior in addition to the occurrence of problem behavior during test and control intervals. Exact correspondence between outcomes was identified for four participants and partial correspondence was identified for the fifth participant. Latency data confirmed shorter latencies to problem behavior during test intervals than control intervals, but did not appear to improve the clarity of trial-based FA outcomes beyond the primary dependent measure (i.e., percentage of trials with problem behavior). Finally, the trial-based FA required an average of 85% less time than the traditional FA (32 min and 3.5 hr, respectively).

Bloom, Iwata, Fritz, Roscoe, and Carreau (2011) replicated and extended trial-based FA procedures with 10 students with developmental disabilities in classrooms of schools for children with developmental disabilities. Researchers modified the procedures used by Sigafoos and Sagers (1995) by extending the length of each test and control interval from 1 to 2 min and by modifying the sequence of test and control intervals to avoid the possibility of carryover.<sup>2</sup> Data on latency to problem behavior were also collected but were not reported because no systematic differences in latency were observed. Correspondence with extended FAs was identified for six

---

<sup>2</sup> Carryover is a type of sequence effect in which responding in one condition is influenced by responding in another condition within the same experimental phase (Hains & Baer, 1989). In the case of the trial-based FA, carryover may occur across the test and control intervals within the same experimental trial.

of 10 participants, with partial correspondence identified for a seventh participant. For two participants who showed non-correspondence, trial-based FAs were modified and repeated, at which point correspondence with extended FAs was observed.

Lambert, Bloom, and Irvin (2012) conducted trial-based FAs and subsequent functional communication training (FCT) interventions in an early childhood preschool center for three young children with developmental delay. The same procedures were used as Bloom et al. (2011), except only 10 trials of each condition were conducted (rather than 20) and trials were implemented by an itinerant special education teacher seeking her masters degree in special education (rather than full-time graduate students). Across participants, at least one function of problem behavior was identified via the trial-based FA and successfully treated using FCT.

Taken together, the small body of research on the trial-based FA appears promising, as this method allows the same evaluation of functional relations as the traditional FA, yet includes a modified set of procedures that may be more practical for natural classroom settings. Compared to traditional FAs, trial-based FA procedures may be incorporated more feasibly into typical classroom routines for several reasons. First, brief trials are embedded throughout the school day during times in which relevant antecedent variables are already in effect, thus producing minimal disruption to ongoing routines. Second, in comparison to extended sessions in which problem behavior may be repeatedly reinforced and increase in frequency, trial-based FAs involve an immediate transition to an abolishing operation (AO; antecedent variables that decrease the effectiveness of a reinforcer; Laraway, Snyckerski, Michael, & Poling, 2003) condition upon the first instance of problem behavior, thus minimizing risk to students and others involved in the assessment. Third, trial-by-trial data collection, in which problem behavior is recorded to have occurred or to have not occurred, is much less complex than the partial

interval data collection techniques common to traditional FAs. Finally, the use of numerous trials per condition may be more likely to provide reliable samples of behavior-environment patterns in the classroom setting while decreasing the total amount of time necessary to complete the analysis. Additional replications of trial-based FAs are needed, however, to determine the utility and impact of this emerging technology.

There are several opportunities for further evaluation and refinement of the trial-based FA. First, although previous studies on trial-based FAs have validated this methodology by evaluating correspondence with traditional FAs conducted in analogue settings, only two studies validated trial-based FA outcomes via a subsequent intervention analysis (i.e., Lambert et al., 2012; Sigafos & Meikle, 1996). Based on evidence suggesting outcomes of analogue FAs may not reflect behavior-environment relations existing in the natural environment, evaluating the effectiveness of interventions based on trial-based FA outcomes may be a preferred method of validation. Second, with the exception of one study (i.e., Lambert et al., 2012), trial-based FAs have only been conducted in separate school or clinical settings for individuals with disabilities. Though teaching staff did implement FA trials in four of the studies (Lambert et al., 2012; LaRue et al., 2010; Sigafos & Meikle, 1996; Sigafos & Sagers, 1995) these staff members were likely to have had specialized training in discrete trial teaching procedures. The teacher who participated in the study by Lambert et al. (2012), for example, was receiving graduate-level training in applied behavior analysis (ABA). No studies to date have included an implementation of the trial-based FA in public elementary school settings by classroom staff who may lack specialized training in ABA and/or discrete trial instructional techniques. Third, none of the studies have included any modifications of the standard attention, tangible, and escape conditions to increase the likelihood that the behavior-environment relations tested in the analysis represent

those naturally occurring in the classroom (i.e., ecological validity). Fourth, with one exception (i.e., Wallace & Knights, 2003), authors evaluating the trial-based FA have displayed the data as total percentages of test and control intervals with problem behavior per condition. This display does not reflect the method of ordering trials (i.e., rapid alternation on a trial-by-trial basis) that makes the trial-based FA an experimental assessment approach. It is unclear why the data have been summarized across conditions to preclude visual analysis of experimental data. Authors also have refrained from identifying the type of experimental design used, perhaps because the ordering of conditions was not apparent from their data displays. Because the trial-based FA is an experimental assessment method, it is important to display the data such that the experimental design is apparent. Finally, with one exception (i.e., LaRue et al., 2010), previous studies on trial-based FAs have included pre-determined numbers of trials per condition, which seems inconsistent with the logic of single-subject design.

The purpose of the current study was to replicate the trial-based FAs in elementary public school settings and determine whether results of similarly-structured intervention trials would validate outcomes of each trial-based FA. The following research questions were addressed:

- (1) Across participants, can functional relations be identified between student problem behavior and environmental stimuli via the trial-based FA methodology when teaching assistants (TAs) implement trials in the student's usual instructional setting?
- (2) For functional relations that are identified, can problem behavior be replaced by an appropriate communication response via the same trial-based methodology when TAs prompt and reinforce appropriate requests?

The study extends previous research on trial-based FAs in several ways. First, outcomes of the trial-based FAs were validated via results of subsequent intervention analyses rather than via results of analogue FAs. Second, TAs implemented all FA and intervention trials to determine the feasibility of the adults who typically support the students implementing these procedures with fidelity within their usual instructional settings. Third, descriptive data on student problem behavior and environmental antecedents and consequences (i.e., quantitative and qualitative direct observation, TA report) were collected prior to each trial-based FA to identify hypotheses and design experimental trials. Fourth, trial-based FA and intervention data were displayed such that the experimental design (i.e., method of ordering conditions) was transparent to facilitate visual analysis of experimental data. Finally, the total number of trials conducted was determined on an individual basis according to response differentiation between test and control intervals among conditions rather than according to a pre-determined number.



## CHAPTER II

### METHOD

#### **Participants and Settings**

Participants included four elementary school students who were receiving special education services and who engaged in high frequencies of problem behavior (as reported by special education teachers and confirmed via direct observation). Four TAs who supported the students during a time of day in which problem behavior occurred frequently also participated in the study. Participants attended two public elementary schools in the southeastern United States. After obtaining school district approval of the research project, the author met with a team of district-employed behavior specialists to identify students who were likely to meet inclusion criteria and benefit from participating in the project (no identifiable information was shared with the author until parent consent was obtained). To participate, students were required to (a) be in grades K-5; (b) have a primary label of autism, intellectual disability, or developmental delay per special education services and/or existing school records; and (c) engage in one or more topographies of problem behavior that met a high-frequency criterion and was likely influenced by social variables (evidenced by teacher report and researcher observation). The high-frequency criterion was identified as five or more occurrences of problem behavior within a 30-min interval in which these behaviors were reported to occur. Eight students were originally consented to participate; one student was excluded due to an extended absence (hospitalization) and three students were excluded because they did not meet the high-frequency problem behavior criterion. Demographic information for student and TA participants are displayed below in Table

1. All student information in Table 1 was collected from each student’s IEP, with information on current medication status confirmed with special education teachers. Descriptions of student problem behaviors, communication skills, and instructional settings are included below.

Table 1

*Student (top) and TA (bottom) participant demographics*

Participant	Gender	Age	Ethnicity	Classification	Diagnosis	IQ	Medication
Abhi	M	9	Asian	Autism	Autism	--	none
Sid	M	5	EA	DD	RS, Hydrocephalus	--	none
Davis	M	8	EA	DD	HMC	--	none
Gretchen	F	9	EA	ID; SI	DS; ADHD	40	FocalinXR

*Note.* M = male; F = female; EA = European-American; DD = developmental delay; ID = intellectual disability; SI = speech impairment; RS = Rhombencephalosynapsis; HMC = Hemimegiencephaly; DS = Down syndrome; ADHD = Attention Deficit Hyperactivity Disorder.

Participant	Gender	Age	Ethnicity	Education	Years Experience	Years w/ student
Mia (Abhi)	F	53	SA	4-year college	15	2
Lorraine (Sid)	F	57	EA	HS	1	1
Darcy (Davis)	F	49	EA	2 years college	8	3.5
Elaine (Gretchen)	F	57	EA	MA	10	1

*Note.* SA = Spanish-American; EA = European-American; HS = High school; MA = Master of Arts.

Abhi was a 9-year-old boy with Autism who engaged in physical aggression (hitting, kicking, pushing, pinching, scratching, biting), disruptions (grabbing and/or ripping materials, throwing, spitting), and elopement. His communication skills included prompted vocal approximations and occasional card exchanges from a Picture Exchange Communication System (PECS). He had no IQ score on file but qualified for the state’s alternate assessment (designed for students with significant cognitive disabilities). Abhi was receiving behavioral services at the

time the study began and his behavior plan included several preventative (visual schedule, first-then board, visual timer, token economy, social stories) and reactive (minimal attention, blocking, prompting picture exchanges) strategies. Abhi spent his morning instructional time with his TA, Mia, in a small instructional room that included a work table, chairs, a filing cabinet, and a computer. The majority of the initial descriptive assessments (observations) and all experimental trials were conducted in this room. Descriptive assessments also were conducted in a speech room during small-group speech instruction, in the library, and in the cafeteria.

Sid was a 5-year-old boy who was diagnosed with Rhombencephalosynapsis (RS; i.e., a rare malformation of the cerebellum) and hydrocephalus. He had strabismus and a history of seizures. His special education classification was developmental delay and he had no IQ score on file. Sid engaged in physical aggression (pinching, scratching, pushing), disruptions (pushing away materials), screaming, elopement, and flopping to the floor. He communicated via single word utterances and, occasionally, short phrases prompted by an adult. Sid was not receiving behavioral services at the time the study began. Initial descriptive assessments were conducted in his special education classroom and his general education kindergarten classroom. All experimental trials were conducted in his special education classroom.

Davis was an 8-year-old boy diagnosed with Hemimegalencephaly (HMC; i.e., enlargement of half the brain) and a visual impairment and had a history of seizures. He received special education services under the classification of developmental delay and had no IQ score on file. Davis was in a wheelchair for most of the school day but could walk with physical assistance from an adult. Davis engaged in physical aggression (scratching, grabbing, pinching, pulling hair, biting, kicking), disruptions (throwing, pushing materials away), elopement (spinning his chair), and screaming. He had limited communication skills, which included head

movements (nodding and shaking head to indicate *yes/no*) and reaching or grabbing. Davis was not receiving behavioral services at the time the study began. Based on special education teacher and TA reports that the highest rates of problem behavior occurred during a subset of his specials classes, initial descriptive assessments were conducted in the gym, art room, music room, library, and his special education classroom. Experimental trials were conducted in the gym or outside (PE class) and art room (Art class), with the exception of three trials that were conducted in the hallway outside each class.

Gretchen was a 9-year-old girl diagnosed with Down syndrome and Attention Deficit Hyperactivity Disorder (ADHD). She received special education services under the classification of intellectual disability (ID) and speech impairment (SI) and had a full-scale IQ of 40 (as measured by the Wechsler Intelligence Scale for Children; WISC). At the time of the study, she was taking 5 mg of dexamethylphenidate (FocalinXR; extended release) in the morning before school. Gretchen engaged in physical aggression (hitting, kicking, pushing), disruptions (taking shoes off, throwing, grabbing items), elopement, and flopping to the floor. She had limited communication skills, which included occasional vocal utterances (approximations of verbal prompts by adults), and reaching or grabbing for items. Gretchen was not receiving behavioral services at the time the study began. Initial descriptive assessments were conducted in her special education classroom and in the hallway during a recycling activity. Experimental trials were conducted in her special education classroom and a separate early childhood classroom when the state's annual standardized testing schedule interfered with the use of her usual classroom.

### **Response Definitions**

**Student behaviors.** The following operational definitions for student problem behaviors were identified. Physical aggression was defined as any instance of (a) forceful contact between

the participant's hands or feet and another person from a distance of 15 cm or greater (i.e., hitting, kicking, pushing, grabbing); (b) forceful contact between the participant's fingers and another person (i.e., scratching, pinching); (c) the participant's open mouth approaching and within 6 cm of another person's clothing or skin (i.e., bite attempts), or (d) another person's hair enclosed in the participant's closed fist (i.e., pulling hair). Disruptions were defined as any instance of (a) forceful contact between the participant's hands or feet and objects from a distance of 15 cm or greater, (b) throwing objects not intended to be thrown, (c) saliva crossing the plane of the participant's lips (i.e., spitting), (d) tearing or ripping instructional materials, or (e) removing one or both shoes. Screaming was defined as any vocalization that was at a higher pitch *and* volume than a conversational level. Flopping was defined as any contact between the participant's bottom or knees and the floor during instructional activities.

All four participants engaged in elopement, which was defined on an individual basis. For Abhi, elopement was defined as any instance of (a) any part of Abhi's body crossing the plane of an open doorway or (b) physical contact between Abhi's hand and the door knob/handle of a closed door. For Sid, elopement was defined as taking three or more steps away from his TA or from the instructional activity setting (e.g., work table). For Davis, elopement was defined as forceful contact between his right arm and the corresponding wheel on his chair during non-transitional times, which resulted in the chair making a spinning motion. For Gretchen, elopement was defined as any instance in which her bottom was not in contact with her seat during instructional activities.

Student replacement behaviors were identified on an individual basis in collaboration with the student's teachers. Abhi's and Gretchen's replacement behavior consisted of a picture exchange, which included picking up a laminated photograph and bringing it into contact with

the palm of the TA's hand. Sid's replacement behavior consisted of the vocal response, "No thank you" and was defined as the audible production of all three words in sequence. Davis's replacement behavior consisted of pressing a button that activated voice output. The button press was defined as contact between any part of Davis's body and the button such that the voice output was activated.

**Teacher behaviors.** Data were also collected on teacher behaviors to record the occurrence of antecedents and consequences of student problem behavior (descriptive assessments) and procedural fidelity of experimental trials (FA and intervention). During descriptive assessments, the vast majority of antecedent and consequent events reflected each TA's interactions with the student, but also included behaviors of other classroom staff (adults) who interacted with the student. Teacher behaviors included attention delivery, tangible restriction, tangible delivery, instruction delivery, and instruction removal. Attention delivery was defined as any vocal or physical interaction directed toward the student (e.g., statements or questions, pats on the back, prompts of any kind). The same definition of attention delivery was also coded for peers.

Tangible restriction was defined as the active removal or denial of a preferred item or activity from the student (e.g., taking a toy and placing it out of reach, saying to the student "All done with computer time," denying a request for more food during snack time). Tangible delivery was defined as providing access to or granting a request for a preferred item or activity (e.g., handing a toy directly to the student or placing it within reach, saying "Now it's time to play computer," granting a request for more food during snack time). In addition, tangible delivery was coded if the student successfully accessed a preferred item that was restricted in some way, even if it was not initiated by the teacher (e.g., student runs across the room, opens

cabinet, and grabs a handful of beads). Preferred items and/or activities coded as restricted or delivered were identified via TA report prior to the onset of descriptive assessments.

Instruction delivery was defined as any prompt (verbal, gestural, or physical) to complete an action or activity, whether academic or non-academic (e.g., “Write the letter A,” “Put your shoes back on,” physically guiding student to sit in seat). To be coded as an instruction, the prompt had to specify an action that could be complied with via some form of observable behavior. That is, announcements such as “It’s time for work now,” warnings such as “One more minute,” or directions such as “Listen to me” were not counted as instruction deliveries. Questions directed to students were coded as instruction delivery if the question facilitated the completion of the current activity. For example, during a coloring activity, the question “What color, blue or green?” would be coded as an instruction, whereas “What did you eat for dinner last night?” would not be coded as an instruction. Instruction removal was defined as an active removal of instructional demands via an announcement (e.g., “Ok, all done with math,” “Let’s take a break”) or the physical removal of instructional materials away from the student. During descriptive assessments only, instruction completion was defined as either (a) any attempt by the student to respond to or comply with the instruction or (b) any teacher-delivered physical prompt to complete the instruction.

## **Data Collection**

**Descriptive assessments.** Descriptive assessments consisted of a series of 10-min observation sessions conducted within a 1-hr period in which problem behavior was reported to occur. An average of three 10-min observations (range, 2-5) were conducted during the scheduled hour per daily observation. During descriptive assessments, a modified 20-s observe, 10-s record partial interval time sampling procedure was used to record student problem

behavior and teacher behavior (antecedent and consequent stimuli) as it naturally occurred. Observers wore ear buds that signaled observe and record intervals and data were collected via paper and pencil (see Appendix A).

Antecedent and consequent stimuli were coded based on the same sets of codes within three stimulus categories. Within the attention category, observers coded the occurrence of (a) teacher attention, (b) peer attention, or (c) diverted attention. If teacher or peer attention occurred at *any* time within a 20-s interval, teacher or peer attention was coded. If neither teacher nor peer attention occurred throughout the 20-s interval, diverted attention was coded. It was therefore possible for both teacher and peer attention to be coded within the same interval, but diverted attention was mutually exclusive of these codes. Within the tangible category, observers coded the occurrence of (a) tangible restriction, (b) tangible delivery, or (c) no tangible. If a preferred item or activity was restricted at *any* time within a 20-s interval, tangible restriction was coded. If a preferred item or activity was provided at *any* time within a 20-s interval, tangible delivery was coded. If neither tangible restriction nor tangible delivery occurred throughout the 20-s interval, no tangible was coded. It was therefore possible for both tangible restriction and tangible delivery to be coded within the same interval, but no tangible was mutually exclusive of these codes. Within the instruction category, observers coded the occurrence of (a) instruction delivery, (b) instruction removal, or (c) no instruction. If instructions were delivered at *any* time within a 20-s interval, instruction delivery was coded. If instructions were removed at *any* time within a 20-s interval, instruction removal was coded. If neither instruction delivery nor instruction removal occurred throughout the 20-s interval, no instruction was coded. It was therefore possible for both instruction delivery and instruction removal to be coded within the same interval, but no instruction was mutually exclusive of these codes.



Problem behavior was coded as present or absent during each 20-s interval. If an instance of problem behavior occurred at *any* time within the 20-s interval, problem behavior was coded as present. If no problem behavior occurred throughout the 20-s interval, problem behavior was coded as absent. For intervals in which problem behavior was absent, only antecedent stimulus categories were coded. For intervals in which problem behavior was present, both antecedent and consequent stimulus categories were coded.

**Experimental trials.** During all FA trials, observers recorded the presence or absence of problem behavior during the test interval and the control interval. The latency to problem behavior (seconds) in the test interval was also recorded, with the maximum latency being 60 s. During all intervention trials, observers recorded the presence or absence of (a) problem behavior, (b) prompted replacement behavior, and (c) unprompted replacement behavior during the test interval and the control interval. The latency to replacement behavior (seconds) in the test interval was also recorded, with the maximum latency being 60 s. These data were recorded via paper and pencil data collection sheets (see Appendices B and C) and stopwatches were used to record latencies to problem and/or replacement behaviors.

**Procedural fidelity.** Across all FA and intervention trials, data collectors recorded whether the programmed antecedents occurred during the test interval and whether the programmed consequences occurred during the control interval using the same antecedent and consequent coding categories used for descriptive assessments. In addition, data collectors indicated whether TAs transitioned from the test to control interval within 5 s of the first instance of problem behavior (FA trials) or replacement behavior (intervention trials). For conditions involving a transition in activities (e.g., from instruction at work table to playing on computer), the TA was required to *announce* the transition (e.g., “Ok, let’s go back to the computer”) within

5 s of the target behavior, as it was not feasible to complete the transition within 5 s. On the data sheets (see Appendices B and C), the latency to problem or replacement behavior was circled when the TA initiated the control interval within 5 s of this latency. To be counted as a trial with fidelity, all relevant stimulus categories across both test and control intervals had to match the programmed stimuli and the transition from the test to control interval had to have occurred within 5 s of (a) problem or replacement behavior or (b) 60 s elapsing without problem or replacement behavior. If any stimulus categories did not match those programmed in either test or control interval, or if the 5-s latency criterion was not met, the trial was coded as having procedural errors. Percentages of correctly-implemented trials were calculated by participant and condition and are displayed in Table 2. Across participants, a mean of 88% (range, 80-100%) of assessment trials were conducted with fidelity, and a mean of 91% (range, 84-100%) of intervention trials were conducted with fidelity.

Table 2

*Procedural fidelity and inter-observer agreement on fidelity of experimental trials*

Participants	% Assessment trials with fidelity	% Assessment trials with IOA (% Agree)	% Treatment trials with fidelity	% Treatment trials with IOA (% Agree)
Mia/Abhi				
Escape+Tangible	100	43 (67)	100	47 (100)
Attention	100	60 (100)	n/a	n/a
Lorraine/Sid				
Escape novel task	88	50 (100)	87	35 (100)
Transition avoid	90	30 (100)	90	45 (100)
Escape+Tangible	91	36 (100)	100	50 (100)
Darcy/Davis				
Attention	89	78 (86)	92	69 (100)
Escape	89	71 (60)	67	33 (100)
Physical prompts	57	78 (86)	n/a	n/a
Elaine/Gretchen				
Attention	77	62 (75)	91	27 (100)
Tangible	92	50 (100)	92	38 (100)
Escape	80	40 (100)	n/a	n/a

**Observer Training and Reliability**

The author was the primary data collector across participants and trained three graduate research assistants to collect inter-observer agreement (IOA) data during a subset of descriptive assessments and experimental trials. Across study phases, each pair of research participants (student and their TA) was observed by only one graduate research assistant in addition to the author. Graduate research assistants were first trained to collect data during descriptive assessments. A 1-hr training meeting was conducted during which the author reviewed response definitions (including examples and non-examples) and data collection procedures as displayed in an observation coding manual. In preparation for collecting data in live classrooms, data collectors practiced collecting descriptive assessment data in a preschool setting including students with and without disabilities. Observations were done in an observation booth such that

agreements and disagreements could be discussed aloud between training sessions. The criterion for reliability in this live setting was three consecutive 10-min sessions with agreement on at least 85% of intervals. (To be counted as an interval with agreement, all antecedent, target behavior, and consequent codes were required to match within that interval.) Based on a general lack of problem behavior occurring during these practice sessions (thus minimizing opportunities to practice coding consequences), an additional training component was required. The author selected a sample video from a previous research project on problem behavior, in which more than half of the intervals included problem behavior and would therefore require coding of both antecedents and consequences. The author created a master code and each observer was required to reach 85% agreement or higher with the master before beginning data collection in the field. (Multiple attempts on this video were allowed.) The final training criterion for descriptive assessments included reaching 85% agreement with the author during three consecutive 10-min sessions in the research setting with the research participants. After this criterion was met, each data collector was considered reliable and subsequent data collection during descriptive assessments counted towards reliability.

Agreement on descriptive assessment data was calculated following each 10-min observation during which a secondary observer was present. Following each 10-min observation, point-by-point agreement was calculated on an interval-by-interval basis. That is, each 20-s interval was scored as either an agreement (i.e., all antecedent stimulus categories, student behavior, and consequent stimulus categories were coded identically) or a disagreement (i.e., one or more antecedent stimulus categories, student responses, or consequent stimulus categories were coded differently). The number of intervals scored as agreements was divided by the total number of intervals and multiplied by 100%. Checking interval-by-interval agreement following

each observation allowed an opportunity to discuss and/or reach consensus on any disagreements. IOA data were collected on a minimum of 30% of descriptive assessments for each participant ( $M = 37\%$ , range 30-45%). The mean percentage of agreement across participants was 89% (range, 87-90%). Percentages of (a) observation sessions and (b) total intervals with IOA and agreement percentages are displayed below by participant in Table 3.

Table 3

*Inter-observer agreement on descriptive assessment data*

Participant	% 10-min sessions	Mean agreement (Range)	% Total intervals	% Total agreement
Abhi	35	90 (80-100)	32	88
Sid	37	89 (75-95)	37	89
Davis	45	88 (75-100)	43	87
Gretchen	30	87 (77-95)	30	87

To train observers to collect data on student and TA behavior during experimental trials, the author led a 1-hr training meeting to describe the experimental trial procedures. The graduate research assistants then role-played example experimental trials, which were video-recorded. In addition to familiarizing the observers with the trial procedures, these videos were subsequently used to practice collecting both primary outcome data (student behavior) and procedural fidelity data (TA behavior). The author created a master code for each recorded trial, and observers practiced coding the trials until two consecutive trials of each type were coded with 100% agreement with the master. A 3-s window of agreement was used for the latency measure. Once this criterion was met, observers practiced coding experimental trials in the research setting with the research participants until 100% agreement with the primary observer (author) was reached

on two consecutive trials of each type. All subsequent data collected during experimental trials by secondary observers counted towards IOA on student behavior and procedural fidelity.

Throughout data collection on experimental trials, the primary and secondary observers compared coding sheets following each trial to allow opportunities to discuss and/or reach consensus on any disagreements. IOA was collected on a mean of 55% of assessment trials and 45% of intervention trials across participants (ranges, 38-77% and 36-58%, respectively). The mean percentage of agreement on student behavior (primary dependent measure) was 97% (range, 89-100%) during assessment trials and 100% during intervention trials. The mean percentage of agreement on TA behavior (procedural fidelity measure) across participants was 88% (range, 79-100%) during assessment trials and 100% during intervention trials. Percentages of IOA on procedural fidelity are displayed in Table 2 by participant, assessment trial condition, and intervention trial condition.

### **Experimental Design**

A modification of the multi-element design (Ulman & Sulzer-Azaroff, 1975) was used to evaluate the effects of test-control trials on problem behavior during the trial-based FA. The multi-element design involves the rapid alternation of test conditions and is the standard for identifying behavior function via FA methodology (Gast, 2010; Kennedy, 2005). The multi-element design is appropriate for assessing behavior function because it allows a relatively rapid comparison of multiple test conditions. Because problem behavior is typically considered readily reversible (i.e., influenced by immediate contextual conditions to a greater extent than prior learning histories), it is expected to be sensitive to rapid alternations of test conditions. Multiple trials were conducted during the same 1-hr period each day, with a minimum of 5 min between trials.

The experimental design used to evaluate the effects of intervention trials was determined based on the results of the trial-based FA. In the case of a single function being identified (Abhi), an A-B-A-B reversal design (Birnbrauer, Peterson, & Solnick, 1974; Campbell & Stanley, 1966) was used with the FA trials as the initial baseline (A) phase. If two functions were identified (Davis and Gretchen), a combination multiple baseline and reversal design was used such that the introduction of intervention trials was staggered across two conditions with a brief reversal on one of the tiers to allow for additional replications. If three functions were identified (Sid), a multiple baseline design was used such that the intervention trials were staggered across the three conditions. When the introduction of intervention trials was staggered across conditions, intermittent probe trials were conducted in the remaining (untreated) conditions.

### **Study Procedures**

**TA interview.** The author conducted an initial interview with each student's TA (and in some cases, the special education teacher also participated) using the Functional Analysis Interview (FAI; O'Neill et al., 1997). This interview provided an opportunity to collect information on topographies of problem behaviors, times of day and/or routines or activities associated with the highest and lowest frequencies of problem behavior, communication skills, preferred items/activities, and potential environmental influences on behavior patterns at school. Interviews lasted between 30 min and 1 hr.

**Descriptive assessments.** Based on the information collected from the interview and the availability of classroom staff, a 1-hr period was identified for each student during which all observations and experimental trials would be conducted. Descriptive assessments consisted of a series of 10-min observation sessions within this 1-hr period that were carried out on a daily basis (subject to any student absences and/or scheduling conflicts). A minimum of five daily

descriptive assessments were completed for each participant, with each descriptive assessment consisting of 20-50 min of direct observation. Teaching staff were instructed to carry on usual instructional routines and to respond to student problem behavior as they would as if the observers were not there. Observers were positioned in close enough proximity to see and hear the student and teacher behaviors being coded but attempted to position themselves in a way that was minimally disruptive to the research participants as well as other teaching staff and students who were present. Data from descriptive assessments were used to verify the information collected from the TA interview and to inform experimental trial procedures such that programmed antecedent and consequent events corresponded with those naturally occurring in the student's classroom.

**TA training on assessment procedures.** Prior to initiation of the trial-based FA, a 1-hr training meeting was held with the TA who supported the student. During this meeting, the author (a) reviewed a summary of the data collected during the descriptive assessments, (b) identified hypotheses regarding the function(s) of the student's problem behavior based on the descriptive assessments and information collected via interview, (c) proposed how these hypotheses may be tested using the test-control trial format, (d) sought TA feedback on whether the experimental procedures corresponded with events the student typically encounters, (e) made adjustments to trial procedures based on TA feedback, and (f) discussed how the trials would be integrated into the student's instructional routine. The author then described the procedures for each trial type and provided a visual aid depicting the test-control trial format. If the procedures were unclear to a TA, the researcher offered to role-play each trial type prior to implementing the trials with the student. The role-playing activity was done with only one of the four TAs (Mia). During implementation of assessment trials, the researcher provided reminders prior to trials on



an as-needed basis. In addition, during all assessment trials, the researcher provided the verbal prompt “switch” to transition from the test to control interval.

**Trial-based FA.** The trial-based FA was initiated following the descriptive assessments and training meeting. A minimum of two trial types were included in each participant’s FA and were selected based on information provided from the TA and direct observation data from the descriptive assessments. The TA who supported the student during the specified 1-hr period implemented all assessment trials. Multiple trials (between 2 and 6) were conducted during this 1-hr window on a daily basis (based on student and TA availability) with at least 5 min between trials. Test conditions included variations of attention, tangible, escape, and combination test-control trials. The first minute of each trial consisted of a test interval in which a hypothesized EO was in effect. The second minute of each trial consisted of a control interval in which hypothesized reinforcer was delivered. If problem behavior occurred at any point during the first minute, the test interval immediately ended and the control interval began. Each trial therefore lasted between 1 and 2 min. Although similar response-reinforcer relations were tested across participants, the trial procedures differed among participants due to experimenter efforts to make the trials correspond with the each student’s usual routines. Descriptions of each assessment trial type per participant are included below.

*Abhi.* Abhi’s usual routine consisted of alternating periods of instruction and computer breaks. The assessment trials included in Abhi’s trial-based FA included (a) attention and (b) escape + tangible trials, both of which were embedded within his regularly scheduled computer breaks. During each type of assessment trial, Abhi’s PECS book was placed on the computer table within his view and reach.

*Attention.* The attention trial was designed to test for positive reinforcement in the form of TA attention. During attention trials, Mia diverted her attention (i.e., stood several feet away from him, did not talk to him or approach him) for the first minute of the trial. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurred first), the researcher provided the verbal prompt, “switch,” and Mia delivered attention for the second minute of the trial (e.g., pat on back, “Oh that looks fun!” [pointing at screen] “Who’s that?”). The form of attention delivered was consistent with Mia’s usual interactions with Abhi when he was engaged on the computer.

*Escape + tangible.* The escape + tangible condition was designed to test for a combination of positive and negative reinforcement in the form of access to computer breaks (escape from instruction and access to computer). This combination trial was designed based on Mia’s report that Abhi was never granted breaks from instruction without access to a preferred activity (which was also supported by researcher observation). During escape + tangible trials, Mia told Abhi computer time was all done, turned off the computer monitor, and instructed him to come to the work table for the first minute of the trial. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurred first), the researcher provided the verbal prompt, *switch*, and Mia allowed Abhi to return to the computer for the second minute of the trial.

*Sid.* Sid’s usual routine consisted of alternating periods of instruction at a table, teacher-facilitated computer time, and breaks during which Sid was free to move about the classroom. The assessment trials included in Sid’s trial-based FA included three types of escape trials: (a) escape from novel instructions, (b) avoidance of transition, and (c) escape + tangible.

*Escape from novel instruction.* The escape from novel instruction trial was designed to test for negative reinforcement in the form of escaping or avoiding unfamiliar instructional materials. Descriptive assessment data revealed variable percentages of instruction completion, and this trial was designed to test Lorraine's report that problem behavior often occurred with novel or unfamiliar instructional activities were introduced. This trial was embedded into an instructional activity that consisted of the TA presenting cards with shapes, colors, objects, and printed words to Sid and asking him to identify them. During the first minute of the trial, Lorraine introduced an unfamiliar set of cards within this instructional activity. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurred first), the researcher provided the verbal prompt, "switch," and Lorraine withdrew the unfamiliar cards and reintroduced the familiar cards for the second minute of the trial.

*Avoidance of transition.* The avoidance of transition trial was designed to test for negative reinforcement in the form of escaping/avoiding transitions from preferred to non-preferred activities. This trial was embedded into Sid's computer time. During the first minute of the trial, Lorraine initiated a transition by telling Sid computer time was over and prompting him to go to the work table. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurred first), the researcher provided the verbal prompt, "switch," and Lorraine terminated the transition and allowed S to return to the computer activity for the second minute of the trial.

*Escape + tangible.* The escape + tangible trial was designed to test for a combination of positive and negative reinforcement in the form of breaks from instruction with access to a preferred item. This trial began when Sid had access to a high-preferred item (i.e., stuffed penguin). During the first minute, Lorraine took the penguin away from Sid, placed it out of

reach but within view (on the opposite side of the work table) and introduced an instructional activity. Initially, the instruction activity involved a color-sorting task, but was later modified to a book reading activity. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurred first), the researcher provided the verbal prompt, “switch,” and Lorraine removed the instructional activity, provided the penguin, and allowed him to leave the work table area for the second minute of the trial.

**Davis.** Davis’s trials were conducted in two different specials classes (PE and Art). In PE, his usual routine consisted of alternating periods of 1:1 instruction on a modified/adapted PE-related activity and walking breaks, during which he walked around the gym with assistance from Darcy. In Art, his usual routine consisted of 1:1 instruction on a modified/adapted art-related activity with occasional walking breaks around the classroom (if other students were free to walk around) or out in the hall. The assessment trials in Davis’s trial-based FA included (a) attention (b) avoidance of physical prompts, and (c) escape from 1:1 instruction trials, all of which were conducted across both specials classes.

**Attention.** The attention trial was designed to test for positive reinforcement in the form of TA attention. Prior to beginning each attention trial, Darcy told Davis she needed to talk to someone else for a minute. During attention trials, Darcy diverted her attention (oriented away from him and spoke to another adult) for the first minute of the trial. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurred first), the researcher provided the verbal prompt, “switch,” and Darcy delivered high-quality attention for the second minute of the trial (e.g., knelt down to eye-level, spoke in an animated tone, asked silly questions, touched the writing on his shirt). The attention was delivered throughout the full 60-s control interval, based on observations of problem behavior occurring soon after attention was withdrawn.

*Avoidance of physical prompts.* The avoidance of physical prompts trial was designed to test for negative reinforcement in the form of avoiding physical prompts. The avoidance trials were embedded into the 1:1 specials-related instructional activity. During the first minute of the trial, Darcy provided instruction using physical (hand-over-hand) prompting. In PE class, this activity involved putting balls into buckets. In Art class, this activity involved drawing/coloring and creating a dream catcher. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurs first), the researcher provided the verbal prompt, “switch,” and Darcy removed the physical prompts but continued instruction using verbal and/or gestural prompting only.

*Escape from 1:1 instruction.* The escape from 1:1 instruction trial was designed to test for negative reinforcement in the form of walking breaks. The escape trials were embedded into the same instructional activities as described above. During the first minute of the trial, Darcy provided 1:1 instruction in the specials-related activity. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurs first), the researcher provided the verbal prompt, “switch,” and Darcy announced that it was time for a break, and she assisted Davis out of his chair and walked either around the gym (PE) or down the hall (Art) for the second minute of the trial. This trial therefore tested the effects of escaping both the instructional activity and his wheel chair.

***Gretchen.*** Gretchen’s usual routine consisted of alternating periods of 1:1 instruction at a table, and breaks with preferred items or activities (which were described as sensory-related activities by her teachers). The assessment trials included in Gretchen’s trial-based FA included: (a) attention, (b) tangible, and (c) escape.

*Attention.* The attention trial was designed to test for positive reinforcement in the form of physical attention. The attention trials were embedded into Gretchen's breaks from instruction, during which she had access to a preferred item (beads). Prior to beginning the trial, Elaine told Gretchen she had some work to do. During the first minute of the trial, the teacher turned away from Gretchen and engaged in paperwork. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurred first), the researcher provided the verbal prompt, "switch," and Elaine delivered attention during the second minute of the trial. The form of attention Elaine provided was initially verbal attention, but was subsequently changed to physical attention in the form of putting lotion on Gretchen's hands (see results section).

*Tangible.* The tangible trial was designed to test for positive reinforcement in the form of accessing a preferred item (beads). The tangible trials were embedded into Gretchen's breaks from instruction, during which she had access to her beads. During the first minute of the trial, Elaine said, "My turn," and attempted to take the beads away from Gretchen. Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurred first), the researcher provided the verbal prompt, "switch," and Elaine said, "Your turn," and returned the beads to Gretchen for the second minute of the trial.

*Escape.* The escape trial was designed to test for negative reinforcement in the form of breaks from instruction. The escape trials were embedded into Gretchen's instruction time. During the first minute of the trial, Elaine prompted Gretchen to engage in a math activity (i.e., matching numbers on a number line). Upon (a) the occurrence of problem behavior or (b) the passage of 60 s (whichever occurred first), the researcher provided the verbal prompt, "switch," and Elaine said, "Let's take a break," and physically removed the instructional materials for the second minute of the trial.

**TA training on intervention procedures.** After the trial-based FA was completed, a second 1-hr training meeting was held with the TA to prepare for the implementation of intervention trials. During this meeting, the researcher (a) summarized the results of the trial-based FA, (b) solicited TA feedback on the topography of replacement behavior most appropriate for the student, and (c) described how the replacement behavior(s) would be prompted and reinforced within the same test-control trial format. During implementation of intervention trials, the researcher provided reminders prior to trials on an as-needed basis in addition to providing instructions for fading prompts across trials. During intervention trials, the researcher no longer provided the verbal prompt “switch” following the replacement behavior, as the replacement behavior itself was a sufficient prompt for the TAs. The researcher did, however, provide any prompts that were time-based (e.g., when 60 s elapsed without the replacement behavior, when a 30-s delay was inserted prior to any prompt to allow for independent replacement behaviors).

**Intervention.** Following the identification of behavior function via trial-based FA, intervention trials were implemented to evaluate whether an appropriate form of communication would replace problem behavior when the appropriate request produced the reinforcing consequence and the problem behavior did not produce the reinforcing consequence. Intervention trials were conducted using the same trial format as in the assessment trials, except that the transition from the test interval to the control interval took place following the first instance of *replacement* behavior or the passage of 60 s. That is, appropriate replacement behaviors were reinforced and problem behaviors were put on extinction during intervention trials. In addition, participants were initially prompted to engage in the replacement behavior at the beginning of the test interval of each intervention trial, and prompts were faded to allow for

unprompted/independent occurrences of replacement behaviors. Although similar contingency reversals were implemented across participants, the replacement behaviors and methods of prompting were identified on an individual basis. Descriptions of each intervention trial type per participant are included below.

*Abhi.* The replacement behavior selected for Abhi was a picture exchange, which was selected based on his prior exposure to PECS and a lack of reliable vocal responses at the time of intervention. Based on the results of Abhi's trial-based FA, the intervention procedures were introduced in the escape + tangible condition.

*Escape + tangible intervention.* The escape + tangible intervention trial was conducted using the same procedures as the corresponding assessment trial, except that a picture of the computer was placed on the front cover of Abhi's PECS book, which was located on the computer table within Abhi's view and reach. During the first two intervention trials, the researcher physically prompted Abhi to pick up the picture of the computer and place it in Mia's hand immediately after she initiated the trial by announcing that computer was all done, turning off the monitor, and instructing him to go to the work table. Immediately following the picture exchange, Mia turned the computer screen back on and said, "Ok, you can have more computer time," allowing him to remain on the computer for the second minute of the trial. During subsequent trials, the prompt was faded to a verbal and gestural prompt (i.e., holding out hand and saying "If you want more computer time, give me the picture"), followed by the gestural prompt only (i.e., holding out hand), and finally initiating the test interval with no prompts to allow for independent picture exchanges.

*Sid.* The replacement behavior selected for Sid was the vocal response, "No thank you." This form of communication was selected based on Sid's reliable vocal requests for preferred



items and activities (e.g., “Pablo” to request the stuffed penguin, “Hi ho” to request a song). Sid was not, however, observed to use words to refuse non-preferred activities or transitions before engaging in problem behavior. Based on the results of Sid’s trial-based FA, the intervention procedures were introduced in all three escape-related conditions.

*Escape from novel instructions intervention.* The escape from novel instructions intervention trial was conducted using the same procedures as the corresponding assessment trial, except that following the initiation of the test interval (i.e., introduction of unfamiliar cards), Lorraine delivered the verbal prompt, “If you don’t want to, say: No thank you.” Contingent on the “No thank you” response, Lorraine immediately removed the unfamiliar cards and reintroduced the familiar cards for the second minute of the trial. If Sid did not say “No thank you,” Lorraine followed through with instruction using the unfamiliar cards, with additional verbal prompts to say “No thank you” interspersed throughout the first minute of the trial. As was the case during assessment trials, the control interval was initiated after 60 s if the target behavior did not occur. If problem behavior began to co-occur with the replacement behavior, Lorraine was instructed to reinforce replacement behaviors only when they occurred in the absence of problem behavior.

*Transition avoidance intervention.* The transition avoidance intervention trial was conducted using the same procedures as the corresponding assessment trial, except that following the initiation of the test interval (i.e., initiating transition to work table), Lorraine delivered the verbal prompt, “If you don’t want to, say: No thank you.” Contingent on the “No thank you” response, Lorraine immediately terminated the transition and allowed Sid to stay at the computer. If Sid did not say “No thank you,” Lorraine followed through with the transition to the work table, with additional verbal prompts to say “No thank you” interspersed throughout the

first minute of the trial. As was the case during assessment trials, the control interval was initiated after 60 s if the target behavior did not occur. If problem behavior began to co-occur with the replacement behavior, Lorraine was instructed to reinforce replacement behaviors only when they occurred in the absence of problem behavior.

*Escape + tangible intervention.* The escape + tangible intervention trial was conducted using the same procedures as the corresponding assessment trial, except that following the initiation of the test interval (i.e., restriction of penguin and introduction of book activity), Lorraine delivered the verbal prompt, “If you don’t want to, say: No thank you.” Contingent on the “No thank you” response, Lorraine immediately removed the book activity and delivered the penguin. If Sid did not say “No thank you,” Lorraine followed through with the book activity, with additional verbal prompts to say “No thank you” interspersed throughout the first minute of the trial. As was the case during assessment trials, the control interval was initiated after 60 s if the target behavior did not occur. If problem behavior began to co-occur with the replacement behavior, Lorraine was instructed to reinforce replacement behaviors only when they occurred in the absence of problem behavior.

*Davis.* The replacement behavior selected for Davis was a button press, which activated recorded speech output. This form of communication was selected based on a lack of reliable vocal responses as well as teacher report that picture exchanges may not be appropriate given his visual impairment. Two buttons were prepared, each producing a distinct spoken request (“Talk to me please” and “Let’s take a walk”). Each button was a different color and was mounted on a box displaying the corresponding message in text. Based on the results of Davis’s trial-based FA, the intervention procedures were introduced in the attention and escape from 1:1 instruction conditions.

*Attention intervention.* The attention intervention trials were conducted using similar procedures as the corresponding assessment trial, except that a button was placed in front of Davis when Darcy told him she needed to talk with someone else. During the first four attention trials, a verbal prompt was also given (“If you want to talk to me, just press the button.”) During subsequent trials, a verbal prompt was only given if 30 s elapsed without a button press. Contingent on the button press (which activated the “Talk to me please” speech output), Darcy immediately delivered high-quality attention for the second minute of the trial. As was the case during assessment trials, the control interval was initiated after 60 s if the replacement behavior did not occur.

*Escape from 1:1 instruction intervention.* The escape from 1:1 instruction intervention trials were conducted using similar procedures as the corresponding assessment trial, except that a button was placed in front of Davis when Darcy initiated the 1:1 instruction to begin the trial. During the first escape trial, a verbal prompt was also given (“If you want to take a walk, just press the button.”) During subsequent trials, a verbal prompt was only given if 30 s elapsed without a button press. Contingent on the button press (which activated the “Let’s take a walk” speech output), Darcy initiated the walking break for the second minute of the trial. As was the case during assessment trials, the control interval was initiated after 60 s if the replacement behavior did not occur.

***Gretchen.*** The replacement behavior selected for Gretchen was a picture exchange. This form of communication was selected based on a lack of observed reliable vocal responses and previous experience with similar picture cards used in instructional contexts. Based on the results of Gretchen’s trial-based FA, the intervention procedures were introduced in the tangible and physical attention conditions. When intervention trials began with Gretchen, we modified the

way they were embedded into her instructional routine. Rather than embedding the tangible and attention trials within her breaks from instruction, we initiated the trials at the end of each interval of instruction such that she was able to practice requesting her preferred break activities on each break.

*Tangible intervention.* The tangible intervention trials were conducted using similar procedures as the corresponding assessment trial, except that a laminated photograph was placed on the table in front of Gretchen while Elaine restricted the beads (held them in her hands). During the first two tangible trials, the researcher physically prompted Gretchen to pick up the picture of the beads and place it in Elaine's open hand at the start of the trial. During subsequent trials, the prompt was faded to a verbal and gestural prompt (i.e., holding out hand and saying "If you want beads, give me the picture"), followed by the gestural prompt only (i.e., open hand), and finally inserting a 30-s delay to the gestural prompt to allow for independent picture exchanges. Contingent on the picture exchange, Elaine immediately provided verbal praise ("You want beads, good asking!") and delivered the beads for the second minute of the trial. As was the case during assessment trials, the control interval was initiated after 60 s if the replacement behavior did not occur.

*Attention intervention.* The attention intervention trials were conducted using similar procedures as the corresponding assessment trial, except that a laminated photograph was placed on the table in front of Gretchen while Elaine restricted her attention and the lotion (held the bottle in her hands and faced slightly away from Gretchen). Based on an independent picture exchange prior to the planned prompt in the first attention trial, a gestural prompt was only delivered in subsequent trials following 30 s elapsing without a picture exchange. Contingent on the picture exchange, Elaine immediately delivered attention in the form of verbal praise ("Good

asking!”) and initiation of the lotion activity for the second minute of the trial. As was the case during assessment trials, the control interval was initiated after 60 s if the replacement behavior did not occur.

**TA questionnaire.** Following the completion of data collection, each TA was asked to complete a brief social validity questionnaire. The questionnaire consisted of two parts. The first part included 10 items on the acceptability of assessment and intervention procedures and was a modification of the Intervention Rating Profile-15 (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985). The second part consisted of two questions for each type of trial the TA implemented. The first question addressed the similarity of the trial to events the student typically encounters in the same setting and the second question addressed the difficulty of trial implementation. TAs responded to each item on the questionnaire by selecting a number on a likert-type scale (1-6). TAs were given the option of completing the questionnaire with or without the researcher being present and were encouraged to ask any questions if any of the items were unclear. All TAs completed the questionnaires individually and the researcher collected them on a follow-up visit.

## **Data Analysis**

**Descriptive assessments.** Data from descriptive assessments were entered into SPSS and summarized across observations for each participant. The total percentage of intervals with each coded event (i.e., estimated base rate, or simple probability) was calculated by dividing the total number of intervals in which each event was coded by the total number of observed intervals (across all descriptive assessments). These percentages represented estimated base rates of various student and teacher behaviors (e.g., student problem behavior, teacher attention, instruction delivery) across observations.

Percentages of the same coded events were calculated for intervals during which problem behavior occurred to determine whether any environmental events were more or less likely to occur immediately surrounding problem behavior (as compared to overall base rates). These percentages were calculated separately for antecedent and consequence codes. For example, the percentage of intervals with problem behavior with teacher attention coded as an antecedent was calculated by dividing the number of intervals in which both events were coded divided by the total number of intervals with problem behavior and multiplied by 100 to yield a percentage. Similarly, the percentage of intervals with problem behavior with teacher attention coded as a consequence was calculated by dividing the number of intervals in which both events were coded by the total number of intervals with problem behavior and multiplied by 100 to yield a percentage.

Bar graphs depicting these percentages were interpreted according to (a) whether a potential EO (diverted attention, instruction delivery, tangible restriction) was coded more often immediately *before* problem behavior than overall, and (b) whether a potential reinforcer (attention delivery, instruction removal, tangible delivery) was coded more often immediately *after* problem behavior than overall. For example, if the likelihood of instruction delivery was higher immediately before problem behavior than across all intervals, instruction delivery would be identified as a hypothesized EO, which would support the inclusion of an escape trial. Or, if the likelihood of attention was higher immediately after problem behavior than across all intervals, attention would be identified as a hypothesized reinforcer, which would support the inclusion of an attention trial. During the TA training meeting on assessment procedures, these patterns of percentages were discussed in terms of how each student's problem behavior may be affecting his or her environment. In addition to these quantitative data, however, feedback from

each TA as well as anecdotal qualitative observations by the researchers (e.g., student affect and/or precursor behaviors in the presence or absence of specific environmental stimuli) were considered in the selection and design of experimental trials.

**Experimental trials.** Data on the occurrence of problem behavior during test and control intervals of FA trials were graphed on a daily basis and interpreted via visual analysis.

Assessment trials continued until response differentiation (i.e., occurrence of problem behavior during test interval and non-occurrence of problem behavior during control interval) in a minimum of one condition was observed. Data on the occurrence of problem behavior, prompted replacement behavior, and independent (unprompted) replacement behavior during intervention trials were graphed on a daily basis and interpreted via visual analysis. Intervention procedures were introduced following stable response differentiation during the assessment trials of each condition. Subsequent introduction of intervention procedures in other conditions were made following (a) the stable occurrence of replacement behavior and nonoccurrence of problem behavior during test intervals of the intervention trials, (b) the stable non-occurrence of replacement behavior during control intervals of the intervention trials, and (c) continued response differentiation during intermittent probes of the assessment (baseline) trials. When applicable, brief withdrawals of intervention procedures were implemented following the stable occurrence of replacement behavior and nonoccurrence of problem behavior during test intervals and the stable non-occurrence of replacement or problem behavior during control intervals of intervention trials. In addition, this data pattern was interpreted as validating the results of the corresponding FA trial.

**Trial durations.** The total time required to complete experimental assessments and interventions was calculated by participant in minutes, days, and weeks. The duration of each

trial was calculated by adding the latency to target behavior to the 60-s control interval, then adding 5 s to account for the transition between intervals. The total number of seconds was summed across trials and divided by 60 to yield total minutes spent conducting assessment and/or intervention trials.



## CHAPTER III

### RESULTS

#### **Descriptive Assessments**

**Abhi.** Results of Abhi's descriptive assessments are displayed in Figure 1. Across observations, teacher attention was delivered in 74% of intervals, instructions were delivered in 43% of intervals, and preferred items were restricted and delivered in 10% and 4% of intervals, respectively. Based on the increased likelihood of attention immediately after problem behavior (compared to the estimated base rate of attention), an attention trial was included in Abhi's trial-based FA. Compared to the estimated base rate of instruction delivery, the likelihood of instruction delivery was higher immediately before problem behavior and lower immediately after problem behavior. In addition, the likelihood of completed instructions was lower immediately surrounding problem behavior than across all intervals. That is, these descriptive data suggest problem behavior was more likely to occur in the context of instruction delivery, and instructions were less likely to be completed or followed through with when problem behavior occurred. Escape was therefore identified as a hypothesized function. Although the base rate of tangible deliveries was very low, the likelihood of tangible restriction was slightly higher immediately before problem behavior compared to the estimated base rate of tangible restriction. Tangible deliveries were rarely captured via the data collection system, however, which was likely due to the computer remaining on for extended intervals (even when Abhi was expected to complete work). Because teacher report and anecdotal observation also supported the tangible hypothesis, access to tangibles was identified as a hypothesized reinforcer for problem behavior.

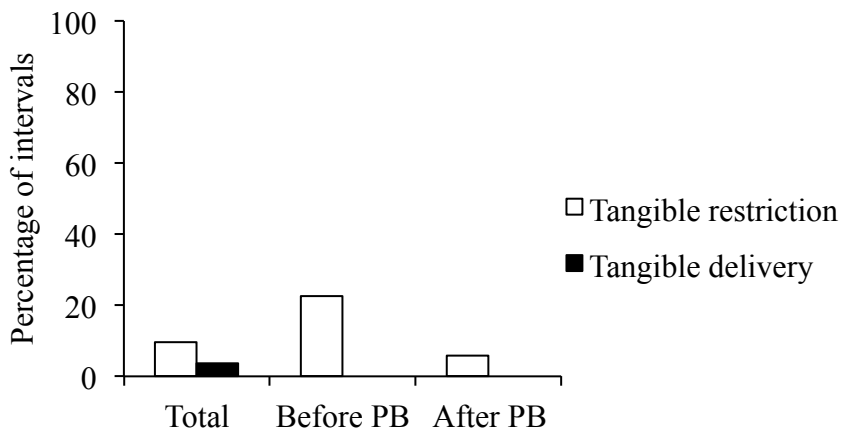
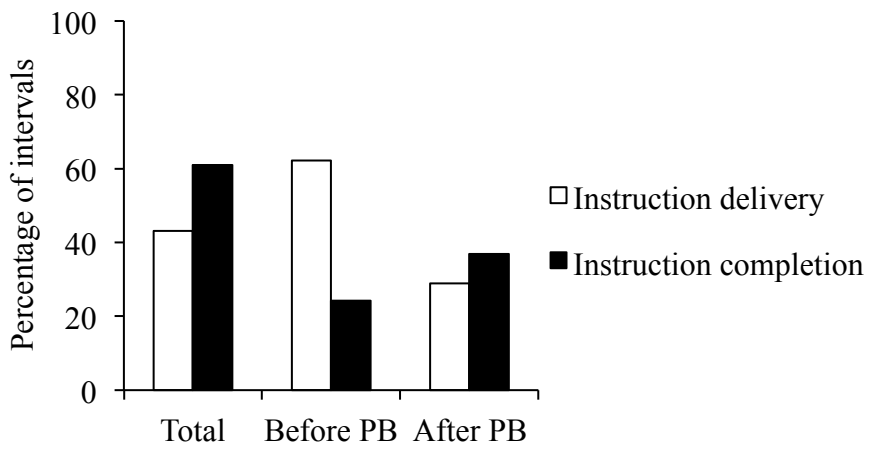
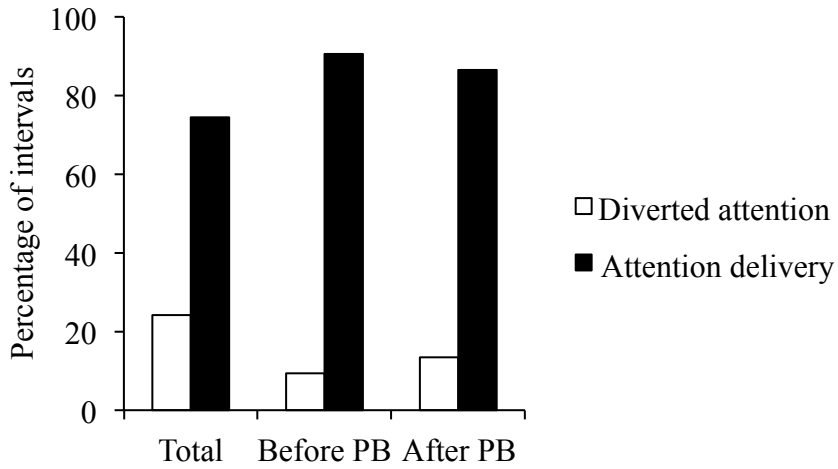


Figure 1. Descriptive assessment results for Abhi.

**Sid.** Results of Sid's descriptive assessments are displayed in Figure 2. Across observations, teacher attention was delivered in 93% of intervals, instructions were delivered in 72% of intervals, and preferred items were restricted and delivered in 7% and 4% of intervals, respectively. Because teacher attention was delivered in the vast majority of intervals, and because Sid was observed to request attention appropriately, attention was not identified as a hypothesized reinforcer for problem behavior. Compared to corresponding estimated base rates, the likelihood of instruction delivery immediately before problem behavior was higher, the likelihood of instruction delivery immediately after problem behavior was lower, and the likelihoods of instruction completion both before and after problem behavior were lower. Escape was thus identified as a hypothesized reinforcer. The base rates of tangible restriction and tangible delivery were low overall, yet there was a higher likelihood of preferred items being restricted immediately before problem behavior and a slightly higher likelihood of these items being delivered immediately after problem behavior. Access to tangibles was therefore identified as a hypothesized function. The escape and tangible hypotheses also were supported by TA report and informal observations of changes in Sid's affect surrounding these events.

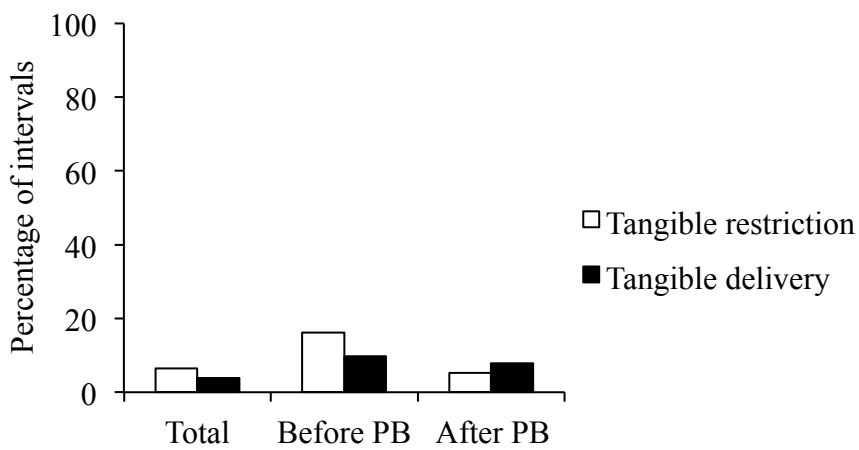
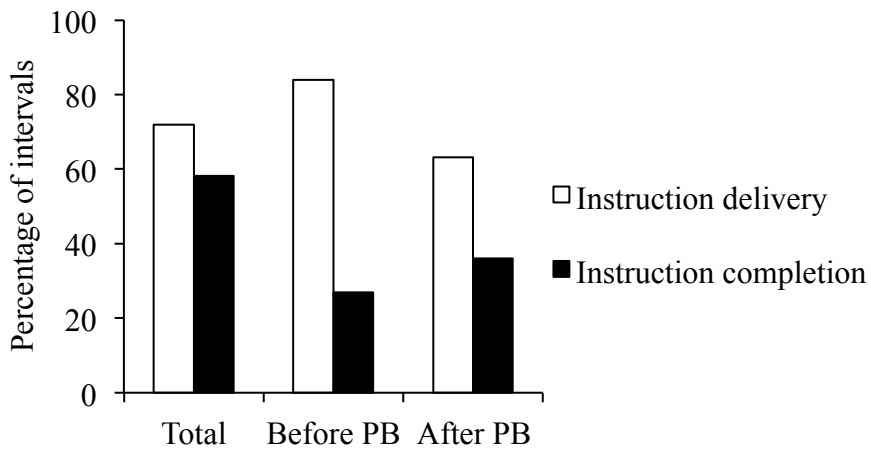
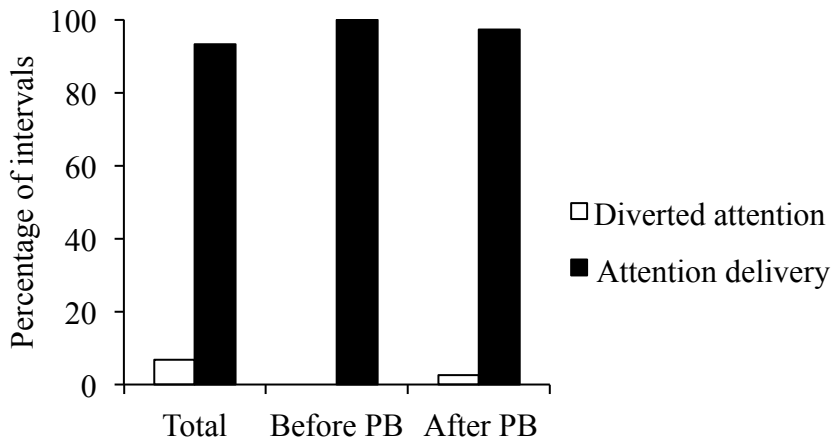


Figure 2. Descriptive assessment results for Sid.

**Davis.** Results of Davis's descriptive assessments are displayed in Figure 3. Across observations, teacher attention was delivered in 78% of intervals, instructions were delivered in 49% of intervals, and preferred items were restricted and delivered in 4% of intervals. Although the likelihood of diverted attention was slightly higher immediately before problem behavior and the likelihood of attention delivery was slightly higher immediately after problem behavior (each compared to relative base rate estimates), these differences were very small and thus difficult to interpret. Because this pattern was supported by TA report and informal observations of changes in Davis's affect when attention was diverted and delivered, attention was identified as a hypothesized reinforcer. Compared to corresponding estimated base rates, the likelihood of instruction delivery was lower immediately before and after problem behavior, with an increased likelihood of instruction completion following problem behavior. Because the likelihood of instruction delivery was lowest following problem behavior, escape was identified as a hypothesized reinforcer. For Davis, the majority of instruction completions included hand-over-hand prompting (rather than compliance), thus it was not clear whether Davis's problem behavior may have been sensitive to this type of prompting. A tangible condition was not included based on the low estimated base rates of tangible restriction and delivery with similarly low likelihoods of each event surrounding problem behavior. Neither TA report nor anecdotal observations supported the inclusion of a tangible condition.

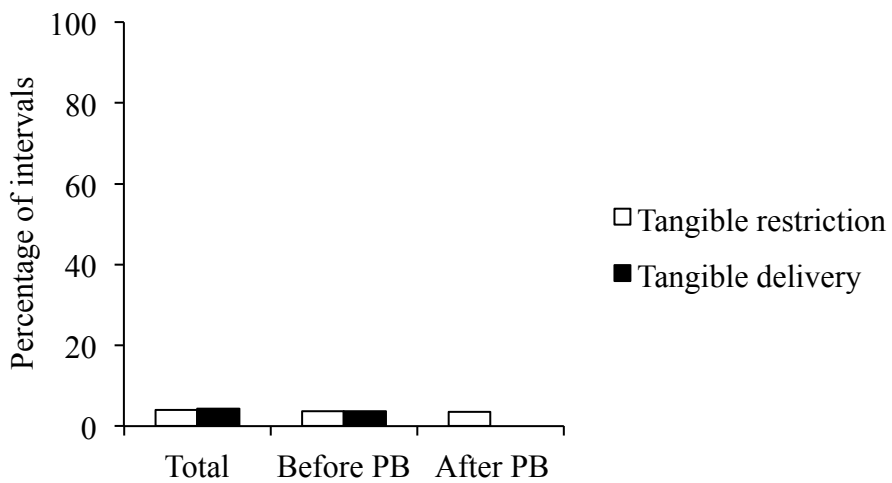
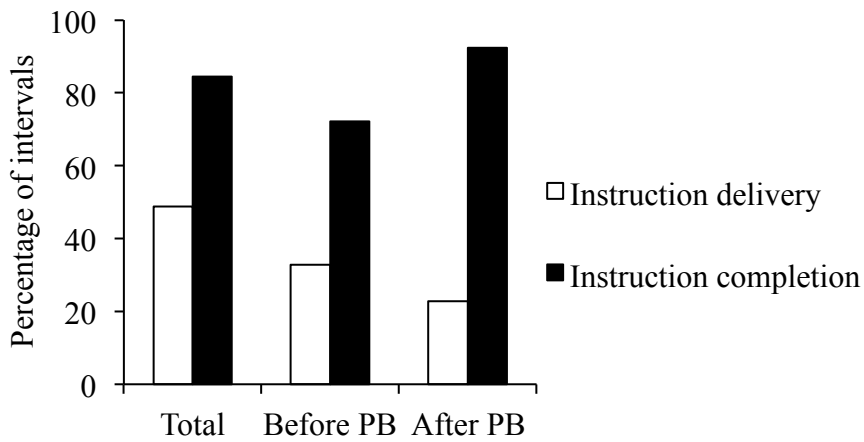
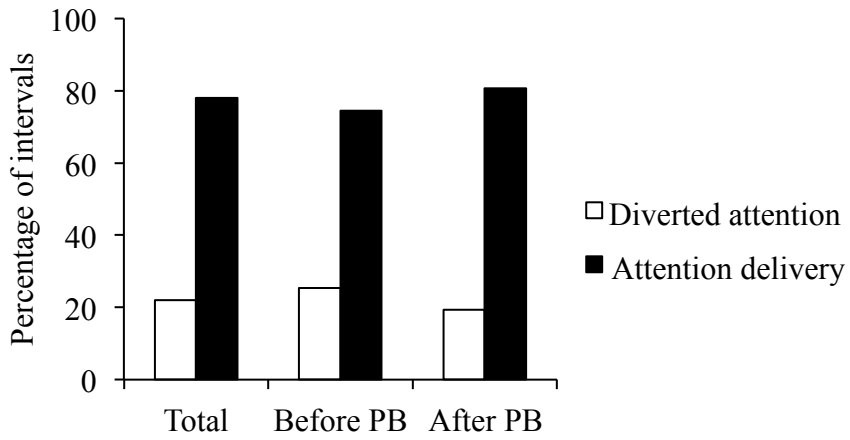


Figure 3. Descriptive assessment results for Davis.

**Gretchen.** Results of Gretchen's descriptive assessments are displayed in Figure 4. Across observations, teacher attention was delivered in 90% of intervals, instructions were delivered in 66% of intervals, and preferred items were restricted and delivered in 18% and 12% of intervals, respectively. Although the estimated base rate of attention delivery was very high, attention delivery did follow problem behavior in nearly all intervals in which problem behavior was coded. Attention was therefore included as a test condition in the subsequent experimental analysis to rule out this function, as Gretchen's teachers reported adult attention did not seem to influence her problem behavior. Based on the slightly increased likelihood of instruction delivery before and after problem behavior relative to the estimated base rate, and decreased likelihoods of instruction completion immediately surrounding problem behavior relative to the estimated base rate, escape from instruction was identified as a hypothesized function. TA report and anecdotal observation (e.g., difficult transitions to the work table and attempts to leave the instructional setting) also supported the inclusion of an escape condition. The estimated base rates of both tangible restrictions and deliveries were relatively low, yet the likelihood of preferred items being restricted was slightly higher before problem behavior than after problem behavior, whereas the likelihood of preferred items being delivered was slightly higher after problem behavior than before problem behavior. Access to tangibles was therefore identified as a hypothesized reinforcer, which was also supported by TA report and anecdotal observation (e.g., frequent attempts to access beads and scooter).

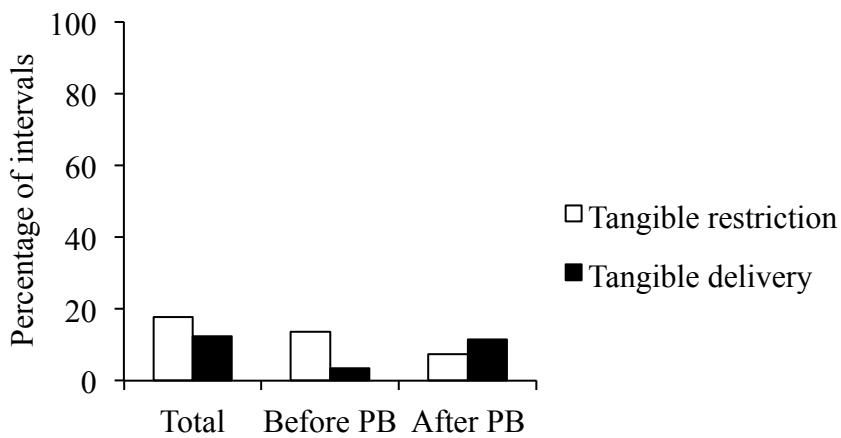
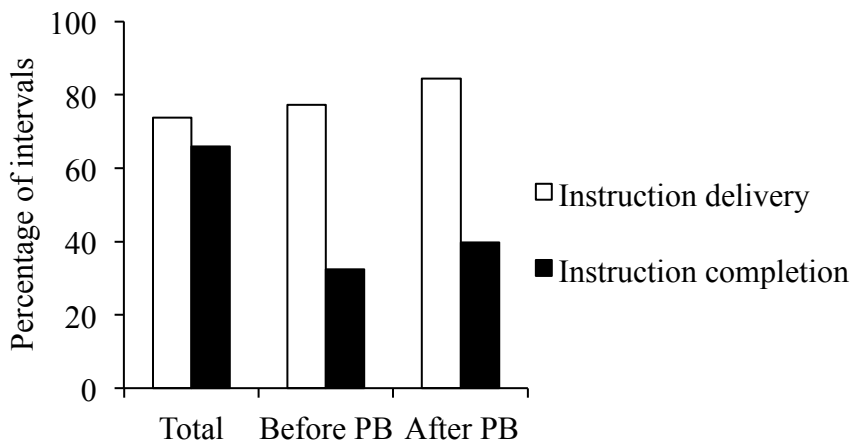
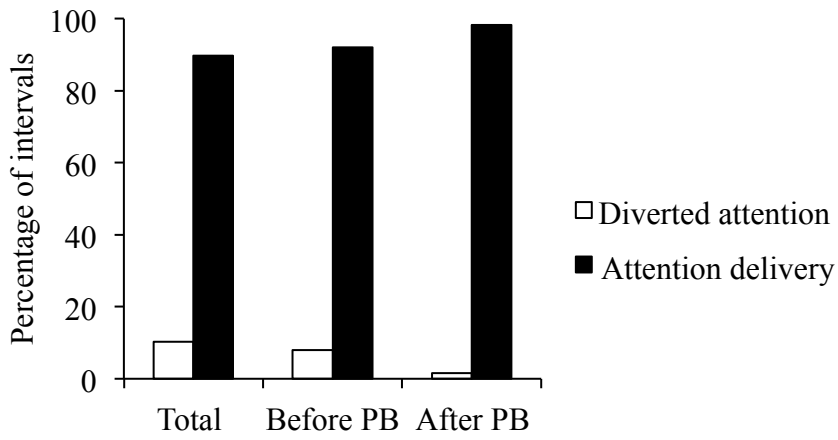


Figure 4. Descriptive assessment results for Gretchen.



## **Trial-based FA**

Results of each participant's trial-based FA are displayed in Figures 5-9. Across participants, open squares represent the occurrence or nonoccurrence of problem behavior during test intervals and  $x$  symbols represent the occurrence or nonoccurrence of problem behavior during control intervals. A steady pattern of response differentiation between test and control intervals (i.e., occurrence of problem behavior during the test interval and nonoccurrence of problem behavior during the control interval) across trials was interpreted as confirming the hypothesized function associated with each condition.

**Abhi.** Results of Abhi's FA are displayed in Figure 5. Abhi's FA included five attention trials and five escape + tangible trials that were rapidly alternated across 3 days of data collection (within a 9-day period). Across attention trials, problem behavior did not occur in test or control intervals. Across escape + tangible trials, problem behavior occurred in all test intervals and did not occur in four of five control intervals. In Trial 5, aggression occurred during the test interval and an elopement attempt occurred after Mia announced Abhi could have more time on the computer but before the computer monitor was turned on. This elopement attempt was therefore documented as an instance of problem behavior during the control interval; however, no problem behavior occurred after the monitor was turned on. The pattern of response differentiation between test and control intervals of the escape + tangible trials supported the hypothesis that problem behavior was maintained by a combination of positive and negative reinforcement in the form of accessing computer breaks.

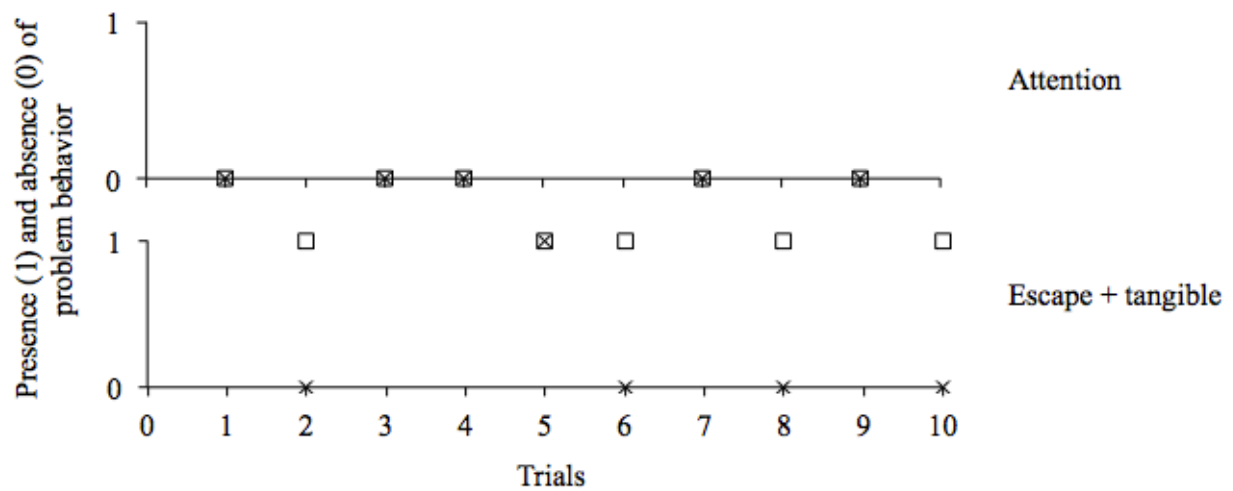


Figure 5. Results of trial-based FA for Abhi.

**Sid.** Results of Sid’s FA are displayed in Figure 6. Sid’s FA included eight trials of each condition that were rapidly alternated across 5 days of data collection (within an 8-day period). Across escape from novel instruction trials, problem behavior occurred during the test interval, and did not occur in seven of eight control intervals. During the final two trials of this condition, Lorraine used a set of unfamiliar cards that were topographically similar to the set of familiar cards (i.e., glossy, included pictures). This change was made to provide a more stringent test of the novelty hypothesis, as it was unclear whether the differentiated response pattern may be due to other physical differences between the two card sets aside from novelty. Because the same differentiated pattern was identified in these final two probe trials, results supported the hypothesis that problem behavior was negatively reinforced by the removal of novel/unfamiliar instructional materials.

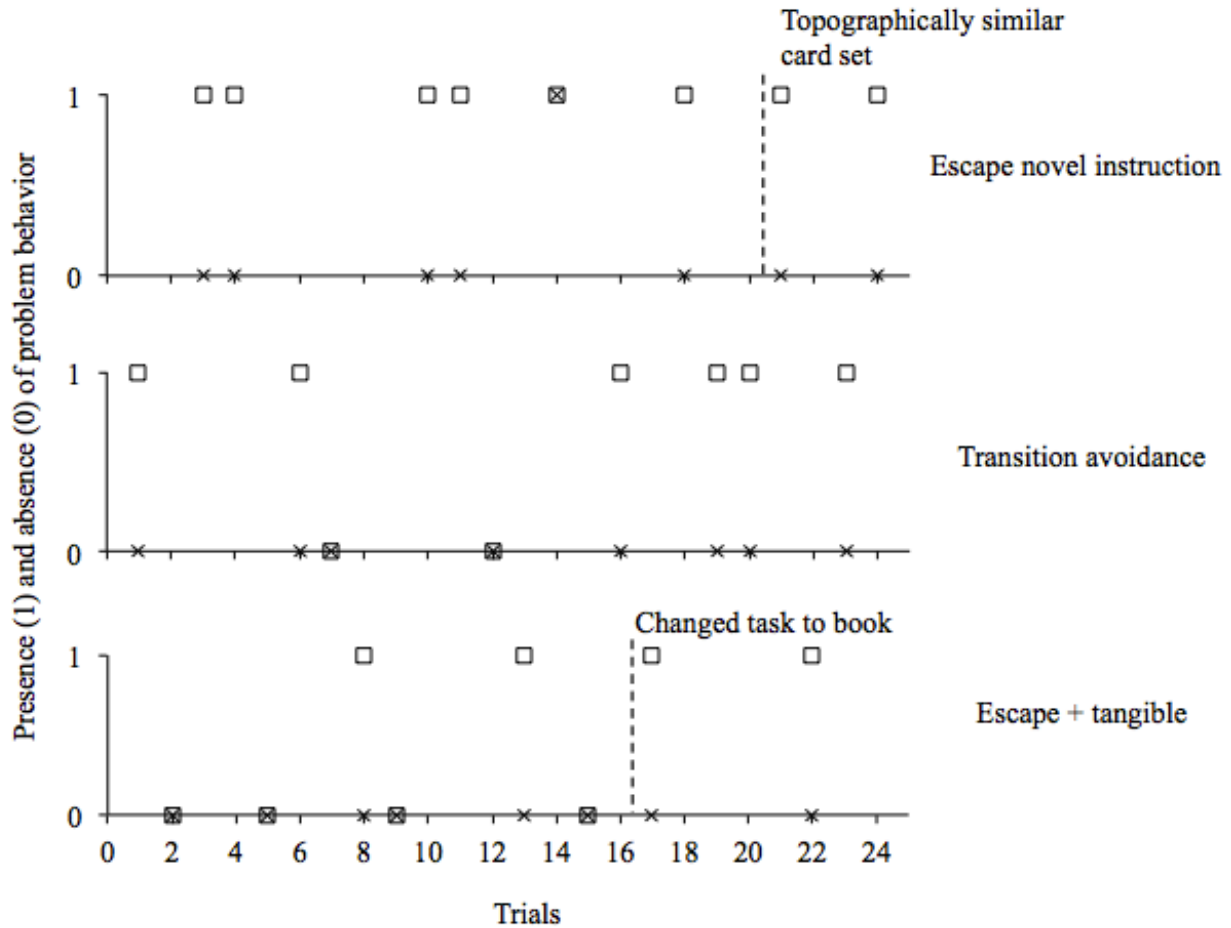


Figure 6. Results of trial-based FA for Sid.

Across transition avoidance trials, problem behavior occurred during six of eight test intervals and did not occur during control intervals. This pattern supported the hypothesis that problem behavior was maintained by negative reinforcement in the form of avoiding or terminating transitions from the computer to the work table. During escape + tangible trials, problem behavior did not occur in either interval of four of the first six trials. At this point, the TA suggested we change the instructional activity from sorting bears to reading a book. Lorraine reported that problem behavior surrounding the restriction of Sid’s penguin typically occurred within the context of non-preferred instructional activities, and was concerned that the bear-

sorting task may be a preferred instructional activity. Thus, beginning at trial 17, we changed the instructional activity to the book, and observed response differentiation between test and control intervals during the next two trials of this condition. As intervention procedures were initiated in the remaining two conditions, we intermittently probed this modified escape + tangible trial for additional replications. Taken together, results of Sid's trial-based FA supported all three negative reinforcement hypotheses.

**Davis.** Results of Davis's trial-based FA are displayed in Figure 7, with closed data points representing those occurring in art class and open data points representing those occurring in PE. Davis's trial-based FA included seven trials per condition that were rapidly alternated across 5 days of data collection (within a 13-day period). Across attention trials, problem behavior occurred during all test intervals and did not occur during any control intervals. This pattern of response differentiation supported the hypothesis that Davis's problem behavior was maintained by positive reinforcement in the form of adult attention. Across physical prompt avoidance trials, problem behavior occurred in three of seven test intervals, but also continued in control intervals of these trials. In the remaining trials, problem behavior either did not occur or occurred during control intervals only. This pattern did not support the hypothesis that Davis's problem behavior was maintained by negative reinforcement in the form of avoiding physical prompts. Across escape of 1:1 instruction trials that were conducted in PE class, problem behavior occurred in four of the five test intervals and did not occur in any control intervals. Only two trials of this type were conducted in art class, in which problem behavior did not occur in either interval. Intermittent probes of this condition were continued after introducing intervention procedures in the attention condition, the results of which supported the hypothesis

that Davis's problem behavior was maintained by negative reinforcement in the form of escaping 1:1 instruction and his chair (at least within the context of PE class).

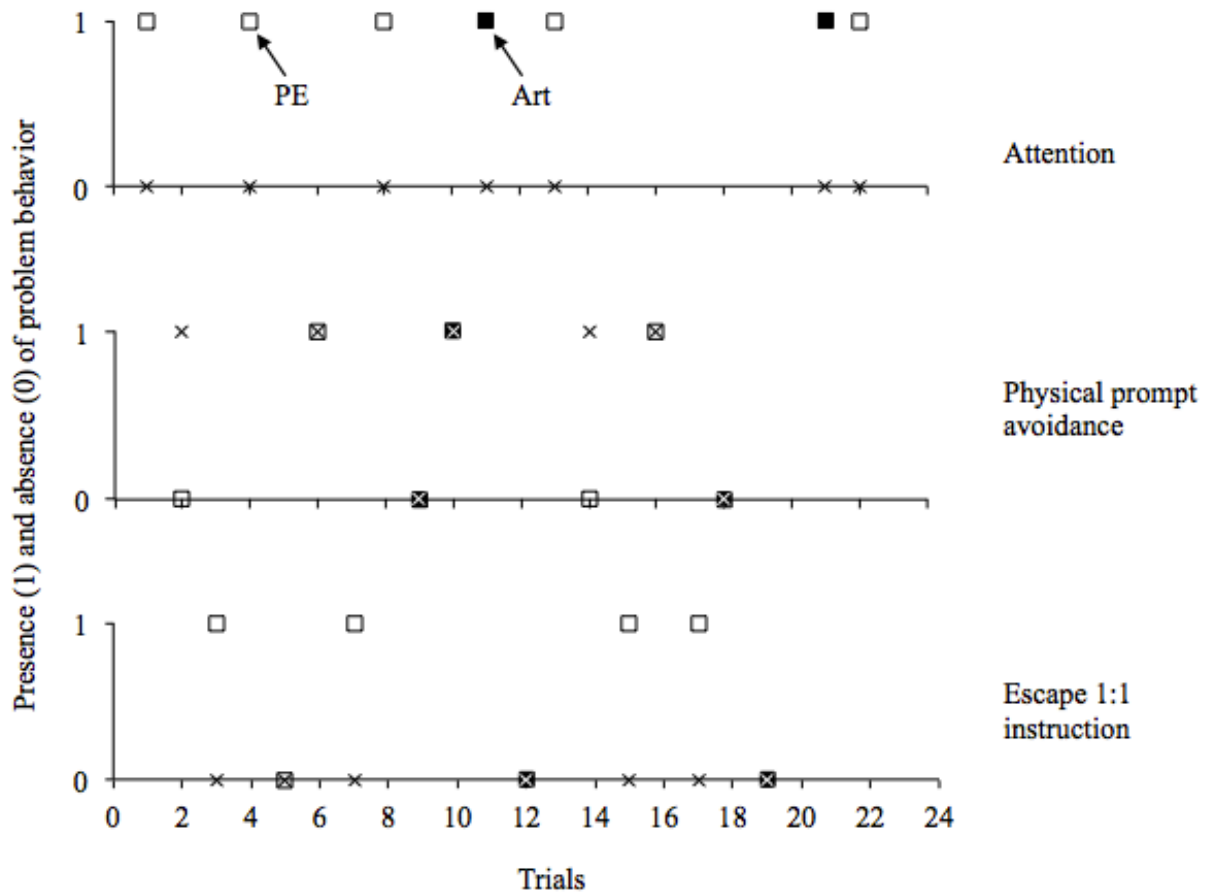


Figure 7. Results of trial-based FA for Davis.

**Gretchen.** Results of Gretchen's trial-based FA are displayed in Figures 8 and 9. In Figure 8, all topographies of problem behavior are included as target behavior. In Figure 9, out-of-seat behavior is not included as target behavior. That is, to be counted as target behavior, a behavior beyond standing up from her seat (e.g., running across the room, flopping to the floor, attempting to climb on Elaine's lap) had to have occurred. The data were re-graphed this way to

explore the possibility that out-of-seat behavior on its own was a member of a separate response class and thus should not have been included in the definition of elopement.

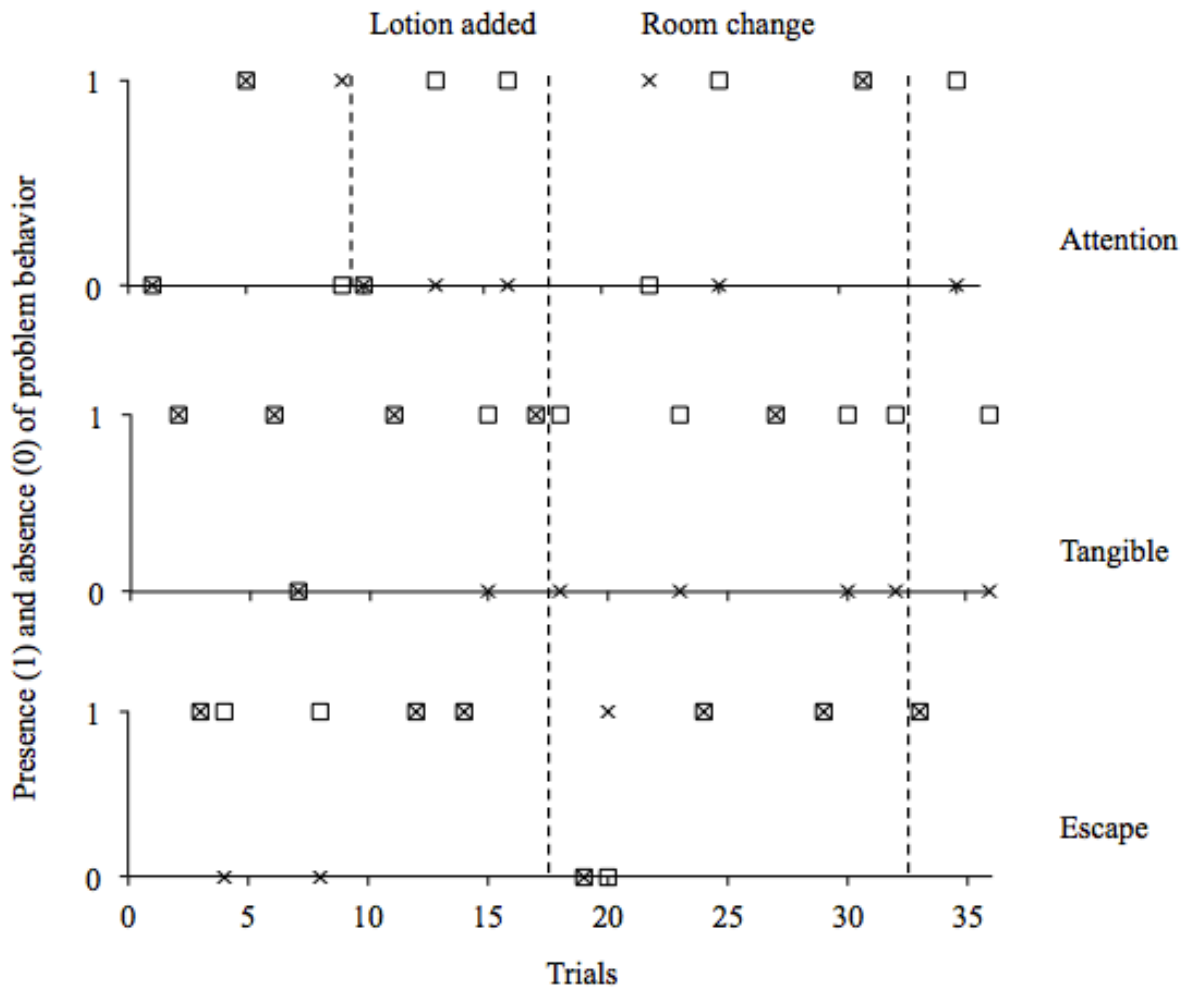


Figure 8. Results of trial-based FA for Gretchen including all problem behaviors.

Gretchen’s trial-based FA included 10 attention and escape trials and 12 tangible trials that were rapidly alternated across 7 days of data collection (within a 12-day period). Trials 18-32 were conducted in an early childhood education classroom (shown on graph) due to standardized testing taking place in Gretchen’s special education classroom. The first three

attention trials included verbal attention delivery during the control interval. Elaine's difficulty providing verbal attention without delivering instructions or asking questions suggested that this type of attention did not match the attention Gretchen typically received from her TA. Based on this observation in addition to the occurrence of problem behavior during control intervals, we sought Elaine's feedback and changed the type of attention to physical attention in the form of the lotion activity. Following the addition of lotion to the control interval, problem behavior occurred in five of seven test intervals and did not occur in five of seven control intervals (though only four of seven trials produced the expected differentiated response pattern). These results, at least to some degree, supported the hypothesis that problem behavior was maintained by positive reinforcement in the form of physical attention. Across tangible trials, problem behavior occurred in 11 of 12 test intervals, but carried over into five of these trials' control intervals. However, the topography of problem behavior from test to control interval was different during several trials. When the out-of-seat behavior was excluded from target behavior, the expected pattern of response differentiation was identified in eight trials, including the final six consecutive trials. Taken together, these results supported the positive reinforcement hypothesis in the form of access to tangibles. Across escape trials, problem behavior occurred in eight of ten test intervals, but carried over into six of these trials' control intervals. Removing out-of-seat behavior did not clarify results, thus results did not support the negative reinforcement hypothesis.

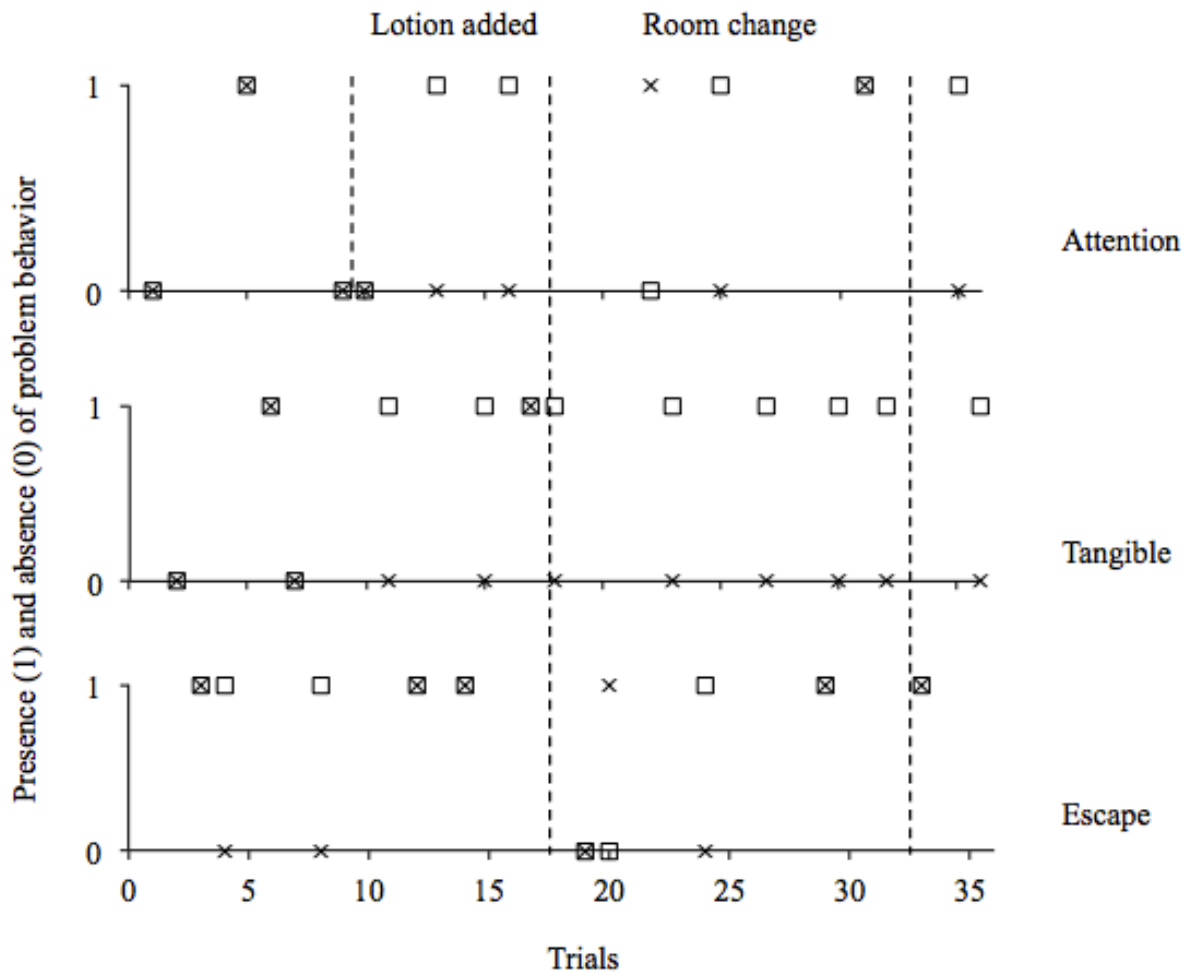


Figure 9. Results of trial-based FA for Gretchen excluding out-of-seat behavior.

### Intervention

Results of each participant's intervention trials are displayed in Figures 10-14. Across participants, squares represent the occurrence or nonoccurrence of problem behavior during test intervals, triangles represent the occurrence or nonoccurrence of replacement behavior during test intervals, and x symbols represent the occurrence or nonoccurrence of problem or replacement behavior during control intervals. Closed triangles indicate prompted replacement behaviors, whereas open triangles indicate independent (unprompted) replacement behaviors.



**Abhi.** Results of Abhi’s intervention trials are displayed in Figure 10, and included 17 trials across 5 days of data collection (within a 12-day period). With the exception of one trial, problem behavior did not occur during intervention trials and was replaced by prompted and independent picture exchanges. A brief reversal of the intervention procedures (i.e., return to assessment procedures) resulted in an immediate return of problem behavior, with an attempted picture exchange in the first reversal trial and no picture exchanges in the second reversal trial (Mia placed her hand on the cover of the book briefly within this trial). When intervention procedures were re-introduced, the replacement of problem behavior by independent picture exchanges was replicated. Results of Abhi’s intervention trials validated the trial-based FA outcome by showing the replacement of problem behavior via the modification of reinforcement contingencies.

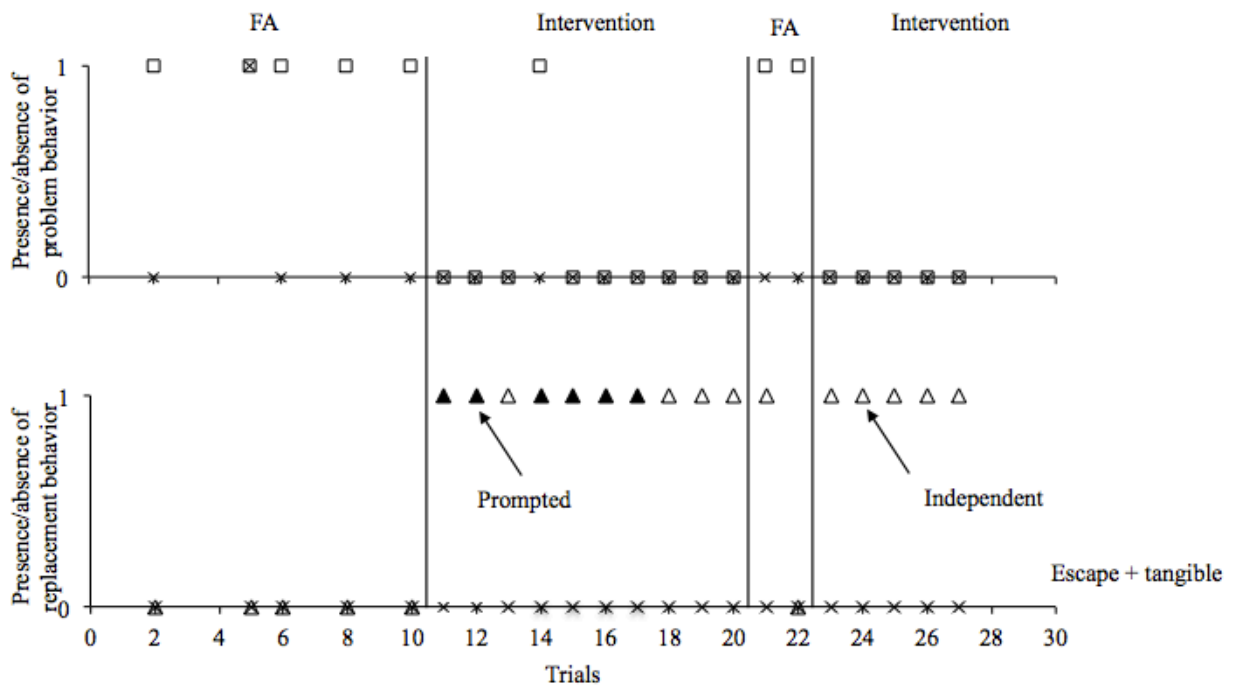


Figure 10. Results of intervention trials for Abhi.

**Sid.** Results of Sid’s intervention trials are displayed in Figure 11, and included 47 trials across 10 days of data collection (within a 14-day period). Following the introduction of intervention procedures in the escape novel instruction condition, the verbally-prompted replacement behavior occurred across all trials, but problem behavior also occurred in five of the first seven intervention trials. Trials continued in this condition until a minimum of three trials occurred without problem behavior, at which point the intervention procedures were introduced in the next condition. Upon introduction of intervention procedures to the transition avoidance condition, verbally-prompted replacement behavior occurred in some, but not all trials and problem behavior continued in the majority of trials. Beginning at Trial 57, we introduced a visual stimulus of the task materials to show Sid what he was being prompted to refuse. Following this procedural adjustment, prompted replacement behaviors occurred across trials and problem behavior stopped occurring. Following a minimum of three trials with prompted replacement behavior and without problem behavior, intervention procedures were introduced in the third and final condition. Because Sid’s teachers and family were planning to fade his stuffed penguin from his school routine all together, we conducted the minimum number of trials necessary to demonstrate the intervention effect. Following the initial trial with both prompted replacement behavior and problem behavior, the following three trials demonstrated replacement behavior without problem behavior. Results of Sid’s intervention trials validated the trial-based FA outcome by showing the replacement of problem behavior via the modification of reinforcement contingencies (though Sid’s “No thank you” response remained dependent on verbal prompts from his TA).

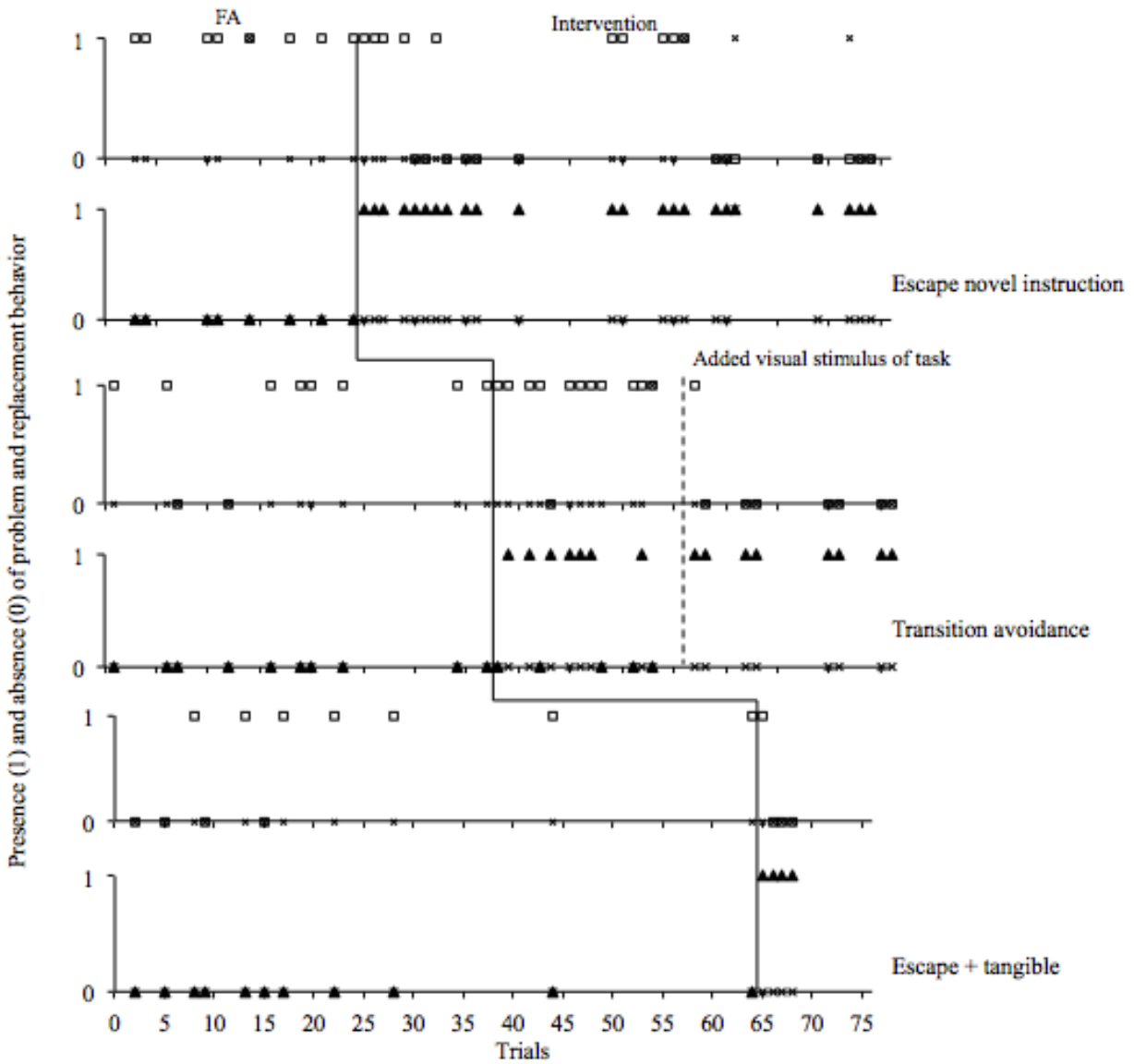


Figure 11. Results of intervention trials for Sid.

**Davis.** Results of Davis’s intervention trials are displayed in Figure 12, with closed data points representing trials conducted in art class, and open data points representing those conducted in PE class. Davis’s intervention included 21 trials across 6 days of data collection (within a 15-day period). When the intervention procedures were introduced in the first attention trial (PE), multiple button presses occurred during both the test and control intervals. Beginning

at the second attention trial, Darcy provided attention and removed the button following the first button press to teach the contingency between the button press and delivery of attention. With the exception of two intervention trials in which problem behavior occurred in the control interval, button presses effectively replaced problem behavior during attention trials. Trials in this condition continued until independent button presses occurred across both PE and Art class contexts.

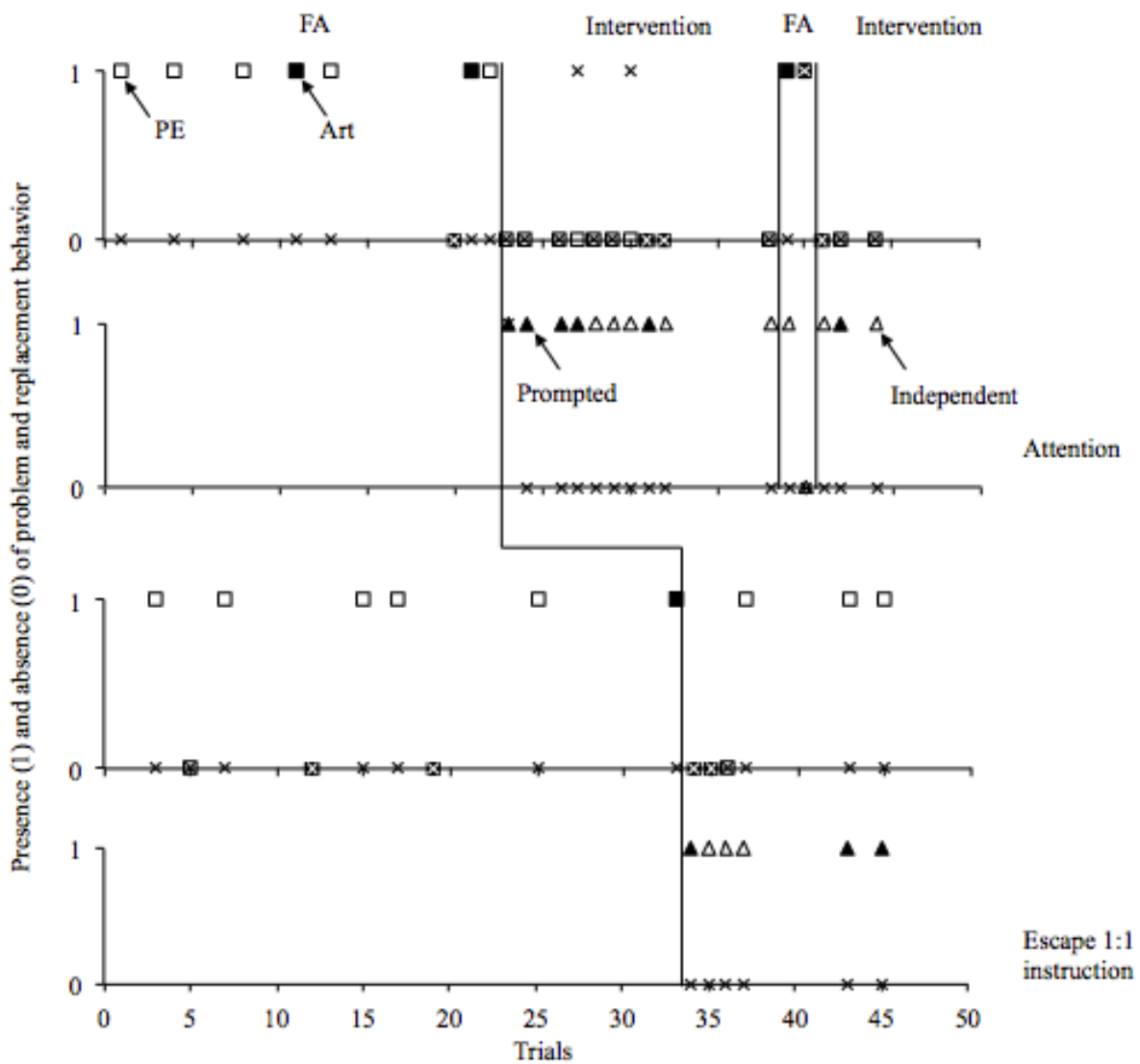


Figure 12. Results of intervention trials for Davis.

Intervention procedures were then introduced in the escape from 1:1 instruction condition, which was done in Art class. It was unclear whether the “Let’s take a walk” button press should be taught in Art class, based on the lack of problem behavior observed in the two assessment trials of this condition that were conducted in Art. However, in the final assessment probe, response differentiation was observed, and we decided to introduce the intervention procedures in this condition. Problem behavior did not occur in three of the first four intervention trials, and three consecutive trials with independent button presses were observed.

We then conducted a brief reversal to assessment procedures in the attention condition to provide an additional opportunity to replicate the treatment effect. Problem behavior immediately returned in the test interval of the first reversal trial and in both intervals of the second reversal trial, with a button press occurring in the first reversal trial but not in the second reversal trial. Intervention procedures were reintroduced, during which problem behavior was again replaced by the button press response. Trials 42-45, however, were conducted on a day that Davis arrived late and was sent straight to PE before eating lunch (Davis typically ate lunch before specials classes). On this last day of data collection, button presses were prompted in the two final escape trials and in one of the two final attention trials. In addition, problem behavior occurred during the test interval of the final two escape trials. Despite these less clear behavior patterns on the final day of data collection, the results of Davis’s intervention trials validated the trial-based FA outcomes. Had it not been so close to the end of the school year, additional intervention data would have been collected.

**Gretchen.** Results of Gretchen’s intervention trials are displayed in Figures 13 and 14. In Figure 13, all topographies of problem behavior are included as target behavior. In Figure 14, out-of-seat behavior was excluded from target behavior as was done for the FA data.

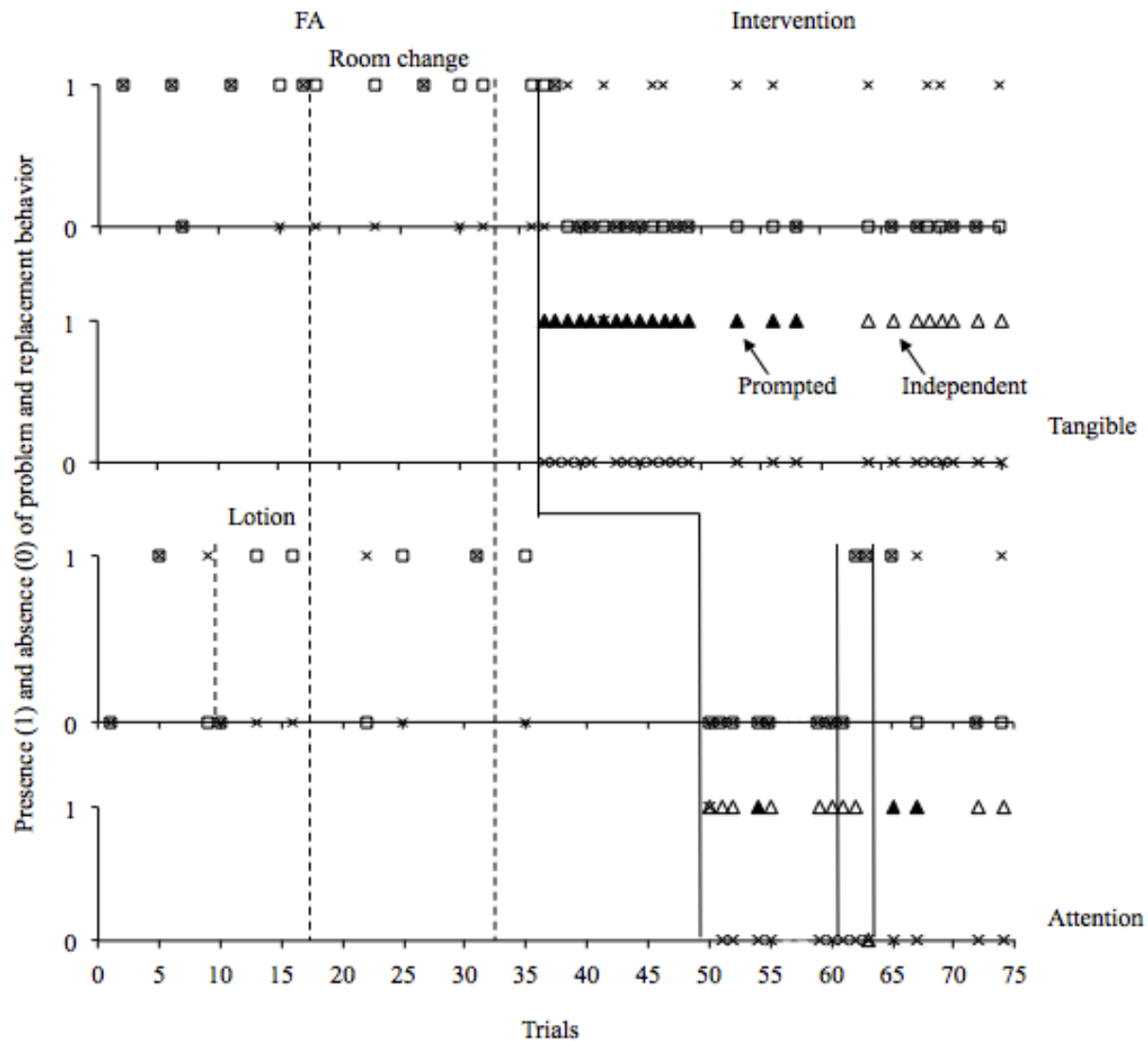


Figure 13. Results of intervention trials for Gretchen including all problem behavior.

Gretchen’s intervention included 39 trials across 9 days of data collection (within a 12-day period). Intervention procedures were introduced first in the tangible condition, as this was the condition with the clearest FA results. During the first two intervention trials, the picture exchange was physically prompted, and problem behavior occurred during these physical prompts. In the next series of intervention trials, the physical prompt was removed and only a verbal and gestural prompt was given. Across trials, the prompts were gradually faded out such

that the picture exchange was occurring without problem behavior during the test interval. As can be seen when comparing Figures 13 and 14, the problem behavior that occurred somewhat frequently during the control interval included the out of seat behavior (i.e., Gretchen would stand up suddenly, be prompted to sit back down, and would comply).

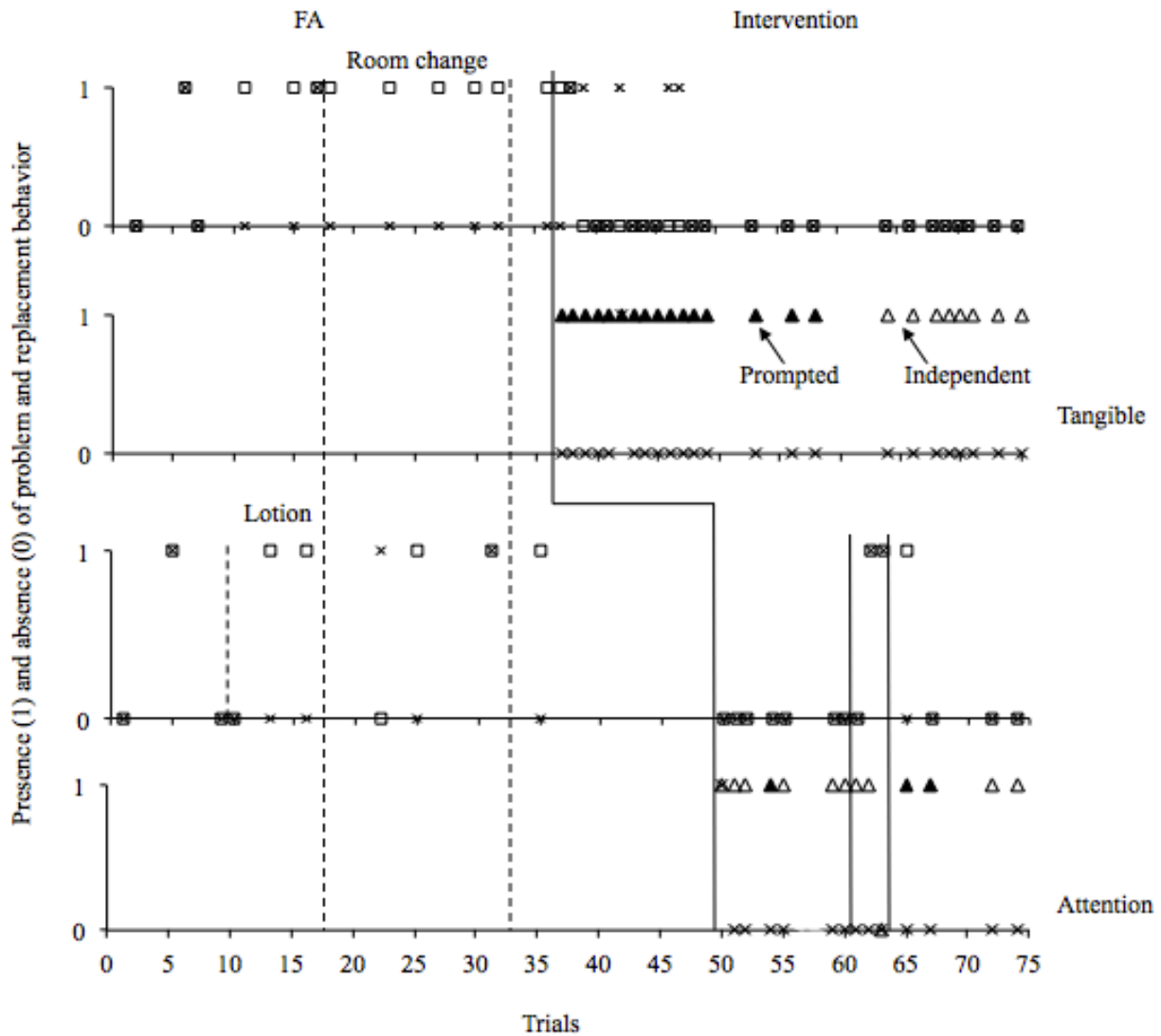


Figure 14. Results of intervention trials for Gretchen excluding out-of-seat behavior.

Following several consecutive trials with a successful picture exchange following a gestural prompt only (opportunities for independent picture exchanges were present via a 30-s delay to the gestural prompt) and without problem behavior in the test interval, intervention procedures were introduced in the attention condition. Interestingly, the first unprompted picture exchange occurred during the first attention intervention trial, and independent and gesturally-prompted picture exchanges occurred without problem behavior in subsequent trials of this condition. Because a final probe trial was not conducted in the attention condition before implementing intervention (a procedural error), a brief reversal was programmed following three consecutive trials with independent pictures exchanges to replicate the intervention effect. Three reversal trials (i.e., return to assessment procedures) were conducted. In the first reversal trial, multiple attempts to exchange the picture occurred without problem behavior. In the second trial, a single attempt to exchange the picture occurred and was followed by problem behavior across test and control intervals. In the third reversal trial, no attempts to exchange the picture were made and problem behavior occurred across intervals. When intervention procedures were reintroduced in this condition, both prompted replacement behavior and problem behavior occurred during the first trial, followed by three trials with picture exchanges without problem behavior. Surprisingly, it was not until the reversal trials in the attention condition were conducted that Gretchen began to independently request access to beads in the tangible condition. Taken together, results of Gretchen's intervention trials validated the trial-based FA outcomes, although clearer patterns of response differentiation were observed when the out-of-seat behavior by itself was not included as problem behavior.



## Social Validity Measures

**TA questionnaire.** TA responses to each questionnaire item are displayed in Table 4. All TAs rated all 10 acceptability-related items as a 5 or above (6 = strongly agree) with the exception of one 4 (i.e., Elaine's response to whether the intervention should prove effective in changing the student's problem behavior). In response to how similar each trial was to events the student typically encounters in the relevant setting (1=very similar, 6=very different), Mia rated the escape + tangible trial a 1 and the attention trial a 4. Darcy rated the attention trial a 1 and the avoidance of physical prompts trial and escape from 1:1 instruction trial as a 2. Lorraine and Elaine rated all of the trials they implemented as a 1. In response to how difficult each trial was to implement (1=very easy, 6=very difficult), all TAs rated all trials a 1 with one exception: Mia's rating of the escape + tangible trial (4). Taken together, with few exceptions, responses across items and TAs indicated high levels of acceptability, similarity between experimental trials and events each student typically encountered with that TA, and low levels of difficulty in terms of implementation. The extent to which TAs' responses were influenced by other variables (e.g., rapport with the researcher, expectations of how their responses would be perceived by others) is unknown.

Table 4

*TA responses on social validity questionnaire*

Item (1=strongly disagree; 6=strongly agree)	Mia	Lorraine	Darcy	Elaine
This was an acceptable assessment and intervention for this student's problem behavior.	6	6	6	5
Most teachers would find this assessment and intervention appropriate for their students' behavior problems.	5	5	5	5
This intervention should prove effective in changing the student's problem behavior.	6	5	5	4
I would suggest the use of this assessment and intervention to other teachers.	6	5	5	6
The student's problem behavior is severe enough to warrant use of this assessment and intervention.	6	6	5	6
I would be willing to use this assessment and intervention in the classroom setting.	5	6	5	6
This assessment and intervention is NOT likely to result in negative side effects for the student.	6	6	5	6
This assessment and intervention is a fair way to handle the student's problem behavior.	6	6	5	5
I liked the procedures used in this assessment and intervention.	6	6	5	5
Overall, this assessment and intervention will be beneficial for this student.	6	6	5	5
How similar was this trial to events the student typically encounters in this setting? (1=very similar; 6=very different)				
Attention trial	4	--	1	1
Tangible trial	--	--	--	1
Escape trial	--	--	2	1
Escape + tangible trial	1	1	--	--
Transition avoidance trial	--	1	--	--
Escape novel instruction trial	--	1	--	--
Avoidance of physical prompts trial	--	--	2	--
How difficult was this trial to implement? (1=very easy; 6=very difficult)				
Attention trial	1	--	1	1
Tangible trial	--	--	--	1
Escape trial	--	--	1	1
Escape + tangible trial	4	1	--	--
Transition avoidance trial	--	1	--	--
Escape novel instruction trial	--	1	--	--
Avoidance of physical prompts trial	--	--	1	--

**Trial durations.** The duration of experimental trials in minutes, days, and weeks are displayed in Table 5. Across participants, the average number of minutes to complete the trial-based FA and intervention trials was 36 min (range, 17-50) and 44 min (range, 21-71), respectively. That is, with the exception of Sid's intervention trials, each assessment and intervention phase required less than 1 hr of total time spent in experimental trials. Of course, the extent to which typical instructional routines can continue during inter-trial intervals impacts the total time required for each evaluation.

Aspects of assessment results, including the extent of response differentiation within and among conditions, and the number of hypotheses tested and confirmed, also impact the total duration of assessment and intervention trials. That is, Abhi's FA and intervention trials required the least amount of time to complete, as his FA trials revealed the clearest pattern of response differentiation, and only one hypothesis was confirmed. Gretchen's FA trials required the most time to complete, as the patterns of response differentiation were less clear. Sid's intervention trials required the most time to complete, as all three hypotheses were confirmed and included in his intervention. Each trial-based FA and intervention evaluation was completed within a span of 2-3 school weeks, with Davis's spanning 6 weeks total due to the schedule of his specials classes (PE two days per week and Art 1 day per week).

Table 5

*Duration of trial-based FAs and interventions by participant*

Participant	Min	Days of data collection	Time span (weeks of school)
Abhi			
FA	17	3	2
Intervention	21	5	3
Sid			
FA	38	5	2
Intervention	71	10	3
Davis			
FA	37	5	3
Intervention	31	6	3
Gretchen			
FA	50	7	2
Intervention	54	9	2

## CHAPTER IV

### DISCUSSION

In the current study, we replicated the trial-based FA procedures in public elementary school settings with TAs implementing all experimental trials. We collected descriptive data on naturally-occurring contingencies and collaborated with teaching staff to plan the experimental assessment and intervention trials. Across all participants, a minimum of one hypothesis generated from the descriptive data and TA feedback was confirmed via the test-control trial procedures and subsequent intervention trials resulted in a socially appropriate communication response replacing problem behavior.

Results of the current study contribute to the small but growing evidence base on trial-based FAs in several ways. First, at least one response-reinforcer relation was identified for each participant and subsequently validated by results of intervention trials consisting of contingency reversals. The test-control sequence of FA trials used in the current study allowed for the identification of response-reinforcer relations, as the hypothesized reinforcer was delivered contingent on target behavior. Second, functional relations between problem behavior and environmental stimuli were identified with relative efficiency using the brief trial format. Despite an anticipated challenge surrounding the brevity of trials (i.e., whether 60-s test intervals would be sufficient to evoke problem behavior), problem behavior reliably occurred during test intervals across participants. The FA and intervention phases each spanned between two and three school weeks across participants, yet the average total time spent in experimental trials per assessment and intervention was between 30 min and 1 hr. As long as typical instructional

routines remain in effect between trials, these results suggest the trials themselves should interfere minimally with the TAs' and students' usual routines. Additionally, because the total number of trials per assessment was based on response differentiation rather than a pre-determined number, we prevented any single participant's trial-based FA from being any longer than necessary before proceeding to the intervention.

Third, a range of behavior-environment relations were successfully identified across participants. That is, the trial-based methodology proved flexible not only in assessing and teaching various topographies of problem and replacement behaviors, but various combinations of environmental stimuli (e.g., escape + tangible, transition avoidance, physical prompt avoidance, escape novel instruction). These adaptations of standard FA conditions were more likely to result in the identification of functional relations that represent naturally-occurring behavior-environment relations for each student. Fourth, TAs were able to implement experimental trials with sufficient levels of fidelity to establish experimental control. A researcher consistently prompted the transition from test to control intervals during assessment trials across participants, but the prompt was no longer necessary to maintain fidelity during intervention trials. Fifth, TAs responded favorably to questionnaire items on (a) the acceptability of procedures, (b) the similarity between experimental trials and events the student typically encounters, and (c) the ease with which trials were conducted.

### **Challenges and Limitations**

Several challenges and limitations of the current study are worthy of discussion, the first of which relates to unexpected procedural constraints. Although one purpose of this study was to evaluate the feasibility of conducting trial-based FAs in natural classroom settings, the students who met inclusion criteria (e.g., high frequency problem behavior) were receiving instruction in

relatively controlled environments with direct support. Although these conditions were already in effect for these students and thus qualified as the student's usual classroom setting, the results may only apply to other similarly structured instructional settings. The question of whether the trial-based FA procedures could feasibly be implemented in settings approximating general education classrooms has yet to be addressed. Another procedural constraint was that each student's trials were distributed within a 1-hr period of the school day (rather than across the entire day), during which the instructional routine was held constant as well as the TA supporting the student. The consistency of these variables (i.e., time of day, contextual variables, implementers) is an advantage in terms of experimental control, but also limits the likelihood that results may be generalized across other times of day, routines, or TAs providing support.

Other challenges relate to the procedural fidelity of trial implementation. Although procedural fidelity was moderately high overall, levels of fidelity varied by participant, experimental phase, and trial type (see Table 2). For example, only 57% (4 of 7) of Davis's physical prompt avoidance trials were implemented without any procedural errors. Specifically, the procedural error included the delivery of some form of physical prompt during the control interval (during which 1:1 instruction *without* physical prompts was programmed). For Davis's TA, it was difficult to deliver instruction for 60 s without any physical prompting, which perhaps reflects a lack of ecological validity for this condition. When evaluating Davis's responding across trials of this condition, the question becomes whether Davis's inconsistent response pattern was due to the inconsistent implementation of procedures within this condition. Levels of fidelity also were higher during intervention trials than assessment trials across participants. Increased levels of fidelity during intervention trials may have been due to (a) additional practice implementing the trials, (b) alternating among fewer types of trials, (c) conducting trials of the

same condition in succession, (d) transitioning from test to control intervals based on a single, discrete, and newly-acquired behavior, or (e) any combination thereof.

Finally, fidelity was calculated according to whether the critical programmed events per test and control interval of each trial were implemented. In many cases, however, other TA behaviors also occurred in one or both of the intervals. The most common example included TAs asking the student a question or delivering a brief instruction during control intervals across conditions. For example, during Sid's escape + tangible trials, Lorraine would correctly remove the penguin and introduce the book activity in the test interval, and correctly deliver the penguin and remove the book activity in the control interval, but might also say, "Give Pablo a hug, Sid!" at some point during the second minute (meeting our definition of instruction delivery). During Gretchen's assessment and intervention trials (especially within the tangible condition), Elaine delivered brief verbal or physical prompts to *sit in seat* during control intervals. Because Gretchen's out-of-seat behavior was a common precursor to running across the room or flopping to the floor, each of which required extended time and effort to return to the original activity, we allowed Elaine to continue providing these prompts on an as-needed basis. Though permitting these prompts resulted in some inconsistencies in whether instruction delivery was coded during control intervals, doing so allowed Elaine to continue instruction during inter-trial-intervals (thus minimizing disruptions to Gretchen's typical routine) as well as conduct multiple trials within the 1-hr block. Had we reported fidelity not only in terms of the critical programmed events, but any other events that met an operational definition within one or more of the original stimulus categories, estimates of procedural fidelity would have been lower. Calculating fidelity based on these definitions, however, seemed stricter than necessary, as experimental control was still preserved across the majority of experimental conditions.



Another potential limitation of the current study is that across participants, several topographies of problem behavior were included as *target behavior*, and were thus treated as a single response class. Behaviors that appeared to occur independent of socially-mediated events (e.g., vocal stereotypies) were not included as target behavior, though this relation was not evaluated experimentally. The question of whether behaviors of more than one response class had been included as target behavior became apparent when interpreting Gretchen's FA and intervention data. The increased response differentiation in the tangible condition of the FA and both intervention conditions when removing out-of-seat from target behavior seemed to suggest that out-of-seat may have been a member of a different response class. Out-of-seat was originally included as a target behavior because it was typically a precursor to flopping to the floor or running across the room. However, as the intervention trials progressed, the form of out-of-seat changed to a sudden 'popping up' to a standing position only (which may have been shaped by Elaine's brief prompts to sit back down throughout intervention).

Although no systematic preference assessments were done prior to designing tangible trials, students' preferred items and activities were reported by TAs during the initial interview and supported via direct observation. It is doubtful that different tangible items would have been incorporated into experimental trials had systematic preference assessments been conducted. However, for students who were taught replacement behaviors to request more than one preferred item or activity, it may have been beneficial to conduct brief preference assessments prior to conducting intervention trials. During a few assessment and intervention trials in the physical attention condition, Gretchen pushed the bottle of lotion away when Elaine attempted to initiate this activity. Especially during the acquisition phase of requesting via picture exchanges, it is critical to ensure an EO is in effect for the item or activity a student is being taught to

request. Intermittent preference assessments (e.g., daily, between trials) may have prevented practicing requesting an item that was, at least momentarily, non-preferred. In addition, a systematic evaluation of Sid's preferences for various instructional activities may have been helpful in designing his FA trials. It was not until we changed the instructional task from sorting colors to reading a book that we confirmed the escape + tangible hypothesis (which also suggests this behavior-environment relation would not generalize across instructional activities).

In relation to the descriptive assessment data, observations were conducted prior to the trial-based FAs to identify hypotheses for behavior-environment reinforcement contingencies in the classroom. Overall, the patterns described by these data did assist in selecting and designing experimental trials, but we also relied on information provided by TAs as well as our own qualitative observations of students in their usual instructional settings. What might be described generally as *clinical judgment* influenced how FA and intervention trials were designed. We considered these various sources of information in an attempt to increase the ecological validity of experimental trials. We did not, however, operationally define *clinical judgment* nor follow a systematic procedure of incorporating non-quantitative information, which may be an important aspect of future studies. Another limitation of the descriptive data relates to inconsistencies between the stimuli coded during descriptive assessments and those programmed during experimental trials in a subset of cases. For example, Gretchen's descriptive data on teacher attention revealed very high levels of attention overall in addition to immediately surrounding problem behavior. Although attention was included as a test condition to rule out this hypothesis, we changed the form of attention from verbal to physical, and identified a reinforcement effect. Had we compared data on physical versus verbal attention during descriptive assessments, a clearer pattern may have been revealed.

## **Future Research**

Results of the current study highlight several avenues for future research on trial-based FAs and subsequent interventions, several of which relate to the strengths and limitations discussed above. Consensus is yet to be reached on the sequence of test and control intervals within experimental trials. Although the first trial-based FA studies used the test-control sequence (Sigafos & Sagers, 1995; Sigafos & Meikle, 1996), researchers more recently have suggested the sequence be reversed to avoid problem behavior carrying over from the test to control interval (Bloom et al., 2011; Lambert et al., 2012). Results of the current study show minimal evidence of this form of sequence effect, with the possible exception of the escape trials of Gretchen's FA. In addition, it seems that in order to test response-reinforcer relations, the response-reinforcer contingency (i.e., shifting from test interval to control interval contingent on problem behavior) should remain intact. If, however, results of test-control trials were found to correspond with results of control-test trials, the latter sequence may be a more socially acceptable methodology, as it avoids the temporary reinforcement of problem behavior during the experimental assessment phase. Bloom et al. (2011) conducted 6-min control-test-control trials but reported results from the first control-test intervals only after identifying more instances of problem behavior in the second control interval than the first control interval for a subset of participants. Although these secondary analyses were not reported, Bloom et al. seem to suggest that analyzing the control-test intervals and test-control intervals separately would have produced similar outcomes, with the control-test intervals showing slightly clearer response differentiation. An alternative method of comparing the two procedures in a future study would be to intersperse test-control and control-test trials within each condition and evaluate whether response differentiation is observed to the same extent across the two procedural variations.

Future research also is needed on the extent to which the trial-based FA procedures may be implemented with precision in more typical classroom settings. The majority of previous studies on this methodology have been conducted in separate schools or centers for students with disabilities (Bloom et al., 2011; LaRue et al., 2010; Sigafoos & Meikle, 1996; Sigafoos & Sagers, 1995). Although the current study extended these procedures to public elementary school settings, the contexts in which trials were conducted did not resemble typical classroom settings (with the exception of Davis's trials, which were conducted in general education PE and Art classes). Strategies for maintaining adequate levels of procedural fidelity in less structured environments should be identified before these procedures may be recommended across various types of school settings. On a related note, systematic evaluations of the training procedures required for classroom staff to implement experimental trials independently and collect data on problem behavior also may be needed. In the current study, data collection during experimental trials was somewhat complex as we collected data on both student behavior and TA behavior to assess procedural fidelity across all experimental trials. Collecting data on student behavior, however, would simply involve indicating whether problem or replacement behavior occurred in the first minute (test interval) versus the second minute (control interval), which may be a reasonable expectation for teaching staff conducting trials. In addition, if teaching staff were trained to reliably record the presence versus absence of problem behavior, this training likely would obviate the need for researcher prompts to *switch* from test intervals to control intervals.

In the current study, descriptive assessments were conducted prior to the trial-based FA only, such that we did not continue collecting descriptive data on behavior-environment patterns throughout the experimental phases. In future studies, however, it may be informative to collect descriptive data on the specific environmental stimuli being evaluated in trials as a measure of

generalization. That is, once response-reinforcer relations have been identified experimentally, these relations may be described outside of the experimental trials to determine whether student and/or TA behavior patterns change outside of trials throughout assessment and/or intervention phases. The methods of data collection and analysis used during descriptive assessments also may be evaluated in future research. Interestingly, roughly similar base rates of stimulus events were identified from the descriptive assessment data across participants. Specifically, attention was typically delivered in over 70% of intervals, instructions were delivered in 40-70% of intervals, and tangibles were restricted and delivered in less than 20% of intervals. Whether these patterns reflect similar school environments across participants (at least during instructional periods with 1:1 support), or whether these patterns were a reflection of potential oversensitivity or under-sensitivity of the data collection procedures, is unclear and may be addressed empirically. To better understand the impact of data collection techniques on describing behavior-environment patterns in classrooms, different techniques (e.g., varying durations of observe intervals) may be applied and compared within participants.

The trial-based FA, as it was conducted in this study, highlighted a tension between experimental control and ecological validity, which may represent another important avenue for additional research. In designing FA trials for each participant, there is a clear tradeoff between isolating environmental variables and maintaining similarity between experimental trials and events the student typically encounters in that setting. For example, Abhi's escape + tangible trial did not allow us to isolate the positive and negative reinforcement effects, but combining these variables into one trial did allow us to better understand how restriction and access to computer breaks (a combination of events experienced regularly in this setting) influenced Abhi's behavior. That is, we could have chosen to design trials that tested the effects of escape from

instruction only (break with no access to preferred activities) and access to tangibles only but if breaks were never delivered without preferred activities, the utility of these results would be unclear.

Similarly, although efforts were made to incorporate naturally-occurring events into trials, the extent to which results extend to other similar (but not identical) contexts is unknown and could be explored in future research. For example, we used a familiar and unfamiliar set of cards to test the escape from novel instruction hypothesis for Sid. Ultimately, however, we can only conclude a functional relation between Sid's behavior and each card set, which was intended to represent familiarity vs. novelty. Programming additional trials with different materials (e.g., familiar book vs. unfamiliar book) would have demonstrated more convincing evidence that it was the *familiarity* with the instructional materials that was responsible for the different patterns of behavior. In addition, distributing these multiple exemplars of each test condition across the school day (rather than in the context of a limited set of activities within a 1-hr time period) may extend the generality of results. For example, Davis's problem behavior was reported to occur at the highest rates when Darcy was supporting him, and the results of his trial-based FA confirmed that his behavior was highly sensitive to Darcy's attention. Had we conducted the same trials with different adults, however, different functions may have been revealed, if problem behavior occurred at all. Though it would require additional sets of trials to be conducted, identifying boundaries of behavior-environment relations (i.e., contexts in which some functions are relevant but others are not) could facilitate designing comprehensive interventions that could be implemented across the school day.

Finally, future refinements of the trial-based FA methodology may be needed for students who engage in lower frequencies of problem behavior that, in some cases, also may be of higher

intensity. For students who engage in low frequencies of problem behavior, we may not expect 60-s test intervals to be sufficient to evoke these behaviors. For students who engage in high-intensity problem behavior (e.g., tantrums that occur for extended periods of time), we may not expect problem behavior to stop occurring upon the momentary transition from a test to control interval. Future research is needed to evaluate the extent to which a trial-based methodology can be extended to different patterns of problem behavior, or whether this methodology is uniquely suited to high-frequency behaviors.

### **Conclusion**

Our results contribute to the validation of trial-based FAs to identify functions of problem behavior within classroom settings. This variation of traditional FA methodology may represent a starting point in the identification of rigorous yet pragmatic procedures to assess and treat problem behavior in the context of students' usual educational settings. Further research is needed to address challenges unique to trial-based FAs, which include (a) maintaining adequate levels of procedural fidelity and (b) evaluating the extent to which response-reinforcer relations generalize across daily activities and routines to (c) develop strategies to translate results of brief test-control trials to comprehensive behavioral interventions.

# APPENDIX A

## DESCRIPTIVE ASSESSMENT DATA COLLECTION FORM

<u>Date:</u>			
<u>Participant ID:</u>	TA: teacher attention	TD: tangible delivered	ID: instruction delivered (/=compliance)
<u>Observation #:</u>	PA: peer attention	TR: tangible removed	IR: instruction removed
<u>Observers:</u>	DA: diverted attention	NT: no tangible	NI: no instruction
<u>Setting:</u>			
<u>Activity:</u>			

INTERVAL	Attention			Tangible			Instruction			Notes	Target Bx		Attention			Tangible			Instruction			Notes
1	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
2	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
3	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
4	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
5	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
6	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
7	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
8	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
9	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
10	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
11	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
12	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
13	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
14	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
15	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
16	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
17	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
18	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
19	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	
20	TA	PA	DA	TD	TR	NT	ID	IR	NI		+	-	TA	PA	DA	TD	TR	NT	ID	IR	NI	



## APPENDIX B

### TBFA FIDELITY DATA COLLECTION FORM

Participant ID: \_\_\_\_\_

Teacher ID: \_\_\_\_\_

Observers: \_\_\_\_\_

Target Behavior(s): \_\_\_\_\_

Instructions: For the test interval and control interval of each FA trial, circle the relevant stimuli and indicate whether target behavior did (+) or did not (-) occur. If target behavior occurs during the test interval, write in the latency to target behavior (seconds) and circle this number if the teacher correctly delivers the relevant consequence within 5 s of target behavior.

DATE	TRIAL #	TRIAL TYPE	TEST INTERVAL					CONTROL INTERVAL					ITI (min)
			ATTN	TAN	INSTR	TARGET BX	LAT	ATTN	TAN	INSTR	TARGET BX		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		
			TA PA DA	TR TD NT	ID IR NI	+ -		TA PA DA	TR TD NT	ID IR NI	+ -		

# APPENDIX C

## INTERVENTION DATA COLLECTION FORM

Participant ID: \_\_\_\_\_  
 Teacher ID: \_\_\_\_\_

Target Behavior(s): \_\_\_\_\_  
 Replacement Behavior(s): \_\_\_\_\_

DATE	TRIAL #	TRIAL TYPE	TEST INTERVAL										CONTROL INTERVAL																											
			ATTN			TAN			INSTR				TB		RB		LAT	ATTN			TAN			INSTR				TB		RB		ITI								
			TA	PA	DA	TR	TD	n/a	ID	IR	NI	+	-	I	P	-		TA	PA	DA	TR	TD	n/a	ID	IR	NI	+	-	I	P	-									

## REFERENCES

- Anderson, C. M., & Long, E. S. (2002). Use of a structured descriptive assessment methodology to identify variables affecting problem behavior. *Journal of Applied Behavior Analysis, 35*, 137-154. doi: 10.1901/jaba.2002.35-137
- Axelrod, S. (1987). Functional and structural analyses of behavior: Approaches leading to reduced use of punishment procedures? *Research in Developmental Disabilities, 8*, 165-178. doi: 10.1016/0891-4222(87)90001-1
- Birnbrauer, J. S., Peterson, C. R., & Solnick, J. V. (1974). Design and interpretation of studies of single subjects. *American Journal of Mental Deficiency, 79*, 191-203.
- Bloom, S. E., Iwata, B. A., Fritz, J. N., Roscoe, E. M., & Carreau, A. B. (2011). Classroom application of a trial-based functional analysis. *Journal of Applied Behavior Analysis, 44*, 19-31. doi: 10.1901/jaba.2011.44-19
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.
- Carr, E. G., Yarbrough, S. C., & Langdon, N. A. (1997). Effects of idiosyncratic stimulus variables on functional analysis outcomes. *Journal of Applied Behavior Analysis, 30*, 673-686. doi: 10.1901/jaba.1997.30-673
- Conroy, M. A., Fox, J. J., Crain, L., Jenkins, A., & Belcher, K. (1996). Evaluating the social and ecological validity of analog assessment procedures for challenging behaviors in young children. *Education and Treatment of Children, 19*, 233-256.

- Desrochers, M. N., Hile, M. G., & Williams-Moseley, T. L. (1997). Survey of functional assessment procedures used with individuals who display mental retardation and severe problem behaviors. *American Journal on Mental Retardation, 101*, 535-546.
- Doss, S., & Reichle, J. (1989). Establishing communicative alternatives to the emission of socially motivated excess behavior: A review. *Journal of the Association of Persons with Severe Handicaps, 14*, 101-112.
- Ellis, J., & Magee, S. K. (1999). Determination of environmental correlates of disruptive classroom behavior: Integration of functional analyses into public school assessment process. *Education and Treatment of Children, 22*, 291-316.
- English, C. L., & Anderson, C. M. (2006). Evaluation of the treatment utility of the analog functional analysis and the structured descriptive assessment. *Journal of Positive Behavior Interventions, 8*, 212-229. doi: 10.1177/10983007060080040401
- Ervin, R. A., Radford, P. M., Bertsch, K., Piper, A. L., Ehrhardt, K. E., & Poling, A. (2001). A descriptive analysis and critique of the empirical literature on school-based functional assessment. *School Psychology Review, 30*, 193-210.
- Gast, D. L. (2010). *Single subject research in behavioral sciences*. New York: Routledge.
- Hains, A. H., & Baer, D. M. (1989). Interaction effects in multielement designs: Inevitable, desirable, and ignorable. *Journal of Applied Behavior Analysis, 22*, 57-69. doi: 10.1901/jaba.1989.22-57
- Individuals with Disabilities Education Act of 1990, 20 U.S.C., § 1400 et seq.
- Individuals with Disabilities Education Act Amendments of 1997, 20 U.S.C., § 1400 et seq.
- 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)

doi: 10.1901/jaba.1994.27-197

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J. R., Vollmer, T. R., Smith, R. G., & Willis, K. D. (1994). The functions of self-injurious behavior: an experimental epidemiological analysis. *Journal of Applied Behavior Analysis*, 27, 215-240. doi: 10.1901/jaba.1994.27-215

Iwata, B. A., Pace, G. M., Kalsher, M. J., Cowdery, G. E., & Cataldo, M. F. (1990). Experimental analysis and extinction of self-injurious escape behavior. *Journal of Applied Behavior Analysis*, 23, 11-27. doi: 10.1901/jaba.1990.23-11

Keller, F. S., & Schoenfield, W. N. (1950). *Principles of psychology*. New York: Appleton-Century-Crofts.

Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston: Allyn & Bacon.

Lalli, J. S., Browder, D. M., Mace, F. C., & Brown, D. K. (1993). Teacher use of descriptive analysis data to implement interventions to decrease students' problem behavior. *Journal of Applied Behavior Analysis*, 26, 227-238. doi: 10.1901/jaba.1993.26-227

Lambert, J. M., Bloom, S. E., & Irvin, J. (2012). Trial-based functional analysis and functional communication training in an early childhood setting. *Journal of Applied Behavior Analysis*, 45, 579-584. doi: 10.1901/jaba.2012.45-579

Lang, R., O'Reilly, M., Machalicek, W., Lancioni, G., Rispoli, M., & Chan, J. M. (2008). A preliminary comparison of functional analysis results when conducted in contrived versus natural settings. *Journal of Applied Behavior Analysis*, 41, 441-445. doi: 10.1901/jaba.2008.41-441

- Lang, R., O'Reilly, M., Lancioni, G., Rispoli, M., Machalicek, W., Chan, J. M., . . . Franco, J. (2009). Discrepancy in functional analysis results across two settings: Implications for intervention design. *Journal of Applied Behavior Analysis, 42*, 393-397. doi: 10.1901/jaba.2009.42-393
- Laraway, S., Snyckerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis, 36*, 407-414. doi: 10.1901/jaba.2003.36-407
- LaRue, R. H., Lenard, K., Weiss, M. J., Bamond, M., Palmieri, M., & Kelley, M. E. (2010). Comparison of traditional and trial-based methodologies for functional analyses. *Research in Developmental Disabilities, 31*, 480-487. doi: 10.1016/j.ridd.2009.10.020
- Lloyd, B. P. (in preparation). A systematic review of experimental functional assessment methods conducted in classroom settings.
- Lydon, S., Healy, O., O'Reilly, M. F., & Lang, R. (2012). Variations in functional analysis methodology: A systematic review. *Journal of Developmental and Physical Disabilities, 24*, 301-326. doi: 10.1007/s10882-012-9267-3
- Mace, F. C., Page, T. J., Ivancic, M. T., & O'Brien, S. (1986). Analysis of environmental determinants of aggression and disruption in mentally retarded children. *Applied Research in Mental Retardation, 7*, 203-221. doi: 10.1016/0270-3092(86)90006-8
- Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology: Research and Practice, 16*, 191-198.

- Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior*, 37, 149-155. doi:10.1901/jeab.1982.37-149
- Najdowski, A. C., Wallace, M. D., Ellsworth, C. L., MacAleese, A. N., & Cleveland, J. M. (2008). Functional analyses and treatment of precursor behavior. *Journal of Applied Behavior Analysis*, 41, 97-105. doi: 10.1901/jaba.2008.41-97
- 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* (2<sup>nd</sup> ed.). Pacific Grove, CA: Brooks/Cole.
- O'Neill, R. E., Horner, R. H., Albin, R. W., Storey, K., & Sprague, J. R. (1990). *Functional analysis of problem behavior: A practical assessment guide*. Sycamore, IL: Sycamore Publishing Company.
- Repp, A. C. (1994). Comments on functional analysis procedures for school-based behavior problems. *Journal of Applied Behavior Analysis*, 27, 409-411. doi: 10.1901/jaba.1994.27-409
- Ringdahl, J. E., & Sellers, J. A. (2000). The effects of different adults as therapists during functional analyses. *Journal of Applied Behavior Analysis*, 33, 247-250. doi: 10.1901/jaba.2000.33-247
- Sigafoos, J., & Meikle, B. (1996). Functional communication training for the treatment of multiply determined challenging behavior in two boys with autism. *Behavior Modification*, 20, 60-84. doi: 10.1177/01454455960201003

- Sigafoos, J., & Sagers, E. (1995). A discrete-trial approach to the functional analysis of aggressive behaviour in two boys with autism. *Australia & New Zealand Journal of Developmental Disabilities, 20*, 287-297.
- Solnick, M. D., & Ardoin, S. P. (2010). A quantitative review of functional analysis procedures in public school settings. *Education and Treatment of Children, 33*, 153-175.
- Tincani, M. J., Castrogiovanni, A., & Axelrod, S. (1999). A comparison of the effectiveness of brief versus traditional functional analyses. *Research in Developmental Disabilities, 20*, 327-338. doi: 10.1016/S0891-4222(99)00014-1
- Ulman, J. D., & Sulzer-Azaroff, B. (1975). Multi-element baseline design in educational research. In E. Ramp & G. Semb (Eds.), *Behavior analysis: Areas of research and application* (pp. 377-391). Englewood Cliffs, NJ: Prentice-Hall.
- U.S. Department of Education. (1995). *To assure the free appropriate public education of all children with disabilities. Individuals with Disabilities Act, Section 618*. Seventeenth Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act.
- Wallace, M. D., & Knights, D. J. (2003). An evaluation of a brief functional analysis format within a vocational setting. *Journal of Applied Behavior Analysis, 36*, 125-128. doi: 10.1901 /jaba.2003.36-125