Journal of
Applied Behavior Analysis

JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2020, **53,** 25–43 NUMBER 1 (WINTER)

Evaluating the boundaries of analytic efficiency and control: A consecutive controlled case series of 26 functional analyses

Joshua Jessel

QUEENS COLLEGE

Rachel Metras

UNIVERSITY OF NORTH TEXAS

Gregory P. Hanley

WESTERN NEW ENGLAND UNIVERSITY

Catherine Jessel

UNIVERSITY OF NORTH TEXAS

Einar T. Ingvarsson

UNIVERSITY OF NORTH TEXAS & CHILD STUDY CENTER

We conducted this study to determine if the efficiency of the functional analysis could be improved without detrimental effects on control. In Experiment 1, we reanalyzed functional analyses conducted for the problem behavior of 18 children. We analyzed rates of problem behavior during the first 5 min and first 3 min of the original 10-min sessions and evaluated if changes in the level of control over problem behavior by the programmed contingency were evident from the analyses of shorter session duration. In Experiment 2, we conducted 8 consecutive functional analyses with 3-min sessions to further evaluate the utility of brief session durations. We found that control over problem behavior was demonstrated when conducting functional analyses with sessions as brief as 3 min.

Key words: analytic control, efficiency, functional analysis, problem behavior

Functional analysis was designed to improve the treatment of severe problem behavior by first demonstrating control over problem behavior by the suspected maintaining environmental variables (Hanley, Iwata, & McCord, 2003). Control during a functional analysis is demonstrated when problem behavior is reliably evoked during a condition of contingent reinforcement and eliminated when those same reinforcers are provided noncontingently in the control condition (Hanley, 2012). Researchers have evaluated the results of behavior analytic treatments in multiple quantitative reviews and have found that an overall greater reduction in problem behavior was achieved when treatment procedures were informed by a functional analysis (Campbell, 2003; Heyvaert, Saenen, Campbell, Maes, & Onghena, 2014; Kahng, Iwata, & Lewin, 2002).

In spite of this empirical support for a pretreatment functional analysis, as well as the numerous replications in the literature (Beavers, Iwata, & Lerman, 2013; Hanley et al., 2003), professional behavior analysts have reported using less effective functional

Rachel Metras is now at Western New England University. Catherine Jessel is now at Long Island ABA. Einar T. Ingvarsson is now at the Virginia Institute for Autism.

We would like to thank Holly Gover, Laura Hanratty, Robin Landa, Adithyan Rajaraman, Kelsey Ruppel, and Jessica Slaton for their expert contributions to the development of this manuscript.

Address correspondence to: Joshua Jessel, Queens College, Department of Psychology, 65-30 Kissena Blvd., Queens, NY 11367. Email: joshua.jessel@qc.cuny.edu

doi: 10.1002/jaba.544

assessment tools in their practice (Oliver, Pratt, & Normand, 2015; Roscoe, Phillips, Kelley, Farber, & Dube, 2015). Oliver et al. (2015) retrieved online surveys from 682 practicing Board Certified Behavior Analysts (BCBAs) and found that 90% of responders indicated that they routinely used various functional assessments. However, those functional assessments were likely to be indirect and descriptive, with responders reporting their use 71% and 83% of the time, respectively. This and other similar reports (e.g., Love, Carr, Almason, & Petursdottir, 2009; Roscoe et al., 2015) suggest that the functional analysis has paradoxically become a widely researched tool for identifying the environmental determinants of problem behaviors (Beavers et al., 2013; Hanley et al., 2003), while being sparsely used by practitioners and clinicians.

Oliver et al. (2015) attempted to identify the reason for this gap between research and practice by asking responders to identify barriers to the use of functional analysis. By and large, responders reported that lack of time for conducting functional analyses served as their biggest barrier. The time required to conduct a functional analysis (i.e., analytic efficiency) has been commonly recognized as a limitation of the functional analysis (e.g., Cooper, Heron, & Heward, 2007; Repp, Singh, Olinger, & Olson, 1990), with some reporting functional analysis to require a mean of 6.5 hr and a range of up to 16.5 hr to conduct (Iwata et al., 1994).

Wallace and Iwata (1999) reanalyzed 46 functional analyses conducted with 15-min sessions. The results of the first 5 and 10 min of the 15-min sessions were evaluated separately to gauge the level of correspondence in identified functions between shorter and longer sessions. A group of six doctoral students with experience conducting and interpreting functional analyses then made judgements of control without knowing the session duration in each instance. Identification of function was not negatively influenced by looking at only the first 10 minutes of each session. With 5-min sessions, lack of correspondence occurred in only three of the 46 analyses. Thus, the authors suggested that sessions might be as short as 5 min without detrimental effects on the interpretations of functional control.

Interpretation of functional control in a single-subject design requires visual inspection of level, variability, and trends with repeated measures (Bourret & Pietras, 2013). Multiple models have been developed to aid in the visual analysis process (e.g., Fisher, Kelley, & Lomas, 2003; Hagopian et al., 1997; Roane, Fisher, Kelley, Mevers, & Bouxsein, 2013), often focusing on the interpretation of the control within functional analyses. Roane et al. (2013) modified previously developed structured criteria (Hagopian et al., 1997) to improve the interpretation of functional analyses of varying lengths. The participants ranged from postbaccalaureate to postdoctoral behavior analysts, and their binary (yes or no) interpretations of functional control were compared to those of expert judges. The structured criteria for visual inspection involved calculating two criterion lines: The upper criterion line was set at one standard deviation above the mean rate of problem behavior during the control condition and the lower criterion line was set at one standard deviation below the mean rate. The functional analysis sessions were 10 min and each functional analysis was determined to have control if more than half of the sessions in the test condition fell above the upper criterion line rather than below the lower criterion line. Roane et al. found that agreement coefficients between the participants and the experts were only above .9 when the participants used the structured criteria as opposed to subjective judgments. Evaluations of the possible variability imposed by shorter session durations may be enhanced by incorporating structured criteria in addition to judgements made by an expert panel.

Current structured criteria for visual analysis might not be sensitive to fluctuations in control introduced by relatively shorter functional analysis sessions, especially when evaluating analyses that tend to involve fewer data points (e.g., two to three data points per condition as in Hanley, Jin, Vanselow, & Hanratty, 2014; Jessel, Hanley, & Ghaemmaghami, 2016). With shorter sessions (e.g., 5 min) in brief analyses, experimental control could degrade due to the proportionally greater effect of variability, data overlap, and suboptimal trends. Nonparametric statistics have been developed for describing the difference between baseline and treatment conditions (i.e., effect sizes) for single-subject design research, but have not been used to aid in the interpretation of functional analysis outcomes. For example, the percentage of nonoverlapping data (PND; Scruggs, Mastropieri, & Casto, 1987) is a conservative approach (Carr, 2015) that identifies the number of points during treatment that do not overlap with the highest baseline point and divides that sum of nonoverlapping points by the total number of treatment points to provide an effect size between 0 and 100%. Effect sizes calculated using simple statistics like PND have been found to be useful in treatment efficacy reviews (e.g., Campbell, 2003; Carr, Severtson, & Lepper, 2009; Heyvaert et al., 2014); however, it is unknown if PND effect sizes between the test and control conditions of a functional analysis would correspond to, or be more or less sensitive to, a structured criteria or visual analysis of functional analysis data.

We attempted to extend the literature on improving analytic efficiency by reanalyzing the data from 18 functional analyses of varying session duration. Previous research (Wallace & Iwata, 1999) included a minimum session duration of 5 min extracted from traditional functional analyses (Iwata et al., 1994) that were originally conducted using 15-min sessions. In the current study, we compared the interpretations of functional control resulting from original 10-min functional analysis sessions to interpretations resulting from the first 3 and 5 min of the original 10-min sessions. Furthermore, the session duration evaluation was conducted with a recently described type of functional analysis (Hanley et al., 2014), referred to as an interview-informed, synthesized contingency analysis (IISCA; Jessel et al., 2016). In addition to the binary judgements of trained panelists, we incorporated structured criteria (Hagopian et al., 1997; Roane et al., 2013) to produce a second set of binary judgements. Because improvements in efficiency may result in degraded control rather than the absolute loss of functional control (Jessel et al., 2016), we also describe and applied a multilevel system involving differing degrees of control (i.e., none, weak, moderate, strong), and applied the PND statistic to the same 18 analyses. In Experiment 2, we conducted eight additional functional analyses with 3-min sessions to determine if the results of the reanalysis were consistent with outcomes of functional analyses relying on the briefest session duration.

EXPERIMENT 1: REANALYSIS OF 18 INTERVIEW-INFORMED SYNTHESIZED CONTINGENCY ANALYSES

Because the purpose of this experiment was to evaluate improvements in efficiency, we conducted a functional analysis specifically designed to be quick and practical. Hanley et al. (2014) described a relatively brief functional analysis of problem behavior format with three children diagnosed with autism. The analysis included a single test condition with synthesized establishing operations and synthesized reinforcement contingencies (e.g., escape to tangibles) designed to simulate typically occurring antecedents and consequences of problem behavior as reported by caregivers. A matched control condition, in which the same reinforcers were freely available, alternated with the test sessions. In a study describing 30 replications of this type of analysis, Jessel et al. (2016) reported that the mean analysis duration was 25 min, with some analyses involving sessions as short as 3 min. In Experiment 1, we describe the results of 18 IISCAs that initially relied on 10-min sessions, and then extracted samples from the beginning of each session so that outcomes from relatively shorter sessions could be reanalyzed.

Method

Participants and settings. All participants received assessment and treatment services for their problem behavior in a 2-week outpatient program supervised by the first and second authors of this study. The data from four of the participants' (Ari, Dace, Annie, and Joe) functional analyses were previously published in Jessel, Ingvarsson, Kirk, Whipple, and Metras (2018a). Their data were included in the current reanalysis because their IISCAs were conducted with 10-min sessions. A fifth participant (Aaron) for whom 10-min sessions were also conducted in Jessel et al. was not included in the current experiment because the delay from the evocative situation (i.e., transitioning to and from different rooms) during the test condition of his functional analysis precluded the type of reanalysis applied in the current experiment. The remaining participants were the next 14 consecutive clients admitted to a clinic for this consecutively controlled case-series (Rooker, Jessel, Kurtz, & Hagopian, 2013). The participant characteristics are presented in Table 1. The median age was 7.5 years and the majority was male (16 males, 2 females). Sixteen of the participants had a diagnosis of autism, with many having additional diagnoses of intellectual disability and/or attention deficit hyperactivity disorder (ADHD). The language abilities of the participants ranged from nonvocal to fluent, with the mode being nonverbal. Sessions were conducted by trained therapists in 3-m x 6-m and 3-m x 4-m treatment rooms.

Response measurement. Observers scored multiple topographies of problem behavior, which varied across participants (see Table 1). Common forms of problem behavior displayed included aggression (e.g., hitting, kicking, scratching, biting others), self-injurious behavior (e.g., hitting, scratching, biting self), disruption (e.g., tearing, throwing, hitting items), tantrums (e.g., dropping and crying or whining for more than 30 s), and loud vocalizations (e.g., yelling, screaming, swearing). Less common forms included inappropriate sexual behavior (ISB) and elopement. All problem behavior was measured as responses per minute and calculated by dividing the total frequency by the duration of the session. Duration of access to reinforcers was also measured. An observer scored the onset (offset) of reinforcement when (a) the antecedent variables intended to evoke problem behavior were removed (presented), (b) the consequent variables were presented (removed), or (c) both (a) and (b) in cases in which positive and negative reinforcement was synthesized.

Interobserver and interrater agreement. Two observers independently scored live sessions or videotaped sessions for each IISCA. The second observer scored a mean of 40% of the sessions (range, 20% to 67%). We segmented each session into 10 s intervals and calculated percentage of agreement by dividing the smaller value by the larger value in each interval, adding the proportions for all intervals together, dividing by the total number of intervals in each session, and multiplying by 100. The mean IOA for problem behavior and reinforcer access across the 18 analyses was 99% (range, 95% to 100%) and 97% (range, 87% to 100%), respectively.

To evaluate reliability of the evaluations of functional control, a second rater used the structured criteria developed by Hagopian et al.

Table 1
Characteristics of Experiment 1 Participants

Participant	Participant Characteristics				
	Age	Sex	Diagnosis	Language Ability*	Problem Behavior
Andy	11	M	ASD, ID	1	Agg, dis, tantrums
Max	7	M	ASD, ADHD	4	Agg, dis, SIB, tantrums
Gene	10	M	ASD	1	Agg, dis, SIB, loud voc, tantrums
Rina	2	F	DS	2	Agg, SIB, tantrums
Eli	5	M	ASD	3	Agg, loud voc, elope, tantrums
Nick	3	M	ASD, ID	1	Agg, dis, loud voc, tantrums
Ari	15	M	ASD, ID, ADHD	4	ISB
Jim	8	M	ASD	1	Agg
Koa	2	M	No diagnosis	2	Agg, dis, loud voc
Jiro	9	M	ASD, ADHD	4	Agg, dis, loud voc, elope
Jin	8	M	ASD, ID, ADHD	1	Agg, dis, SIB, disrobe, elope
Corey	9	M	ASD, ID	2	Agg, dis, loud voc
Annie	5	F	ASD, ADHD, ID	3	Tantrums
Tiff	6	M	ASD	1	Agg
Dace	11	M	ASD, ID	2	Agg, dis, SIB, tantrums
Job	5	M	ASD, GAD	2	Dis, SIB, tantrums
Smith	15	M	ASD	4	Agg, dis, loud voc
Joe	5	M	ASD	1	Agg, dis, tantrums

Note. ASD is autism spectrum disorder. ID is intellectual disability. ADHD is attention deficit/hyperactivity disorder. GAD is generalized anxiety disorder. DS is Down syndrome. SIB is self-injurious behavior. ISB is inappropriate sexual behavior.
* 1 = non-verbal; 2 = 1-word utterances; 3 = short diffluent sentences; 4 = full fluency.

(1997) and modified by Roane et al. (2013), and independently rated all of the 54 analyses (18 analyses at three different durations each). An agreement was scored if the primary and secondary raters both scored an analysis as having or not having control. Otherwise, a disagreement was scored. The agreements were summed and divided by the total number of analyses evaluated by both raters; IOA was 100%.

The secondary rater also scored levels of control using the multileveled criteria described below for all 54 analyses. An agreement was scored if the level of control (i.e., none, weak, moderate, strong) for an analysis recorded by the primary rater matched that of the secondary rater and a disagreement was scored if the recordings did not match. The number of agreements was divided by the total number of scorings, and IOA was 100%.

Experimental design. The test and control conditions of the IISCAs were rapidly alternated (i.e., multielement design). The usual number of sessions was five (i.e., control, test, control, test, test), but additional sessions were conducted when differentiation was not evident from the five-session analysis.

Procedure. The functional assessment procedures were similar to those described by Hanley et al. (2014) and involved an interview, brief informal contingency probe, and functional analysis. The functional assessment process was designed to discover and then evaluate a single, contextually-relevant reinforcement contingency for each participant (i.e., evaluating sensitivity to distinct classes of reinforcement was not the aim of the functional assessment process).

First, the therapist conducted an open-ended interview with the caregivers to identify problem behaviors that co-occurred with the problem behavior of primary concern, and to identify the antecedent and consequent events that seemed relevant to the problem behavior (see appendix in Hanley, 2012, for the interview). The therapist used the questions from the interview as a guide and asked only questions in each section until enough categorical information was obtained for the therapist to (a) operationally define problem behavior and (b) design the test condition of the functional analysis. No modifications were made to the open-ended interview; however, the therapist did ask follow-up questions to have the caregivers expand on an answer as needed. The caregivers were asked questions regarding the specific topographies of problem behavior, which problem behavior they found most concerning, and which problem behaviors were likely to co-occur. The therapist used the responses to identify the topographies of problem behavior that were of utmost concern and those that appeared to be functionally related, but less severe.

The information obtained from caregivers during the interview regarding the antecedent and consequent events was used to develop a unique contingency reflecting the situations in which problem behavior was reportedly likely to occur for each participant. In other words, general classes of reinforcement were not evaluated and only a single test condition of a synthesized contingency implicated by the caregivers was included in the subsequent analysis. Had the information provided during the interview implicated multiple contexts of disparate contingencies (e.g., problem behavior maintained by access to interactive play at home and escape from independent work completion at school), multiple IISCAs would have been conducted specifically evaluating each unique behavior–environment relation. However, this was not reported for any of the participants. The interview required 15 to 30 min to conduct.

Second, the therapist conducted a brief contingency probe in the reportedly problematic context. Data were not collected during this time. The contingency probe sometimes helped to refine the operational definitions of problem behavior and sometimes enhanced understanding of the situations that evoked problem behavior. For example, Rina's caregivers reported that Rina enjoyed playing with others with particular leisure activities. However, if the playmate were to tell Rina that playtime was finished and to come with them to a different location, Rina would exhibit problem behavior. The experimenter would then have the caregivers watch while the therapist allowed Rina to choose an activity and played with her, and periodically told her the "toys were all finished" and prompted her to leave the area. If problem behavior were to have occurred, the transition would have been discontinued and the interactive play reinstated. When problem behavior did not occur during those situations, the caregivers were asked for additional information to help make the situation more similar to the situations occurring in the home (e.g., suggestions on the specific instructions, suggestions on preferred activities). In other words, the probe was used to calibrate the ecological precision of the contingency before the systematic evaluation during the IISCA. The contingency probe lasted from 5 to 30 min across participants and continued until the therapists felt they had enough information from the interview and contingency probe to conduct the analysis.

Third, the therapist conducted the IISCA. The putative reinforcers were provided for 30-s access following each instance of problem behavior in the test condition and continuously in the control condition. The specific reinforcers included in the test and control conditions for each participant are presented in Table 2. Sixteen of the 18 analyses included the synthesis of positive and negative reinforcement. The other two analyses included a synthesis (Nick) or isolated (Ari) form of only positive reinforcement. In addition, the number of functional analysis iterations required before differentiated results were obtained is included

Table 2
Experiment 1 Functional Analysis Information

Participant	Iteration	Test Condition	Control Condition
Andy	First	Escape from adult-directed play to free play with mom	Continuous free play with mom and no adult directions
Max	First	Escape from academic instructions to independent play	Continuous independent play and no academic instructions
Gene	First	Escape from instructions to independent play	Continuous independent play and no instructions
Rina	First	Escape from transitions to interactive play	Continuous interactive play in the preferred location
Eli	First	Escape from instructions to interactive play	Continuous interactive play and no instructions
Nick	First	Access to interactive play	Continuous interactive play
Ari	First	Access to iPad	Continuous access to iPad
Jim	Second	Escape from diverted attention to interactive play	Continuous interactive play with full attention
Koa	First	Escape from transitions to interactive play with mom	Continuous interactive play with mom in the preferred location
Jiro	First	Escape from caregiver instructions to independent play	Continuous independent play and no caregiver instructions
Jin	First	Escape from gross motor instructions to interactive play	Continuous interactive play and no gross motor instructions
Corey	First	Escape from adult access to independent access of iPad	Continuous independent access of iPad
Annie	First	Escape from adult-directed to child-directed play	Continuous child-directed play and no adult directions
Tiff	First	Escape from blocked access to free access to leisure items	Continuous free access to leisure items
Dace	Second	Escape from instructions to independent play	Continuous independent play and no instructions
Job	Third	Escape from caregiver-directed play to child-directed play	Continuous child-directed play and no caregiver directions
Smith	Second	Escape from academic instructions to independent play	Continuous independent play and no academic instructions
Joe	First	Escape from transitions to iPad	Continuous access to iPad in the preferred location

in Table 2. If the IISCA failed to implicate a socially mediated function, the therapist returned to asking the caregivers more open-ended questions about potential discrepancies between the IISCA and their typical experience. The information was used to modify the contingency probe and functional analysis. Fourteen out of 18 analyses (78%) did not require additional iterations (i.e., redesigned analyses), whereas three participants (17%) required one set of modifications to the analysis, and one required two sets of modifications to the analysis.

Data analysis. We reanalyzed the data from the first 5 and 3 min of each 10-min session to evaluate the impact of shorter sessions on the interpretation of behavioral function. We calculated the rate of problem behavior during the 5-min IISCA reanalysis by only including the problem behavior from the first 5 min of each 10-min session and dividing that total by five. The rate of problem behavior from the 3-min IISCA reanalysis was calculated by only including the problem behavior from the first 3 min of each 10-min session and dividing that total by three. We used four methods to evaluate functional control across the analyses of varying session duration. We also calculated the analysis duration of each application by multiplying the session duration (i.e., 3 min, 5 min, 10 min) by the number of sessions conducted. We considered five sessions of 3-min each, creating an analysis duration of 15 min, to be the most efficient possible application.

Binary panelist criterion. We developed a criterion of agreement between multiple panelists to determine if functional control was demonstrated in a given functional analysis, similar to

Wallace and Iwata (1999). The panel consisted of six BCBAs who were doctoral students with experience conducting research in functional analysis. In addition, the panelists completed training on visual analysis prior to their review of the data. The visual analysis training was not specific to the IISCA and included the review of multiple strategies on evaluating functional control in different experimental designs. The data from each of the 18 IISCAs were depicted in three different ways (i.e., data from 3-min, 5-min, and 10-min sessions). Therefore, the BCBAs conducted a total of 54 evaluations of functional control. The BCBAs were sent an email with a PowerPoint® and were provided with the instructions:

> Each slide will contain the results of an individual functional analysis. Visually inspect each individual data set and indicate in the notes section whether you believe the results support the identification of a functional relation or if you believe the results are ambiguous (yes if functional relation, no if ambiguous). Complete your review independently.

The 54 IISCAs of varying session durations were presented in a mixed order, and there was no indication of session duration. We set the criterion for demonstration of functional control as at least five or six of the ratings indicating yes.

Binary structured criteria. The structured criteria were developed by Hagopian et al. (1997) and modified by Roane et al. (2013). We examined each figure individually and used the rate of problem behavior during the control condition to calculate an upper criterion line (i.e., one standard deviation above the mean) and a lower criterion line (i.e., one standard deviation below the mean). We then counted the number of test sessions that fell above the upper criterion line and the number of test sessions that fell below the lower criterion line. To calculate a percentage, we subtracted the number of data points below the lower criterion line from the number of data points above the upper criterion line and divided the difference by the total number of test sessions. Any functional analysis with a quotient at or above 50% was considered to be differentiated and to have control. If the mean rate of problem behavior during the control condition was 0, the upper and lower criterion lines were set at 0. During these cases, 50% of the test sessions had to be above 0 to be considered differentiated.

Multilevel structured criteria. The multileveled structured criteria were considered to be an extension of the binary structured criteria developed by Hagopian et al. (1997). We categorized the IISCAs according to four levels of control: strong, moderate, weak, and none (see examples in Figure 1). The IISCAs categorized as having *strong* control did not have any overlap in the data across the test and control conditions and did not have any occurrences of problem behavior in the control condition (i.e., immediate, sustained differentiation between the test and control condition). The IISCAs that were categorized as having *moderate* control had some overlap in the data across the test and control conditions *or* some occurrences of problem behavior in the control condition. The IISCAs that were categorized as having *weak* control had some overlap in the data across the test and control conditions *and* some occurrences of problem behavior in the control condition. Lastly, the IISCAs with substantial overlap in data paths or that failed to replicate effects in an experimental design were categorized as having *no* control.

Percentage of non-overlapping points. To calculate PND, we counted the number of data points in the test condition that were above, and did not overlap with, the highest datum in the control condition. This sum was then divided by the total number of test sessions conducted and multiplied by 100 to get a percentage. A PND ranging from 0 to 100% was calculated for each IISCA.

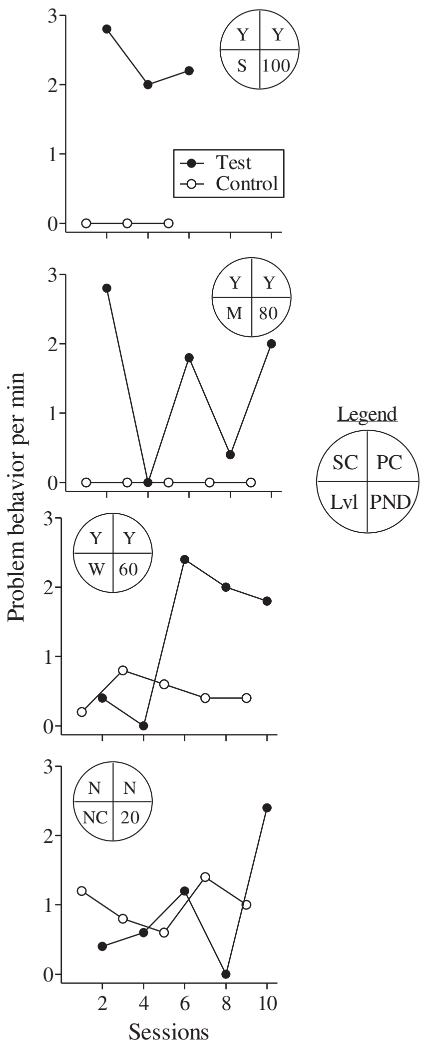

Figure 1. Hypothetical data of functional analyses meeting criteria of strong (no overlap and zero occurrences in the control condition), moderate (some overlap *or* some occurrences in the control condition), weak (some overlap *and* some occurrences in the control), and no control (substantial overlap). SC refers to structured criteria interpretation of control (Y = Yes, N = No). PC refers to panelist criterion. Lvl refers to the level of control (S = strong, M = moderate, W = weak, NC = no control). PND refers to percentage of nonoverlapping data.

Results and Discussion

The results for all 18 participants are presented in Figures 2–4. Elevated rates of problem behavior were observed during the test condition across participants in the 10-min (*M* = 1.2 RPM; *SD* = 0.4), 5-min (*M* = 1.0 RPM; *SD* = 0.5), and 3-min analyses (*M* = 1.0 RPM; *SD* = 0.5). The mean analysis duration of the 10-min, 5-min, and 3-min IISCAs was 56 min, 28 min, and 17 min, respectively.

The results of the binary evaluations of functional control are presented in the top half of the pie charts of each IISCA in the figures. The evaluations completed using the panelist criterion (top right quadrant of pie charts) resulted in an overall high percentage of IISCAs designated as demonstrating functional control; however, there was a slight decrease in demonstration of control when the analyses included briefer sessions. All 18 IISCAs with the full 10-min sessions were identified as having control, whereas 94% and 83% of the 5-min and 3-min IISCAs, respectively, met the panelist criterion.

The evaluations completed using the binary structured criteria are presented in the top left quadrant of the pie charts. The binary structured criteria resulted in 100% of the analyses being designated as demonstrating functional control, regardless of session duration.

The results of the multilevel evaluation of control are presented in the bottom half of the pie charts in each IISCA. The majority of the IISCAs were determined to have *strong* control, with the percentage decreasing across the 10-min (89%), 5-min (78%), and 3-min analyses (67%). Similar percentage of analyses with *moderate* control were obtained from the 10-min (11%) to the 5-min (11%) and 3-min (17%) analyses. There was only one IISCA that had *weak* control (Tiff, 5-min sessions). The shorter durations also resulted in one 5-min IISCA (6%) and three 3-min IISCAs (17%) being designated as having *no control*.

Similar patterns in PND effect sizes (bottom right quadrant of pie charts) were obtained with minimal overlap in the test and control conditions during the 10-min analyses (*M* = 98%; *SD* = 6.8) and a decrease in PND

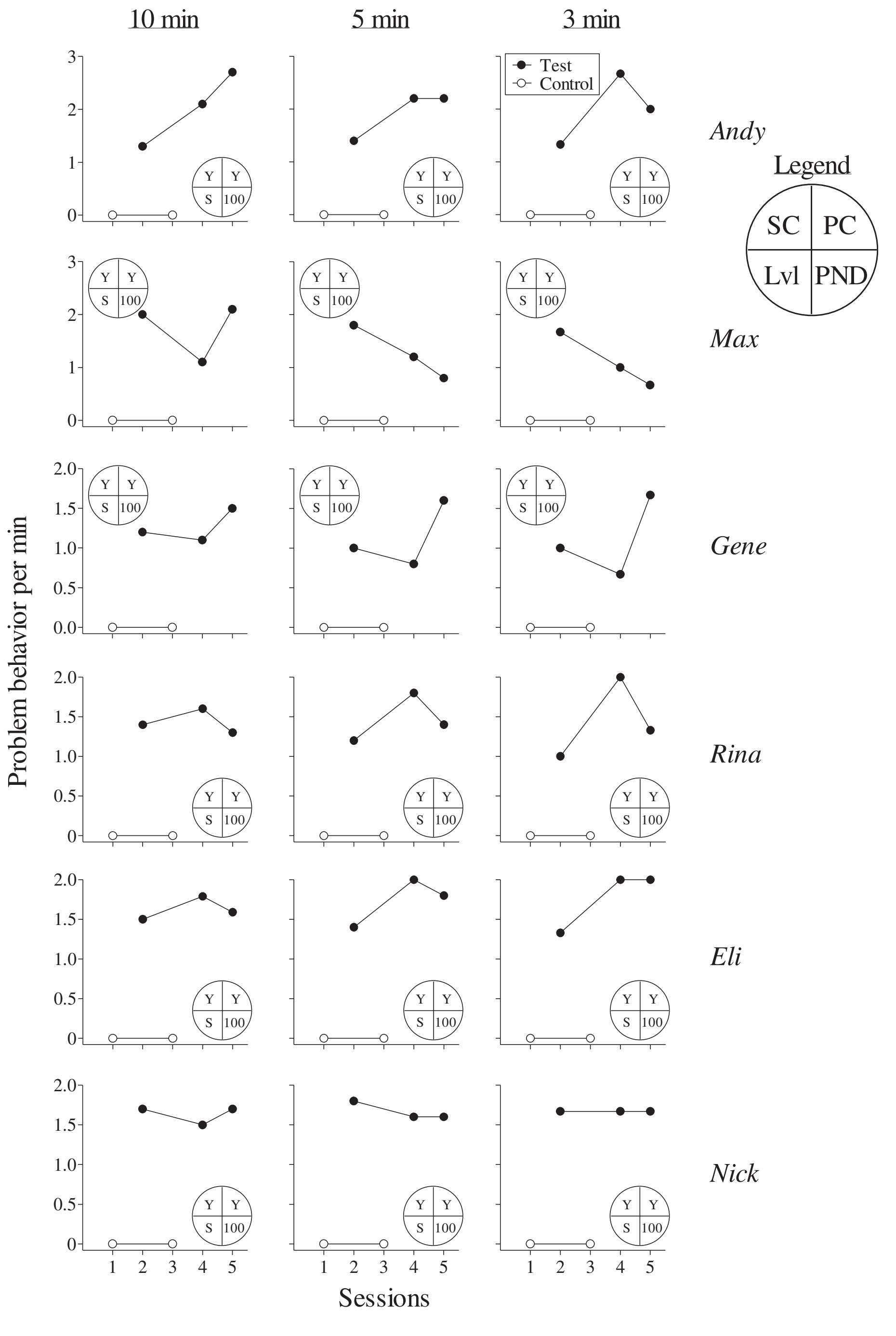

Figure 2. The reanalysis of the first 3 min (right) and 5 min (middle) of a 10-min IISCA (left) for six participants from Experiment 1. SC refers to structured criteria interpretation of control (Y = Yes, N = No). PC refers to panelist criterion. Lvl refers to the level of control (S = strong, M = moderate, W = weak, NC = no control). PND refers to percentage of nonoverlapping data.

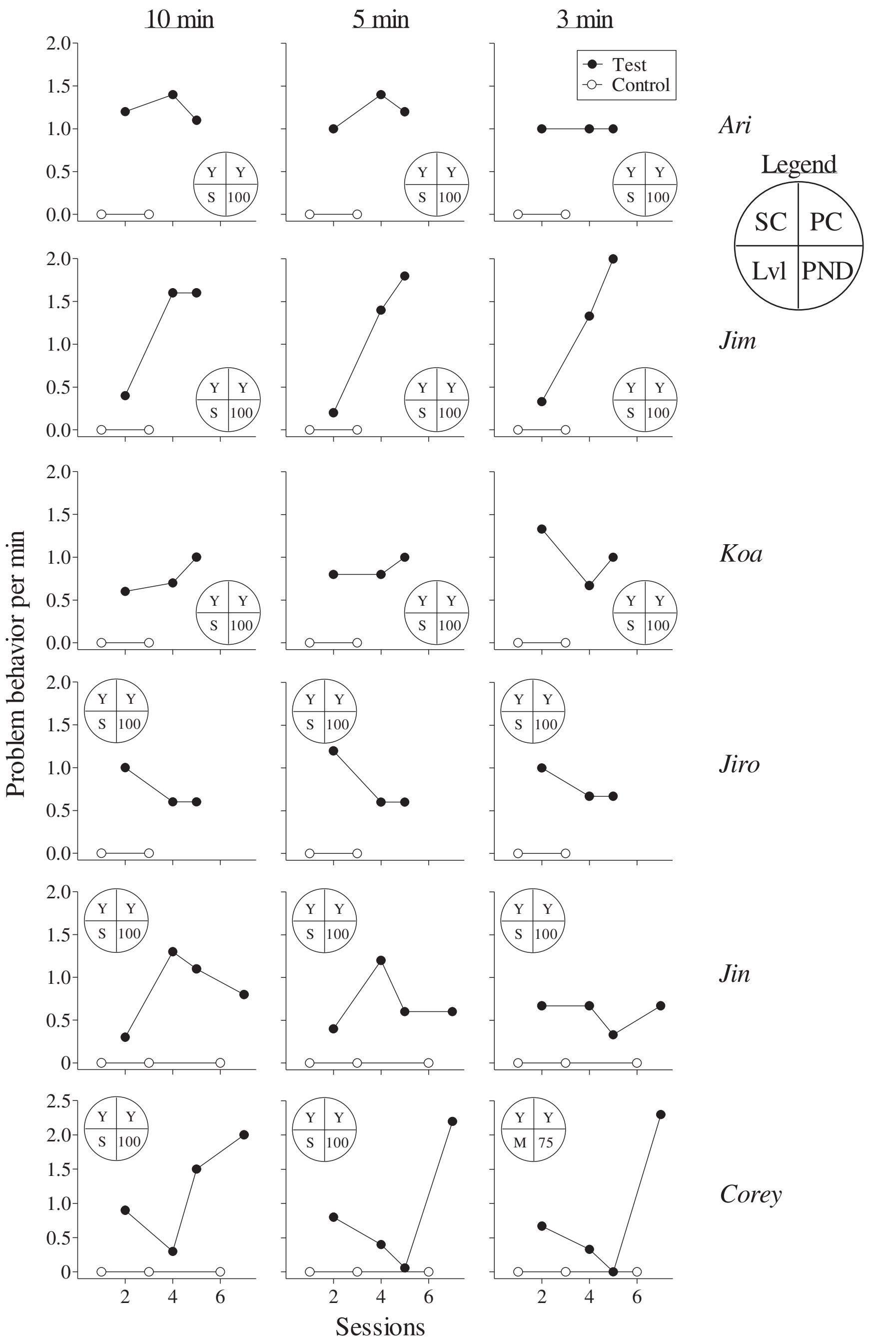

Figure 3. The reanalysis of the first 3 min (right) and 5 min (middle) of a 10-min IISCA (left) for six additional Experiment 1 participants. SC refers to the structured criteria interpretation of control (Y = Yes, N = No). PC refers to panelist criterion. Lvl refers to the level of control (S = strong, M = moderate, W = weak, NC = no control). PND refers to percentage of non-overlapping data.

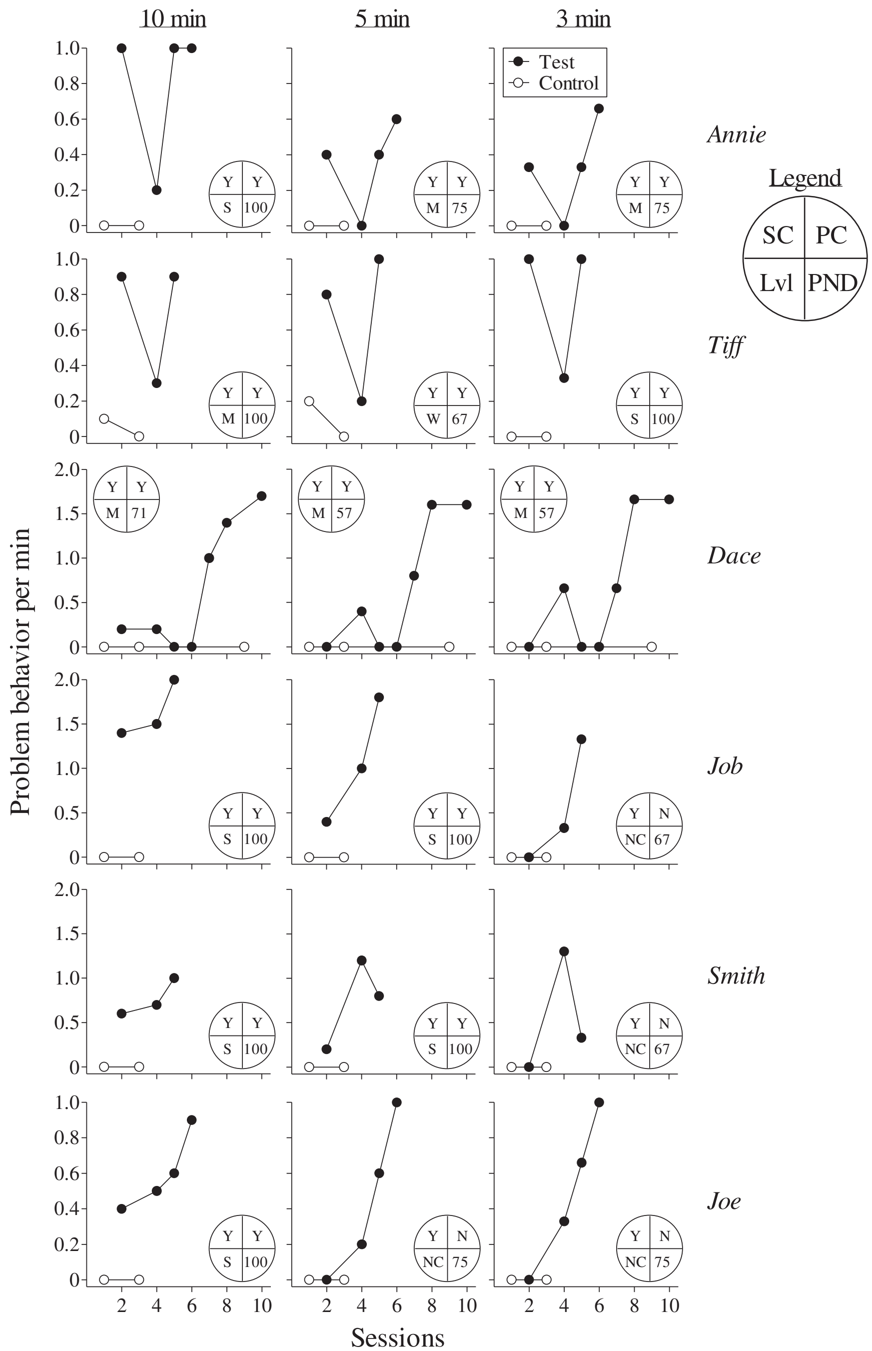

Figure 4. The reanalysis of the first 3 min (right) and 5 min (middle) of a 10-min IISCA (left) for six additional Experiment 1 participants. SC refers to the structured criteria interpretation of control (Y = Yes, N = No). PC refers to panelist criterion. Lvl refers to the level of control (S = strong, M = moderate, W = weak, NC = no control). PND refers to percentage of nonoverlapping data.

during the 5-min (*M* = 93%; *SD* = 14.0) and 3-min analyses (*M* = 90%; *SD* = 15.4).

The reanalysis of 5-min and 3-min IISCAs indicates that shorter sessions can be conducted with a high likelihood of obtaining differentiated results, independent of the manner in which the data were analyzed. Furthermore, the effects of reducing session duration on the level of control were minimal. These findings suggest that a functional analysis can be efficient without sacrificing control, which is important for clinicians who may have time constraints. Based on these results, a practitioner may obtain differentiated outcomes in as little as 15 min and the entire assessment period (i.e., interview through IISCA) could be completed within 35 min under optimal circumstances.

EXPERIMENT 2: EIGHT REPLICATIONS OF 3-MIN IISCAS

In Experiment 1, we reanalyzed data from IISCAs conducted with 10-min sessions to determine if brief session durations would affect the interpretability of the results. However, the reanalysis in Experiment 1 shares the same limitation as the reanalysis of session durations of the standard functional analysis conducted by Wallace and Iwata (1999). Whereas particular time windows (e.g., the last 5 or 7 min in each session) can be removed from the data analysis, the participants still experienced the contingencies in effect during these time windows, and this history likely influenced responding in subsequent sessions. Differentiated outcomes with the briefer session durations may thus be inflated due to extended contact with the arranged contingencies. In Experiment 2, we conducted additional IISCAs with 3-min sessions, with additional clients, to address this limitation.

Method

Participants and settings. The first eight consecutive clients who were admitted to the outpatient clinic following the 18 participants from Experiment 1 served as participants in Experiment 2. Their median age was 6 years, and seven participants were boys. Additional participant characteristics can be located in Table 3.

Response measurement and interobserver agreement. We calculated the rate of problem behavior per minute across sessions during the IISCA. We calculated IOA for problem behavior and reinforcement during a mean of 40% (range, 20% to 60%) of each IISCA application using a partial agreement coefficient (see Experiment 1). The mean IOA obtained for

Table 3
Characteristics of Experiment 2 Participants

	Participant Characteristics				
Participant	Age	Sex	Diagnosis	Language Ability*	Problem Behavior
Omi	8	M	ASD, ID	3	Agg, dis, SIB, tantrums
Hina	4	F	ASD, ID	2	Agg, tantrums, SIB, loud voc
Levi	4	M	ASD, GAD	4	Tantrums, loud voc
Dan	4	M	ASD, ID	3	Agg, dis, tantrums, loud voc
Matt	10	M	ASD, ID, FASD, ADHD	4	Agg, dis, loud voc
Lei	8	M	ADHD	4	Agg, dis, tantrums, loud voc
Val	3	M	ASD, ID	1	Agg, dis, tantrums, SIB, undressing, loud voc
Duke	13	M	ASD	4	Agg, dis, tantrums, loud voc

Note. ASD is autism spectrum disorder. ID is intellectual disability. ADHD is attention deficit/hyperactivity disorder. GAD is generalized anxiety disorder. DS is Down syndrome. SIB is self-injurious behavior. ISB is inappropriate sexual behavior.
* 1 = non-verbal; 2 = 1-word utterances; 3 = short diffluent sentences; 4 = full fluency.

Table 4
Experiment 2 Functional Analysis Information

Participant	Iteration	Test Condition	Control Condition
Omi	First	Escape from adult access to independent access of iPad	Continuous independent access of iPad
Hina	First	Escape from adult-directed to child-directed play	Continuous child-directed play and no adult directions
Levi	First	Escape from adult-directed to independent play	Continuous independent play and no adult directions
Dan	First	Escape from adult-directed to child-directed play	Continuous child-directed play and no adult directions
Matt	First	Access to water play	Continuous independent water play
Lei	First	Escape from difficult academic instructions to child-directed play	Continuous child-directed play and no academic instructions
Val	First	Escape from blocked access to free access to leisure items and snacks	Continuous free access to leisure items and snacks
Duke	First	Escape from adult-directed to child-directed math completion	Continuous child-directed math completion with no adult directions

problem behavior and reinforcement was 97% (range, 92% to 100%) and 99% (range, 94% to 100%), respectively. In addition, we calculated interrater agreement from all of the IISCAs for the evaluation of control and level of control. The interrater agreement was 100% for both analyses.

Procedure and design. The IISCA procedures were identical to those of Experiment 1 with the exception that session durations were 3 min for each participant. The test condition of each IISCA was conducted in rapid alternation with the control condition in a multielement design. Additional information on the individual contingencies and the number of analysis iterations for each participant can be found in Table 4. Seven of the eight analyses included the synthesis of positive and negative reinforcement, whereas only a single analysis evaluated positive reinforcement alone. No analyses required modifications and differentiated results were obtained on the first attempt for all.

Data analysis. We used the same two sets of binary criteria (i.e., panelist and structured) from Experiment 1 to determine if the IISCA had control. For the panelist criterion, we gave the six trained panelists new slides including each individual IISCA. We evaluated the degree of control using the same multilevel criteria as in Experiment 1 that provided four possible outcomes: strong, moderate, weak, and no control. Lastly, we calculated effect sizes for each IISCA using PND.

Results and Discussion

The results of the eight IISCAs are presented in Figure 5. Elevated rates of problem behavior were observed during the test condition (M = 1.62 RPM; SD = 1.04) in comparison to the control condition (M = 0.02 RPM; SD = 0.08). Duke was the only participant to have any overlapping data points. In addition, problem behavior was completely eliminated during the control condition across participants with the exception of Val. The average total analysis duration was 16.1 min.

The results of the binary evaluations of functional control (top half of pie charts) were similar to that of Experiment 1. All but one IISCA (88%) were considered to show functional control using the panelist criterion (top right quadrant of pie charts), whereas all eight IISCAs were determined to have control based on the structured criteria (top left quadrant of pie

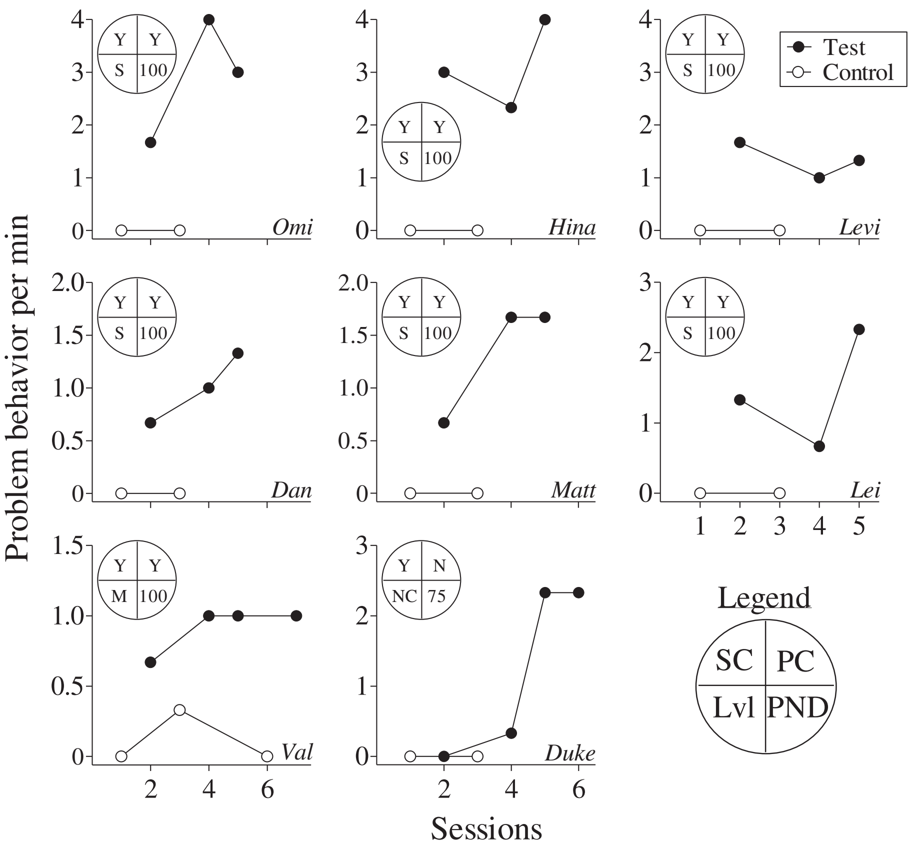

Figure 5. The results of the 3 min analyses for the eight participants of Experiment 2. SC refers to the structured criteria interpretation of control (Y = Yes, N = No). PC refers to panelist criterion. Lvl refers to the level of control (S = strong, M = moderate, W = weak, NC = no control). PND refers to percentage of nonoverlapping data.

charts). The evaluation of levels of control using the multilevel structured criteria identified six IISCAs (75%) with *strong* control, one IISCA with *moderate* control, and one IISCA with *no control.* All but one of the PND statistics was 100%. Duke's data set had minimal overlap (PND = 75%).

The majority of the 3-min IISCAs resulted in differentiated analyses. In addition, those that were differentiated generally had strong demonstrations of experimental control. These results were similar to those of Experiment 1, suggesting that the control shown in the 3-min analyses of Experiment 1 were not largely influenced by participant experiences represented in the excised data. Based on these results, we suggest that practitioners might obtain useful assessment results using the IISCA with 3-min sessions in most cases.

GENERAL DISCUSSION

Some boundaries of analytic efficiency and control were evaluated in the current investigation. A decrease in functional analysis session duration to as short as 3 min had limited impact on interpretations of functional control (see Figure 6 for a summary of Experiments 1 and 2). Overall, there was no detriment to experimental control in 67% of the IISCAs,

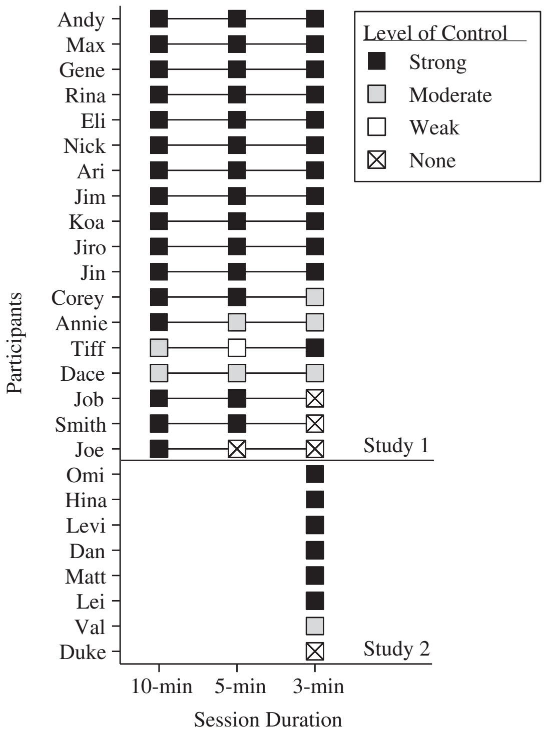

Figure 6. Summary of outcomes across all participants from Experiments 1 and 2 regarding level of control (strong, moderate, weak, none) across session durations.

and only minimal detriments in the level of control when reducing the sessions to 3 min (17%) and 5 min (11%). The findings from Experiment 1 were supported by the eight consecutive IISCAs conducted with 3-min sessions in Experiment 2. The majority of the eight IISCAs (75%) were considered to have strong control. Based on these results, we suggest that practitioners may be able to conduct functional analyses with sessions as brief as 3 min with minimal negative impact on their interpretations of functional control.

The level of correspondence between each of the four criteria used to evaluate functional control across the 62 distinct analyses[1] from Experiments 1 and 2 (Figure 7) was generally high. The two binary criteria (top panel) had a high level of correspondence with only 8% disagreement. The disagreements were specific to the panelist criterion and may be an artifact of how the conditions during the IISCA were arranged. The panelists reported that control could not be properly identified because differentiation was only observed after two control sessions and the first test session had been conducted. Thus, by only looking at the first 3 min in each condition, one cannot rule out the possibility that an unknown confounding variable influenced problem behavior across sessions, independent of the control and test conditions. The structured criteria were insensitive to this important aspect of experimental design. Interestingly, both the panelist criterion and multilevel criteria of control agreed with the panelist determinations of lack of control in all five analyses.

High levels of correspondence in the multilevel structured criteria and PND (bottom panel) were also observed, with 96% of the IISCAs that had no overlapping data being interpreted as having strong levels of control. Only one IISCA was interpreted as having weak control and although this corresponded to a smaller PND, smaller PNDs were obtained in two IISCAs that had moderate control. These minor disagreements may be due to the multilevel criteria being a more stringent determinant of fluctuations in control (e.g., for the two disagreements, the multilevel criteria underrated the obtained PND).

Both the multilevel criteria and PND provided a more nuanced interpretation of control in comparison to the binary evaluations, and both may be useful for predicting general therapeutic outcomes of the function-based treatments. This study did not examine treatment efforts, but treatments developed from and

[1]There were 18 IISCAs analyzed at three different session durations (i.e., 10-min, 5-min, 3-min) creating 54 distinct analyses in Experiment 1. Eight additional analyses with 3-min sessions were conducted in Experiment 2. This produces a total of 62 distinct analyses.

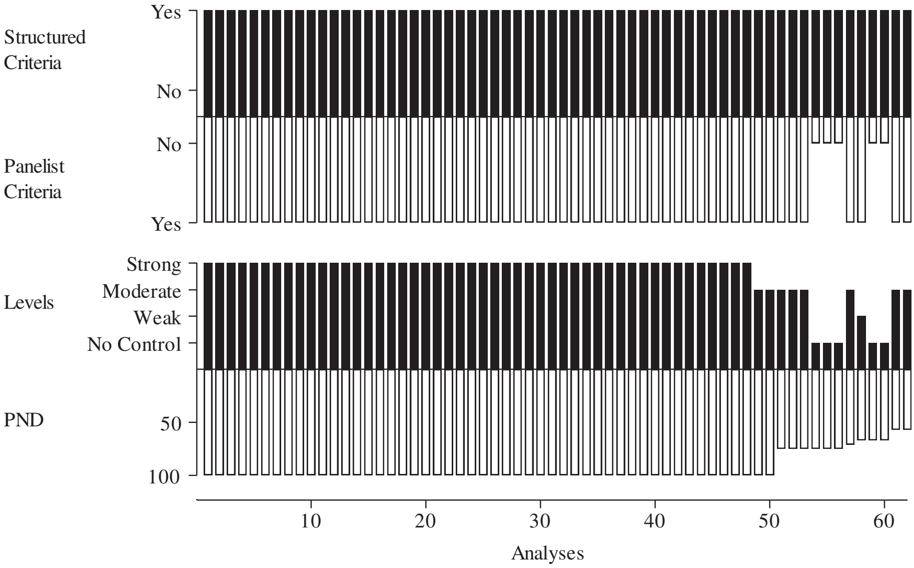

Figure 7. Summary of each of the four evaluations of functional control from Experiments 1 and 2. Each pair of white and black bars represents an individual analysis. There were 54 distinct analyses in Experiment 1 (18 IISCAs each represented in 10-min, 5-min, and 3-min sessions) and 8 analyses in Experiment 2, producing a total of 62 analyses.

informed by the IISCA have been independently evaluated in previous research (e.g., Beaulieu, Van Nostrand, Williams, & Herscovitch, 2018; Ghaemmaghami, Hanley, & Jessel, 2015; Hanley et al., 2014; Herman, Healy, & Lydon, 2018; Jessel, Ingvarsson, Kirk et al., 2018; Jessel, Ingvarsson, Metras et al., 2018; Rose & Beaulieu, 2018; Santiago, Hanley, Moore, & Jin, 2016; Slaton, Hanley, & Raftery, 2017; Strand & Eldevik, 2018). Evaluating the treatment relevance of the PND and multilevel criteria of control could involve two different courses of investigation.

First, this study could be replicated and extended to the point that each of the four evaluations of functional control are represented by multiple participants. Subsequent function-based treatment outcomes could be compared and the predictions noted above evaluated against those outcomes. Second, studies of existing function-based treatments can be collated in a literature review with the multilevel criteria retrospectively applied to the functional analyses that informed those treatments. The results of either study would indicate if strong control during functional analyses is likely to correspond with better treatment outcomes.

The degree of experimental control obtained in a functional analysis is important because inadequate control could lead to at least two complications in subsequent treatment. First, weak control might be predictive of the use of punishment or arbitrary reinforcement contingencies because problem behavior continues to occur despite the elimination of the putative establishing operation (Jessel et al., 2016). Second, moderate control (e.g., functional analyses with inconsistent levels of problem behavior) may also make it challenging to identify reliable evocative situations for differential reinforcement. Therefore, strong demonstrations of control (i.e., clear and stable differentiation in the functional analysis) may be better for treatment planning than weaker demonstrations of control. Differentiation with variability allows for the inference that a controlling variable has indeed been identified. However, variability also suggests there may exist other important controlling variables (Sidman, 1960).

REFERENCES

Beaulieu, L., Van Nostrand, M. E., Williams, A. L, & Herscovitch, B. (2018). Incorporating interview-informed functional analyses into practice. *Behavior Analysis in Practice, 11*, 385-389. https://doi.org/10.1007/s40617-018-0247-7

Beavers, G. A., Iwata, B. A., & Lerman, D. C. (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis, 46*, 1-21. https://doi.org/10.1002/jaba.30.

Bourret, J. C., & Pietras, C. (2013). The five pillars of the experimental analysis of behavior. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, & K. A. Lattal (Eds.), *APA handbook of behavior analysis* (pp. 33-64). Washington, DC: American Psychological Association. https://doi.org/10.1037/13937-002

Campbell, J. M. (2003). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: A quantitative synthesis of single-subject research. *Research in Developmental Disabilities, 24*, 120-138. doi.org/10.1016/S0891-4222(03)00014-3

Carr, M. E. (2015). A sensitivity analysis of three non-parametric treatment effect scores for single-case research for participants with autism. *Review Journal of Autism and Developmental Disorders, 2*, 67-78. https://doi.org/10.1007/s40489-014-0037-2

Carr, J. E., Severtson, J. M., & Lepper, T. L. (2009). Noncontingent reinforcement is an empirically supported treatment for problem behavior exhibited by individuals with developmental disabilities. *Research in Developmental Disabilities, 30,* 44-57.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson.

Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis, 36*, 387-406. https://doi.org/10.1901/jaba.2003.36-387.

Ghaemmaghami, M., Hanley, G. P., & Jessel, J. (2015). Contingencies promote delay tolerance. *Journal of Applied Behavior Analysis, 49*, 548-575. https://doi.org/10.1002/jaba.333

Hagopian, L. P., Fisher, W. W., Thompson, R. H., Owen-DeSchryver, J., Iwata, B. A., & Wacker, D. P. (1997). Toward the development of structured criteria for interpretation of functional analysis data. *Journal of Applied Behavior Analysis, 30*, 313-326. https://doi.org/10.1901/jaba.1997.30-313.

Hanley, G. P. (2012). Functional assessment of problem behavior: Dispelling myths, overcoming implementation obstacles, and developing new lore. *Behavior Analysis and Practice, 5*, 54-72. https://doi.org/10.1007/BF03391818

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147-185. https://doi.org/10.1901/jaba.2003.36-147

Hanley, G. P., Jin, C. S., Vanselow, N. R., & Hanratty, L. A. (2014). Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments. *Journal of Applied Behavior Analysis, 47*, 16-36. https://doi.org/10.1002/jaba.106

Herman, C., Healy, O., & Lydon, S. (2018). An interview-informed synthesized contingency analysis to inform the treatment of challenging behavior in a young child with autism. *Developmental Neurorehabilitation, 21*, 202-207. https://doi.org/10.1080/17518423.2018.1437839

Heyvaert, M., Saenen, L., Campbell, J. M., Maes, B., & Onghena, P. (2014). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: An updated quantitative synthesis of single-subject design. *Research in Developmental Disabilities, 35*, 2463-2476. https://doi.org/10.1016/S0891-4222(03)00014-3

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. https://doi.org/10.1901/jaba.1994.27-215

Jessel, J., Hanley, G. P., & Ghaemmaghami, M. (2016). Interview-informed synthesized contingency analyses: Thirty replications and reanalysis. *Journal of Applied Behavior Analysis, 49*, 576-595. https://doi.org/10.1002/jaba.316

Jessel J., Ingvarsson, E. T., Kirk, H., Whipple, R., & Metras, R. (2018a). Achieving socially significant reductions in problem behavior following the interview-informed synthesized contingency analysis: A summary of 25 outpatient applications. *Journal of Applied Behavior Analysis, 51*, 130-157. https://doi.org/10.1002/jaba.436

Jessel, J., Ingvarsson, E. T., Metras, R., Whipple, R., Kirk, H., & Solsbery, L. (2018b). Treatment of elopement following a latency-based interview-informed, synthesized contingency analysis. *Behavioral Interventions, 33*, 271-283. https://doi.org/10.1002/bin.1525

Kahng, S. W., Iwata, B. A., & Lewin, A. B. (2002). Behavioral treatment of self-injury, 1964-2000. *American Journal of Mental Retardation, 107*, 212-221. https://doi.org/10.1901/jaba.1997.30-267

Love, J. R., Carr, J. E., Almason, S. M., & Petursdottir, A. I. (2009). Early and intensive behavioral interventions for autism: A survey of clinical practices. *Research in Autism Spectrum Disorders, 3*, 421-428. https://doi.org/10.1016/j.rasd.2008.08.008

Oliver, A. C., Pratt, L. A., & Normand, M. P. (2015). A survey of functional behavior assessment methods used by behavior analysts in practice. *Journal of Applied Behavior Analysis, 48*, 817-829. https://doi.org/10.1002/jaba.256

Repp, A. C., Singh, N. N., Olinger, E., & Olson, D. R. (1990). The use of functional analyses to test causes of self-injurious behaviour: Rationale, current status and future directions. *Journal of Mental Deficiency Research, 34*, 95-105.

Roane, H. S., Fisher, W. W., Kelley, M. E., Mevers, J. L., & Bouxsein, K. J. (2013). Using modified visual-inspection criteria to interpret functional analysis outcomes. *Journal of Applied Behavior Analysis, 46*, 130-146. https://doi.org/10.1002/jaba.13

Rooker, G. W., Jessel, J., Kurtz, P. F., & Hagopian, L. P. (2013). Functional communication training with and without alternative reinforcement and punishment: An analysis of 58 applications. *Journal of Applied Behavior Analysis, 46*, 708-722. https://doi.org/10.1002/jaba.76

Roscoe, E. M., Phillips, K. M., Kelley, M. A., Farber, R., & Dube, W. V. (2015). A statewide survey assessing practitioners' use and perceived utility of functional assessment. *Journal of Applied Behavior Analysis, 48*, 830-844. https://doi.org/10.1002/jaba.259

Rose, J. C., & Beaulieu, L. (2018). Assessing the generality and durability of interview-informed functional analyses and treatment. *Journal of Applied Behavior Analysis*. Advance online publication. https://doi.org/10.1002/jaba.504

Santiago, J. L., Hanley, G. P., Moore, K., & Jin, C. S. (2016). The generality of interview-informed functional analyses: Systematic replications in school and home. *Journal of Autism and Developmental Disorders*, *46*, 797-811. https://doi.org/10.1007/s10803-015-2617-0

Scruggs, T. E., Maestropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education, 8*, 24-33. https://doi.org/10.1177/074193258700800206

Sidman, M. (1960). *Tactics of scientific research*. US: Authors Cooperative, Inc.

Slaton, J. D., Hanley, G. P., & Raftery, K. J. (2017). Interview-informed functional analyses: A comparison of synthesized and isolated components. *Journal of Applied Behavior Analysis, 50*, 252-277. https://doi.org/10.1002/jaba.384.

Strand, R. C. W., & Eldevik, S. (2017). Improvements in problem behavior in a child with autism spectrum diagnosis through synthesized analysis and treatment: A replication in an EIBI home program. *Behavioral Interventions, 33*, 102-111. https://doi.org/10.1002/bin.1505

Wallace, M. D., & Iwata, B. A. (1999). Effects of session duration on functional analysis outcomes. *Journal of Applied Behavior Analysis, 32*, 175-183. https://doi.org/10.1901/jaba.1999.32-175

Received June 13, 2018
Final acceptance December 11, 2018
Action Editor, Linda LeBlanc