Published in final edited form as:

J Appl Behav Anal. 2020 January ; 53(1): 44–66. doi:10.1002/jaba.583.

Efficiency in Functional Analysis of Problem Behavior: A Quantitative and Qualitative Review

Valdeep Saini,

University of Nebraska Medical Center's Munroe-Meyer Institute, Brock University.

Wayne W. Fisher,

University of Nebraska Medical Center's Munroe-Meyer Institute

Billie J. Retzlaff,

University of Nebraska Medical Center's Munroe-Meyer Institute

Madeleine Keevy

University of Nebraska Medical Center's Munroe-Meyer Institute

Abstract

Originating in the mid-1960s, functional analysis (FA) has become the gold standard method for understanding the environmental variables that come to shape and maintain problematic behaviors such as aggression, self-injury, and property destruction. Over the decades, a number of studies have refined FA methods, attempting to improve the overall efficiency of the analysis through experimental design and procedural modifications. In the present review, we used ongoing visual-inspection criteria and basic probability theory to compare and analyze levels of efficiency across FA types. The multielement design and synthesized contingency analyses were about equally efficient with respect to the mean number of sessions conducted per function tested, and the trial-based and synthesized contingency analyses were the 2 most efficient with respect to the mean duration per function tested. We discuss the implications of these findings in the broader context of efficiency and provide recommendations for maximizing efficiency during an FA. We also discuss other qualitative procedural details that may influence the overall efficiency of an FA.

Keywords

efficiency; functional analysis; functional assessment; reviews

According to Skinner (1953), a *functional relation* is one in which a change in an independent variable produces an orderly and predictable change in a dependent variable. Skinner suggested that the process through which one identifies the extent to which environmental variables are functionally related to behavior change be termed a *functional analysis*.

In the assessment and treatment of problem behavior, a functional analysis (FA) is a type of assessment in which environmental events hypothesized to maintain problem behavior are systematically introduced and withdrawn under highly controlled conditions (Betz & Fisher, 2011; Hanley, Iwata, & McCord, 2003). An FA of problem behavior typically includes (a) a condition in which an environmental outcome is contingent on problem behavior (e.g., the presentation of a reinforcer) and (b) a condition in which the environmental event is provided independent of the occurrence of problem behavior (i.e., a control condition; Thompson & Iwata, 2005). Behavior analysts examine differences in levels of problem behavior observed in one (or more) test conditions relative to the control condition using single-subject experimental designs.

In the 1960s and 70s, several researchers systematically manipulated individual environmental variables (e.g., escape) to determine their influence on problem behavior (e.g., Berkson & Mason, 1963; Carr, Newsom, & Binkoff, 1976, 1980; Lovaas, Freitag, Gold, & Kassorla, 1965; Lovaas & Simmons, 1969). However, Iwata, Dorsey, Slifer, Bauman, and Richman (1982/1994) dramatically advanced our field when they published their seminal study describing an operant method to assess the most common prevailing environmental variables hypothesized to reinforce self-injurious behavior. After repeatedly exposing nine participants to four experimental conditions (i.e., social disapproval, academic demand, being alone, and unstructured play), Iwata et al. reported that higher levels of self-injury were consistently associated with specific stimulus conditions for six of nine participants, thus demonstrating a functional relation between self-injury and specific environmental consequences.

Perhaps the most important feature of FA has been its systematic identification of the reinforcer for problem behavior for individual patients, which allows the behavior analyst to manipulate that consequence to decrease problem behavior and replace it with more prosocial behavior (Hanley et al., 2003; Tiger, Hanley, & Bruzek, 2008). That is, FA has emerged as the predominant method of prescribing effective behavioral treatments for persons who display problem behavior (Repp, 1994). A number of investigations have directly compared behavioral interventions based and not based on an FA, and results consistently favored the FA-based approach (e.g., Bachmeyer et al., 2009; Kuhn, DeLeon, Fisher, & Wilke, 1999; Smith, Iwata, Vollmer, & Zarcone, 1992). In addition, the results of a large-scale meta-analysis indicated that behavioral treatments based on FA tend to be more effective than pharmacological interventions (Didden, Duker, & Korzilius, 1997).

Since the publication of the methods described by Iwata et al. (1982/1994), research on improving FA methods has grown markedly, likely due to the successful outcomes of behavioral interventions based on FA (Kuhn et al., 1999; Smith et al., 1992). That is, the primary purpose of behavioral assessment is to ultimately inform treatment, and FA has historically been effective at informing function-based interventions (Hanley, 2012). As a result, behavioral scientists have focused much attention toward advancing FA methods.

Improvements to FA methods have taken many forms including, but not limited to, strategies to increase relevance to the natural environment (e.g., Mace & Lalli, 1991; Thomason-Sassi, Iwata, & Fritz, 2013), improve accuracy in detecting functional relations (e.g., Vollmer,

Marcus, Ringdahl, & Roane, 1995), increase applicability across populations (e.g., Cooper et al., 1992) or to individuals at remote locations (e.g., Wacker et al., 2013), increase applicability across topographies of problem behavior (e.g., Beavers, Iwata, & Lerman, 2013), as well as identify idiosyncratic or complex antecedent–behavior–consequence relations (e.g., Fisher, Lindauer, Alterson, & Thompson, 1998; Roscoe, Schlichenmeyer, & Dube, 2015). Refinements also have included data-analysis methods (e.g., Derby et al., 1994) as well as training individuals to visually inspect and accurately interpret FA results (e.g., Hagopian et al., 1997; Roane, Fisher, Kelley, Mevers, & Bouxsein, 2013).

One area that has received considerable attention in the FA literature has been improving the efficiency of FA such that the assessment process requires less time and effort, allowing treatment to be initiated more quickly. Although many studies have sought to improve FA efficiency, researchers have not reached consensus on a clear and objective definition of FA *efficiency*. Colloquially, efficiency might refer to an FA that saves time, effort, or resources, or one that increases the rapidity with which functional relations are identified. Thus, FA efficiency can be viewed on a multidimensional axis on which multiple variables may influence overall, or net, efficiency. For instance, an FA that is implemented with two therapists instead of three therapists might be considered more efficient (i.e., saves resources). An FA that could be conducted in the child's typical environment might be considered more efficient than one requiring the child to travel to a treatment clinic (i.e., reduces effort). Recently, Jessel, Hanley, and Ghaemmaghami (2016b) defined efficiency as the speed with which an FA provides a convincing demonstration of the variables that influence problem behavior. Based on this definition, speed might refer to the amount of time required (i.e., net run time) or number of sessions required to identify a functional relation. Given the difficulty in quantifying all of the various aspects of efficiency, an analysis of "speed" (i.e., time and sessions) seems to be a reasonable starting point for comparing the efficiency of various FA methods. Therefore, efficiency in FA might be best defined as the minimum amount of time and/or sessions required to objectively identify a functional relation.

Using this definition, a researcher could then develop structured criteria to determine the minimum amount of time required to identify a functional relation for a given FA. However, the "speed" of an FA is influenced by variables other than efficiency, such that researchers might continue conducting an FA beyond an initial demonstration of a functional relation. In other words, studies in the FA literature are not necessarily designed to provide the minimum amount of data sufficient to show a functional relation. Instead, the authors might be more concerned with providing a "definitive" rather than a "sufficient" demonstration of functional control (Baer, Wolf, & Risley, 1968). Studies also might involve comparisons between analysis models that require a larger number of sessions conducted for one or both methods in order to provide comparable data sets (e.g., Thomason-Sassi, Iwata, Neidert, & Roscoe, 2011). That is, FAs designed to convincingly isolate a specific function of problem behavior or to evaluate a procedural variation against a more thorough standard (e.g., Iwata et al., 1982/1994) are difficult to compare in terms of efficiency. Simply counting the number of sessions conducted during the FA as a measure of its efficiency, as Jessel et al. (2016b) did, could produce misleading results in those situations in which the researchers

extended the FA beyond what was necessary for a minimal demonstration of experimental control.

A critical feature of a comparative analyses, such as one that compares FA efficiency across experimental designs, requires selection of equally powered procedures that are designed to answer the same experimental question (Johnston, 1988; Van Houten, 1987). That is, given the definition of efficiency provided above, a comparative analysis of FA efficiency would require procedures that are explicitly designed to determine the minimum amount of data required to identify a functional relation. Saini, Fisher, and Retzlaff (2018) recently developed structured ongoing-visual inspection (OVI) criteria (based on an adaptation of the procedures described by Roane et al., 2013) that appear to be well suitable for such an analysis. Ferron, Joo, and Levin (2017) developed a similar ongoing visual-inspection approach; however, they designed their method specifically for interpreting multiple-baseline designs, and thus their procedure could not be applied to the prevailing types of FAs reported in the literature.

The OVI procedure developed by Saini et al. (2018) involved the application of the visual-inspection rules developed by Roane et al. (2013) but applied on an ongoing basis as the FA unfolded. This allowed Saini et al. to determine the point at which differentiation between one or more test conditions and the control condition could be objectively identified (i.e., the earliest point at which a functional relation could be demonstrated).

Saini et al. (2018) applied these criteria to a sample of 100 multielement FAs randomly selected from 1075 FAs in the extant literature that met their inclusion/exclusion criteria. For each of these 100 FAs, they applied the Roane et al. (2013) criteria on an ongoing basis starting with the first three sessions from each test and control condition. If they did not identify an operant function with these data samples, they then applied the criteria to the first four sessions of each test and control condition, and if necessary, the first five sessions, and so on. They continued this process until they either identified at least one operant function (e.g., attention) or they applied the criteria to the entire FA, whichever came first. If results met the Roane et al. criteria for at least one operant function (e.g., attention) when just examining a subset of the FA data, Saini et al. then applied an additional criterion to determine whether to stop there or to continue the OVI. That is, they evaluated whether it would be mathematically possible to identify an additional operant function (e.g., escape) if they added one more session per condition. If they could potentially identify an additional operant function, they continued the OVI; if it was not mathematically possible to identify an additional function, they ended the OVI.

Results of the Saini et al. (2018) study showed strong correspondence between the FA interpretations made using OVI and interpretations made using the Roane et al. (2013) criteria on the full FA (exact agreement = 93%). They reported even higher correspondence when the criterion variable was the consensus between (a) applying the Roane et al. (2013) criteria to the full FA and (b) the authors' interpretations of the FA reported in the paper (exact agreement with the consensus = 95%). In addition, Saini et al. achieved these high levels of predictive validity using 41% fewer sessions, on average, than the full FAs. As a result, OVI represents an appropriate method for a comparative analysis of FA efficiency

because it is designed to identify functional relations in FA data using only the minimum amount of required data. Using OVI in this manner circumvents limitations associated with variations in the lengths of FAs for a number of reasons unrelated to a minimal demonstration of functional control.

One limitation of the OVI procedures developed by Saini et al. (2018) with respect to an efficiency analysis of all FA types is that they developed a specific OVI for use with multielement and pairwise data displays, which represent the most common experimental designs used with FAs (Beavers et al., 2013). As such, their specific OVI method could not be used to compare the efficiency of a multielement FA with a trial-based FA, for example. Therefore, to compare the efficiency of all major FA methods, it would be necessary to develop a set of parallel OVI methods that could be used as a common metric for evaluating the efficiency of a variety of types of FAs such as trial-based FAs (Bloom, Iwata, Fritz, Roscoe, & Carreau, 2011), brief FAs (e.g., Northup et al., 1991), and FAs conducted using a reversal design (e.g., Piazza et al., 2003).

The expanded OVI methods we developed for evaluating the efficiency of various FA methods, which we describe in detail below, are based on two premises. The first is that when we visually inspect FA results, they show a clinically meaningful difference between one or more test conditions and the control condition. The second premise is that when we apply basic probability theory, we can reasonably conclude that the results are not due to chance fluctuations in the data (i.e., that the null hypothesis is false; Baer, 1977; Hays, 1988). Probability theory examines the likelihood of various outcomes when selected at random (e.g., the probability of drawing four aces in a game of five-card stud using a well-shuffled deck of cards). For the purpose of evaluating the efficiency of an individual FA in the current paper, we combined visual inspection with basic probability theory on an ongoing basis to identify the fewest number of sessions or trials necessary to satisfy both of these premises.

The purpose of the present study was to provide a systematic review of efficiency in the FA of problem behavior. We reviewed the FA literature and applied OVI and probability theory to objectively extract measures of efficiency across studies and FA types. In doing so, we sought to (a) evaluate the efficiency of different FA models, (b) identify barriers to improving FA efficiency, and (c) inform future research specific to FA efficiency. We discuss the implications of our efficiency analysis for the assessment of problem behavior as well as other variables that may contribute to efficiency.

Method

Article Identification

We selected all studies reviewed by Hanley et al. (2003) and Beavers et al. (2013; citations obtained from authors of those two studies). We then conducted an initial search of the published literature from 2011 through 2016 using search criteria based on those described by Hanley et al. Namely, we searched *PsycINFO*, *ERIC*, the *Journal of Applied Behavior Analysis*, and *Google Scholar*, using the keywords *functional analysis*, *functional assessment*, and *behavioral assessment*. We also examined the reference section of each

obtained article to identify additional studies not identified during the initial search. We subsequently updated our search using the same criteria described above to include published literature from 2017 and 2018.

Inclusion and Exclusion Criteria for Dataset Identification

We first organized all data sets according to the type/experimental design used (i.e., brief FA, trial-based FA, reversal design, multielement design, pairwise design, synthesized contingency analysis). For each FA type, we chose inclusion and exclusion criteria that ensured the FA demonstrated a functional relation and that also allowed us to apply basic probability theory (i.e., we included only FAs with at least 5 total data points).

For FA types for which we identified fewer than 100 datasets in the literature (e.g., trial-based FAs), we included all of the identified data sets that met the inclusion/exclusion criteria described below. For the one FA type for which we identified more than 100 data sets in the literature (i.e., multielement FAs), we randomly sampled studies from our initial data set (i.e., studies published up to 2016) that met the inclusion/exclusion criteria described below until we had analyzed 100 data sets. That is, we replicated the OVI procedures used by Saini et al. (2018) with the caveat that we included only studies not included in the prior Saini et al. investigation. In addition to these initial 100 data sets selected from studies published up to 2016, we also randomly selected approximately 10% of the multielement FAs published from 2017–2018.

Brief FA.—We included brief FAs if they contained: (a) an initial assessment of a control condition and at least one test condition; (b) a description of the duration of each session in the method section of the article (which was necessary to calculate efficiency); and (c) at least three replications of the test condition and two replications of the control condition (e.g., if the investigators tested for an escape function, they included an initial control and an escape session, plus a contingency reversal with a minimum of three sessions).

Trial-based FA.—We included trial-based FAs that allowed us to determine whether or not the participant emitted the target response during each trial (i.e., trial-by-trial data). In order to access this information, we sent an email to the corresponding author for each study published within the last 7 years that included at least one trial-based FA. We chose this time frame because research ethical standards typically require researchers to retain their raw data for 7 years (e.g., APA, 2007). We included all trial-based FAs for which we obtained trial-by-trial data.

Reversal design.—We included FAs using a reversal design if the authors (a) specified the duration of each session in the method section, (b) included a control condition, (c) conducted at least three sessions in each phase, and (d) included at least two phases of the control condition and two phases of each test condition. If the investigators replicated only one test condition, we analyzed the efficiency of the FA only for that specific function.

Multielement design.—We included FAs using a multielement design if the authors (a) specified the duration of each session in the method section, (b) included a control condition, (c) included at least two test conditions, and (d) conducted at least three sessions for each

condition. In addition, as mentioned above, we excluded all 100 of the multielement FAs analyzed using OVI by Saini et al. (2018).

Pairwise design.—We included FAs using a pairwise design if the authors (a) specified the duration of each session in the method section, (b) included a control condition, (c) included only one test condition per phase, and (d) conducted at least three test and at least two control sessions. Each phase of a pairwise FA was considered to be an application (e.g., if the authors first conducted a pairwise comparison of attention versus control, and then conducted a pairwise comparison of escape versus control, we treated each comparison as a separate application of a pairwise FA).

Synthesized contingency analysis.—We included synthesized contingency analyses if the authors (a) specified the information used to construct the test and control conditions, (b) specified the duration of each session in the method section, and (c) conducted at least three test and at least two control sessions.

Data Extraction

Except for trial-based FAs, we used PlotDigitizer (Huwaldt, 2016) to extract data from images of the graphs copied in PDF format. We first identified the x and y coordinates and the scales of the axes by entering values for the origin and the maximum values of the x - and the y - axes. We then extracted the value of every data point in the analysis by clicking on the center of each datum, which generated numerical x and y coordinates. This method proved applicable for all types of measures (e.g., response rate, percentage of intervals). We entered the data values produced by PlotDigitizer into a Microsoft Excel spreadsheet, created graphs for each FA, and applied the structured criteria described below based on the type of the FA.

Procedure

We developed OVI criteria paralleling those used by Saini et al. (2018) for the interpretation of each FA type. These criteria combined visual inspection with basic probability calculations to determine whether we could reasonably conclude that the apparent differences between the test and control conditions were not due to chance fluctuations in the data (i.e., that the null hypothesis was false; with $p < .05$). We provide a description of each set of structured criteria below.

OVI for brief FAs.—We first calculated the mean of the dependent variable (e.g., rate, percentage of intervals) for all sessions in the initial assessment (e.g., summed the number of responses across all of the alone, attention, escape, and toyplay sessions in the initial assessment, and then divided by the total number of minutes for those sessions). Next, we drew a horizontal criterion line through the figure corresponding to that mean. We identified a potential function as present (e.g., an escape function) if: (a) the data points for the target response always or almost always fell above the criterion line in the corresponding test sessions (e.g., in the escape sessions) and below the criterion line in the control and contingency-reversal sessions (e.g., when the alternative response produced escape); and (b) the probability that this observed pattern of responding occurred by chance was less than one in twenty ($p < .05$). To determine the second criterion, we counted the number of times the

target response fell above and below the criterion line in a manner consistent with the presence of a function (i.e., above the criterion line during test sessions and below the criterion line during control or contingency-reversal sessions). We then compared this number to the total number of sessions using the binomial distribution, which prior investigators have similarly applied and rigorously tested in structured, visual-inspection research (e.g., Fisher, Kelley, & Lomas, 2003). Table 1 shows the number of data points in the functional analysis and the corresponding number of data points that needed to be consistent with the presence of a function in order to consider the analysis differentiated (i.e., conclude the presence of a function).

The top panel of Figure 1 shows a hypothetical brief FA that, based on visual inspection, meets the first criterion of showing a clinically meaningful difference between one test condition (i.e., escape) and the control condition. To evaluate the second criterion, we: (a) calculated the number of times the target response (i.e., aggression) fell above and below the criterion line in a manner consistent with the hypothesized escape function and (b) compared those numbers using Table 1. That is, the data points for aggression fell above the criterion line each time aggression produced escape ($n = 2$ of 2) and below the criterion line in the control condition ($n = 1$ of 1) and during the contingency-reversal session when an alternative response (i.e., signing) produced escape ($n = 2$ of 2). Thus, the data points for aggression fell above or below the criterion line in a manner consistent with an escape function for 5 of the 5 opportunities. Thus, for this example, (a) visual inspection of the data revealed clear, and clinically meaningful differences between the test and control conditions consistent with an escape function and (b) probability analysis indicated that it is highly unlikely that these results could be attributed to chance. These results met our structured OVI criteria for demonstrating an escape function, and we would conclude that this application identified a function of problem behavior in 7 sessions.

OVI for trial-based FAs.—We determined that a trial-based FA identified a function of the target response if (a) the participant consistently emitted the target response during one or more of the test conditions, (b) the participant consistently omitted the target response during the control condition, and (c) the probability that this observed pattern of responding occurred by chance was exceedingly low ($p < .05$). We evaluated the first two criteria using visual inspection. We evaluated the third criterion by calculating the number of test and control trials that produced results consistent with a given function (e.g., attention) and then compared that number with the total number of trials using the binomial test as described above. We began by applying these criteria to the first five trials of the FA because a minimum of 5 of 5 trials consistent with a hypothesized function represents the fewest needed to reject the null hypothesis using the binomial equation (i.e., 4 of 4 trials consistent with a hypothesized function produces a p -value of .06; 5 of 5 produces a p -value of .03). If we did not identify a function, we repeatedly reapplied the criteria with one additional series of test and control trials. We continued this process until we identified a function or included all of the available trials, whichever came first. If we applied the criteria with all of the trials included and we did not identify a function, we scored the trial-based FA as undifferentiated. If a given trial-based FA tested for more than one function, we terminated our OVI of that FA after (a) we applied the OVI criteria with all of the available trials included or (b) we

identified at least one function and we ascertained that extending the trial-based FA by another series (i.e., one additional series of test and control trials) could not result in the identification of an additional function.

The second panel of Figure 1 shows a hypothetical trial-based FA that, based on visual inspection, meets the first two criteria and suggests an attention function. To evaluate the third criterion, we: (a) calculated the number of times the target response (i.e., disruption) occurred during the first three test and control trials. We observed disruption in the attention trials during 3 of 3 trials and did not observe disruption in 2 of the first 3 control trials. We then compared the number of trials consistent with an attention function ($n = 5$) to the total number of trials ($n = 6$). Given the nonsignificant result, we would then add the next test and control trials and reapply the binomial equation to the first eight trials of the trial-based FA (i.e., $n = 4$ of 4 for test trials and $n = 3$ of 4 for control trials). In this case, when the number of trials consistent with an attention function ($n = 7$) is compared to the total number of trials ($n = 8$) using the binomial test, the probability of the observed pattern occurring by chance is exceedingly low. Therefore, after 8 trials, the trial-based FA would be considered differentiated, indicating an attention function for the target behavior.

OVI for reversal designs.—For reversal designs, we applied the conservative dual-criteria method described by Fisher et al. (2003) because the investigators similarly based this method on the binomial equation and rigorously tested the method using a Monte Carlo study. We applied these criteria independently for each potential function that was evaluated within the FA (e.g., if the FA evaluated escape and attention as possible reinforcers, we applied the dual-criteria method to compare each test condition with the control condition independently).

To complete the calculations, we used the Excel workbook provided by the first author of Fisher et al (2003). If the FA included the minimum number of sessions required by the dual-criteria method in each test phase (i.e., a minimum of 5 sessions), we applied the criteria independently to each adjacent pair of phases in the FA (in an ABAB design, we compared A_1 with B_1 , B_1 with A_2 , and A_2 with B_2). If one or more phases had fewer than 5 sessions, we combined the corresponding test and control phases and then applied the dual-criteria method.

We began by applying the criteria with the first five test sessions (in or across phases) in comparison with the first five control sessions. If we did not identify a function, we added an additional series of test and control sessions (i.e., the first six sessions per condition) and then reapplied the criteria. We repeated this process until we identified a function for a given test condition (e.g., attention) or we applied the criteria to the entire dataset.

The third panel of Figure 1 shows a hypothetical reversal FA that, based on visual inspection, meets the first two criteria and suggests an attention function. To evaluate the third criterion, we created the mean criterion line and trend criterion line as described by Fisher et al (2003). Given this hypothetical FA has less than 5 sessions, we would combine the test and control sessions across phases. The first 5 test sessions were each above both the mean and trend criterion lines. Using the binomial distribution, the probability of 5 out of 5

sessions being above each line is exceedingly low, and therefore we would consider this FA differentiated, demonstrating an attention function, after the 10 sessions (i.e., the first 5 toy play and the first 5 attention sessions). In this case, the data points shaded in gray (i.e., the 6th toy play and attention sessions) would be considered unnecessary using the OVI criteria.

OVI for multielement FAs.—For multielement FAs, we used the OVI criteria described by Saini et al. (2018), which applied the formulae and guides described by Roane et al. (2013) on an ongoing basis as the FA unfolded. We used a Microsoft Excel workbook that automatically inserted into the graph (a) the upper criterion line (UCL) at one standard deviation above the mean for the control condition and (b) the lower criterion line (LCL) at one standard deviation below the mean of the control condition, or at zero if the standard deviation was larger than the mean. We then applied the Roane et al. criteria for each potential function in each FA (e.g., if the FA dataset tested for a tangible and an escape function, the experimenter applied the Roane et al. criteria to each of those two potential functions). For each potential function, we began by examining the first three data points per condition ($n = 3$ in the equation below). The experimenter then (a) summed the number of data points above the UCL (u), (b) summed the number of data points below the LCL (l), (c) summed the number of test sessions conducted for that function (n), and (d) used the values for u , l , and n to evaluate whether the following equation held true.

$$u - 1 \geq n/2$$

We also applied rules from the appendix of Roane et al. for FAs with trends in the data, low-magnitude effects, or response patterns consistent with either multiple maintaining reinforcers or automatic reinforcement.

If the equation held true for one or more test conditions or the data met criteria for a function based on a special rule, we concluded that a functional relation existed between the test condition(s) and problem behavior. If we identified one or more functions based on the first three data points per condition (e.g., escape), we then evaluated whether it would be mathematically possible to identify one or more additional functions (e.g., attention) by including one additional data point per condition (i.e., $n = 4$). For example, the bottom panel of Figure 1 shows a hypothetical multielement FA with escape and attention as the test conditions and play as the control condition. As can be seen, the escape condition has three data points above the UCL and no data points below the LCL, whereas the attention condition has two data points above the UCL and one below the LCL. Thus, the above equation holds true for the escape condition ($3 - 0 \geq 3/2$) but not for the attention condition ($2 - 1 < 3/2$). However, if we included one additional data point per condition ($n = 4$) and the fourth attention datum fell above the UCL, then we also would identify an attention function (because $3 - 1 \geq 4/2$). Thus, in this case we would continue the multielement FA and end it only when we reached a point at which (a) it was mathematically impossible to identify an additional function by including one additional data point per condition, (b) we identified an operant function for each condition tested, or (c) we included all data points from the FA.

OVI for pairwise FAs.—We used the OVI criteria described above for multielement FAs to analyze pairwise FAs with one difference. Because the pairwise FA evaluates a single function per comparison, we ended the OVI when the above equation held true for that function or after we conducted the analysis with all available data. If a researcher tested the effects of multiple, individual contingencies in a pairwise fashion, we treated each test and control combination (e.g., attention v. control) as a separate application of a pairwise FA.

Synthesized contingency analysis.—We used the OVI criteria described above for pairwise FAs for synthesized contingency analyses because they represent a special variation of a pairwise FA. That is, a synthesized contingency analysis combines multiple contingencies into a single test condition in comparison with a single control condition.

Measurement of FA Efficiency

In addition to recording the identified function(s) of problem behavior using the various structured criteria described above, we also quantified the number of sessions and amount of time required to identify one or more functions of problem behavior or to determine that the FA produced undifferentiated results. To measure assessment duration, we multiplied the number of sessions required to meet the FA termination criteria by the reported duration of each session. For trial-based FAs, we calculated the maximum duration of time the assessment could have taken because authors reported the maximum trial duration but not the duration of shorter trials (i.e., ones in which the trial ended early due to a short latency to the target behavior). Some FA types test for the presence or absence of multiple potential functions of behavior (e.g., multielement design), whereas others test for the presence or absence of one potential function (e.g., pairwise). In order to equate the number of functions tested, we also (a) divided the mean number of sessions required by the number of functions tested and (b) divided the mean duration of time required by the number of functions tested. We made these calculations to allow for a comparison of efficiency while accounting for the number of functions tested because testing for multiple functions typically takes longer than testing for a single function.

Interrater Agreement

Two independent raters evaluated a minimum of 33% of the FAs of each type (e.g., 46.7% of the brief FAs). For each pair of ratings, we scored an exact agreement if the raters agreed on the function or functions of the target behavior and the number of sessions required to determine the function. The independent raters showed exact agreement for 100% of the brief FAs, 100% of the trial-based FAs, 93.7% of the FAs using reversal designs, 100% of the pairwise FAs, 88.6% of the multielement FAs, and 93.3% of the synthesized contingency analyses. For all disagreements, the two raters discussed and reapplied the OVI criteria until they reached a consensus agreement, and this consensus was used in the analysis.

Results

Table 2 shows the mean number of sessions and mean duration of assessment for each type of FA, as well as the mean number of sessions and mean duration required per function

tested. Figure 2 shows the mean duration on the y axis for the assessment (black bars) and per function tested (white bars) across all FA types.

The sample of brief FAs included 15 brief FAs in five research articles. Each brief FA contained sufficient replications to test for only one potential function (e.g., in the initial assessment the investigators included an attention and an escape condition; however, they conducted the replication or contingency-reversal phase only for the attention function). The OVI criteria agreed with the structured criteria applied to the full data sets for 100% of the applications and with the post-hoc author interpretation for 86.7% of the applications (see Table 3 for exact agreement between OVI and the structured criteria applied to the full dataset and with the post-hoc author interpretations for all FA types). Overall, the brief FAs included in this investigation required a mean of 7.2 sessions (standard deviation [SD] = 0.6) and a mean of 72.0 min (SD = 5.6) for both the total assessment when we used OVI and when tested on a per-function basis (i.e., because each brief FA tested for only one potential function). Using OVI with these brief FAs reduced the mean number of sessions conducted by 4.0% (from the 7.5 sessions conducted by the authors to 7.2 sessions when applying OVI).

The sample of trial-based FAs included 10 data sets published in three research articles. The OVI criteria agreed with the structured criteria applied to the full data set for 90.0% of the applications and with the post-hoc author interpretation for 90.0% of the applications. Overall, the trial-based FAs included in this investigation required a mean of 41.6 trials (SD = 23.5) and a mean of 59.6 min (SD = 23.3) for the total assessment when we used OVI. Using OVI with these trial-based FAs reduced the mean number of trials conducted by 35.4% (from the 64.4 trials conducted by the authors to 41.6 trials when applying OVI).

When we analyzed the data on a per-function basis, trial-based FAs had a mean of 13.0 trials (SD = 4.3) and a mean of 19.5 min (SD = 5.5) per function. In addition, these time estimates represent the maximum time the trial-based FAs could have required because the authors reported the maximum trial lengths but not the actual, measured trial lengths. However, given the small sample size for trial-based FAs, these data should be interpreted with caution.

The sample of FAs that used reversal designs included 48 applications published in 17 research articles. The OVI criteria agreed with the structured criteria applied to the full data sets for 100% of the applications and with the post-hoc author interpretation for 93.8% of the applications. Overall, the FAs using a reversal design required a mean of 27.6 sessions (SD = 12.3) and a mean of 281.7 min (SD = 171.6) when we used OVI to interpret the results and determine the endpoint of the analysis. Using OVI with the FAs that used reversal designs reduced the mean number of sessions conducted by 13.0% (from the 31.7 sessions conducted by the authors to 27.6 sessions when applying OVI). When we analyzed the data on a per-function basis, the FAs using reversal designs had a mean of 15.4 sessions (SD = 5.6) and a mean of 163.3 min (SD = 140.9) per function.

The sample of multielement FAs included 105 applications published in 35 research articles. The OVI criteria agreed with the structured criteria applied to the full data sets for 94.3% of

the applications and with the post-hoc author interpretation for 84.8% of the applications. Overall, the included FAs using a multielement design required a mean of 16.6 sessions ($SD = 10.2$) and a mean of 150.6 min ($SD = 165.4$) when we used OVI to interpret the results and determine the endpoint of the analysis. Using OVI with these multielement FAs reduced the mean number of sessions conducted by 19.0% (from the 20.5 sessions conducted by the authors to 16.6 sessions when applying OVI). When we analyzed the data on a per-function basis, the included multielement FAs had a mean of 5.3 sessions ($SD = 3.4$) and a mean of 47.9 min ($SD = 55.5$) per function.

The sample of pairwise FAs included 72 applications published in 22 research articles. Each pairwise FA tested only for one potential function of behavior. The OVI criteria agreed with the structured criteria applied to the full data sets for 98.6% of the applications and with the post-hoc author interpretation for 91.7% of the applications. Overall, the included FAs using a pairwise design required an mean of 8.9 sessions ($SD = 6.0$) and a mean of 82.2 min ($SD = 70.7$) for both the total assessment as well as on a per-function basis (because we treated each pairwise analysis as a separate FA). Using OVI with these pairwise FAs reduced the mean number of sessions conducted by 28.2% (from the 12.4 sessions conducted by the authors to 8.9 sessions when applying OVI).

The sample of synthesized contingency analyses included 53 applications published in eight research articles. Each synthesized contingency analysis compared only one test condition with one control condition. The OVI criteria agreed with the structured criteria applied to the full data sets for 100% of applications and agreed with the post-hoc author interpretation for 92.5% of the applications. Overall, the included synthesized contingency analyses required an mean of 5.2 sessions ($SD = 0.7$) and a mean of 25.1 min ($SD = 11.7$) for both the total assessment as well as on a per-function basis (because the authors tested only a combined contingency). Using OVI with these synthesized analyses reduced the mean number of session conducted by 11.9% (from the 5.9 sessions conducted by the authors to 5.2 sessions when applying OVI).

Discussion

We extended the OVI procedures developed by Saini et al. (2018) for use with a variety of FA types, including brief FAs, trial-based FAs, multielement FAs, pairwise FAs, FAs using reversal designs, and synthesized contingency analyses. Similar to the results obtained by Saini et al., OVI produced high levels of exact agreement for interpretation of all FA types when compared with (a) FA interpretation based on structured criteria similar to Roane et al. (2013) on the full FA data set and (b) FA interpretation provided by the authors of the original studies. In addition, when we applied OVI to these data sets, we identified one or more functions of problem behavior in fewer sessions (or trials) than the number of sessions (or trials) conducted by the original authors. However, the reduction in the number of sessions needed to demonstrate a functional relation varied greatly across the six FA types, with trial-based FAs, pairwise FAs, and multielement FAs showing relatively large reductions (i.e., reductions of 35.4%, 28.2%, and 19.0% in the number of trials or sessions required to accurately identify one or more functions, respectively) and brief FAs showing a negligible reduction (i.e., 4.0% decrease in the number of sessions required). These results

indicate that OVI represents an accurate and relatively efficient approach to interpreting FA results across all major FA types, which can substantially increase the efficiency of some of the FA types.

Another primary purpose of this investigation was to use OVI to objectively compare the efficiency of each of the six major FA types in terms of the minimum amount of time required to identify one or more operant functions of problem behavior while maintaining high levels of exact agreement with the full FA data sets. Similar to the findings reported by Jessel et al. (2016b), our results showed that the synthesized contingency analysis required the fewest number of sessions and least amount of time overall. However, in contrast with the findings of Jessel et al., when we divided the total number of sessions required to identify one or more functions by the number of contingencies tested to determine the number of sessions required per contingency, the synthesized contingency analysis and the multielement FA tied for the lowest number of sessions required. In addition, the trial-based FA required the least amount of time after correcting for the number of contingencies tested.

Our findings also differed from those of Jessel et al. (2016b) in the magnitude of differences observed in the overall number of sessions or duration of time required by various FA methods because we used different methods to assess efficiency than Jessel et al. did. That is, Jessel et al. counted the total number of sessions conducted for each FA without considering the point in time during the FA where one or more functions became evident. For example, Hagopian, Toole, Long, Bowman, and Lieving (2004) conducted 50 FA sessions with the participant named Sally even though the identified escape function became evident using OVI after 15 sessions (i.e., 70% fewer sessions than the number actually conducted). If Jessel et al. included the data for Sally in their analysis, they would have included all 50 sessions as their measure of FA efficiency. By contrast, we used only the 15 required to accurately identify the function of problem behavior using OVI. We argue that this latter number (15) represents a more accurate measure of efficiency of this FA (and FAs in general), because the other 35 sessions were not needed to accurately ascertain that escape from nonpreferred demands functioned as reinforcement for Sally's problem behavior. Overall, using our OVI criteria, we found multielement FAs to be more than twice as efficient (averaging 151 min to identify one or more functions) than Jessel et al. did (averaging 313 min when counting every session), and we found trial-based FAs to be more than four times as efficient (averaging 54 min) than Jessel et al. did (averaging 232 min).

One limitation of the present study is that, in part, due to our strict inclusion criteria (which was necessary for appropriate application of OVI and basic probability theory), we did not acquire 100 representative applications for each of the various FA type. Common reasons for excluding FA applications from this analysis included (a) not including a traditional control condition (i.e., a condition with no establishing operation present; e.g., Carr, Yarbrough, & Langdon, 1997; Day, Horner, & O'Neil, 1994; Lee, Sugai, & Horner, 1999) and (b) not describing the exact duration of sessions for individual participants (e.g., Arron, Hall, Sloneem, Forman, & McClintock, 2006; Asmus et al., 2004). Despite this limitation, results of the present study provide a relatively strong and objective comparison of efficiency across FA types. Therefore, this quantitative review presents several general findings that may be

useful for understanding the variables that affect how behavior analysts should proceed when conducting an FA.

Implications of Main Findings

Improving efficiency using OVI.—One method behavior analysts can use to improve the efficiency of many of these FA types is to conduct the FA using OVI as the FA unfolds and to end the analysis after reaching the OVI termination criteria. Using OVI may have the largest effect on efficiency for trial-based FAs, pairwise FAs, and multielement FAs, and the least effect on efficiency for brief FAs and synthesized contingency analyses. Given brief FAs and synthesized contingency analyses were developed, in part, to increase efficiency of FA, these designs already tend to include only the necessary number of sessions for a believable account.

Implications for brief FAs.—The OVI criteria used in the present study may provide additional support for interpretation of brief FAs. In order to test for the presence of a specific individual function, we recommend that at least five sessions related to each operant function are conducted (e.g., three test and two control conditions), such that the OVI criteria can be applied.

Implications for trial-based FAs.—Our estimate of the duration of time required to identify one or more functions of problem behavior using trial-based FAs represents the maximum time it could have taken (if the individual consistently emitted the target response in the last second of the trial). When calculating efficiency relative to the number of operant functions tested, the trial-based FA required the least time per function. Given our conservative method of estimating the efficiency of the trial-based FA, it is likely that the efficiency of trial-based FAs relative to the other FA methods may be greater than the current results indicate.

Another potential advantage of trial-based FAs relative to a multielement or FAs using reversal designs is that trial-based FAs may be more conducive to typical environments (e.g., classrooms), because they may be more easily implemented by typical caregivers (Austin, Groves, Reynish, & Francis, 2015; Gerow, Rispoli, Gregori, & Sanchez, 2019). Given that trial-based FAs tend to require fewer staff and less time (per function) than the other FA methods (see Bloom et al., 2011 for an exception), they might be less disruptive in classroom settings where teachers are required to manage several children at a time (cf. Kodak, Fisher, Paden, & Dickes, 2013).

Implications for FAs using a reversal design.—Reversal designs proved to be the least efficient type of FA. A primary feature of the more efficient design types (e.g., multielement FA, synthesized contingency analysis) is rapid alternation between test and control conditions, which typically allows for a more efficient analysis compared to a reversal design. The reversal design tends to be less susceptible to internal validity threats (e.g., multiple treatment interference) and likely should be used primarily when efficiency must be sacrificed at the cost of accuracy (e.g., cases where an initial FA of a different type produces inconclusive results).

Implications for multielement FAs.—When considering efficiency with respect to the individual functions analyzed in an assessment, the multielement design proved to be equally efficient as those that involved fewer sessions but also evaluated fewer or just one contingency. We highlight this point because the multielement design is the most commonly used experimental design when conducting an FA (Beavers et al., 2013). Results of the present review suggest that an FA arranged in a multielement design is a relatively efficient method to assess the reinforcing effects of multiple putative reinforcers on problem behavior.

The vast majority of multielement FAs involved sessions lasting between 5 and 15 min. Thus, it is likely that the efficiency of the multielement FA could be increased substantially simply by reducing the duration of each session. For example, behavior analysts might begin a multielement FA using 3-min test and control sessions, and increase the duration of the sessions only if differentiation between one or more of the test conditions and the control condition is not evident or emerging after conducting three to four sessions per condition.

Implications for pairwise FAs.—Pairwise FAs are more efficient than those using a multielement design when the total assessment time is considered; however, this advantage is lost if more than one potential operant function is examined. Therefore, pairwise designs may be best suited for situations in which other information (e.g., descriptive assessment outcomes) suggests the target behavior is likely maintained by a specific operant contingency.

Implications for synthesized contingency analyses.—The synthesized contingency analysis appears to obtain its efficiency by combining programmed EOs, discriminative stimuli, and consequences into a single test condition for problem behavior (Hanley, Jin, Vanselow, & Hanratty, 2014). During a synthesized test condition, Hanley et al. simultaneously introduced multiple EOs identified by either an interview or structured observation (e.g., withdrawal of preferred toys and introduction of nonpreferred demands), and then reintroduced those putative reinforcers (e.g., escape plus access to preferred toys) contingent on problem behavior. They compared this synthesized test condition with a control condition in which they made all of the putative reinforcers available to the participant continuously and noncontingently (thereby removing all of the putative EOs). Given that this method included just one test and one control condition, its overall efficiency tended to be high (e.g., mean length of about 23 min in Hanley et al.). One advantage of this strategy is that the treatment process begins quickly. For example, Jessel, Hanley, and Ghammaghami (2016a) reported that the synthesized contingency analysis had a mean of 25 min to conduct across 30 cases.

Although combining contingencies in this manner efficiently identifies one or more of the synthesized contingencies reinforces problem behavior, it cannot be determined whether one, some, or all of the individual contingencies are necessary or sufficient for reinforcement of problem behavior. For example, in a test condition with a synthesized contingency involving both contingent escape and access to toys, it is impossible to know whether problem behavior is maintained by (a) negative reinforcement as an individual contingency, (b) positive reinforcement as an individual contingency, or (c) only when negative and

positive reinforcers are synthesized into a single contingency. As such, additional research is needed involving within-subject comparisons of the results of the synthesized contingency analysis with those of other types of FAs to determine whether synthesizing contingencies in this manner obtains its efficiency at the cost of accuracy (potentially increasing false positive or negative outcomes; cf. Fisher, Greer, Romani, Zangrillo, & Owen, 2016; Greer, Mitteer, Briggs, Fisher, & Sodawasser, provisionally accepted; Retzlaff, Fisher, Akers, & Greer, in press).

Other Variables that Affect FA Efficiency

As described previously, the overall time required is only one variable among many that might affect FA efficiency. That is, it is likely that multiple variables interact to affect the rapidity with which a functional relation can be identified. For a thorough understanding of efficiency, one should consider these qualitative variables as well.

Format and Design.—One method to increase efficiency is to screen for automatically reinforced problem behavior before assessing potential social functions of problem behavior. Querim et al. (2013) used the typical alone (or ignore) condition of the FA as a screening tool for problem behavior maintained by automatic reinforcement. They conducted a series of consecutive, 5-min sessions and observed whether problem behavior extinguished (indicative of problem behavior maintained by social reinforcement) or persisted (indicative of problem behavior maintained by automatic reinforcement). The authors found that this screening procedure predicted whether social or automatic consequences reinforced problem behavior in 28 of 30 cases. In addition, they found that the mean duration of the screening for automatic reinforcement lasted 21.5 min, suggesting that it is one of the more efficient FA methods when comparing the methods on a per-function basis.

The advantage of the Querim et al. (2013) screening procedure with respect to efficiency is that when it identifies an automatic function for problem behavior, the behavior analyst may initiate treatment immediately thereafter. However, one potential limitation of this method is that it may fail to identify cases wherein problem behavior is multiply controlled by automatic and social reinforcement (albeit this has been reported infrequently in the literature; Beavers & Iwata, 2011). This limitation also applies to those FA methods that only test for social reinforcement (e.g., brief FA, synthesized contingency analysis). That is, because these FAs do not test for automatic reinforcement, they cannot rule out the possibility that high levels of the target response in the test conditions may, in part, result from automatic reinforcement, and low levels of the target response in the control condition may, in part, result from the availability of competing sources of stimulation (e.g., attention, preferred tangible items).

Some authors have suggested that pre-assessment information can be used to improve the efficiency of an FA (e.g., Hanley et al., 2014). That is, testing for conditions that are unlikely to function as reinforcers for the given problem behavior, or that have not historically been correlated with the target response, may increase the overall assessment duration unnecessarily. This latter factor may be especially pertinent to the inclusion of a test for problem behavior maintained by tangible reinforcement when no pre-assessment

information exists indicating that the target behavior has produced contingent access to tangible items in the past. In addition, including a tangible test condition when there is no indication that access to tangible items had historically been correlated with the target response increases the risk of creating a tangible function when one did not exist in the typical environment prior to the assessment (cf. Rooker, Iwata, Harper, Fahmie, & Camp, 2011; Shirley, Iwata, & Kahng, 1999). As such, the tangible condition should only be included in an FA if the results of other assessments (e.g., interviews with caregivers, descriptive assessments) suggest a possible tangible function.

Data display.—The efficiency of an FA also may be affected by how functional relations are identified through visual inspection. For instance, in addition to examining response rate per session, an alternative measure to consider is within-session patterns of responding (Roane, Lerman, Kelley, & Van Camp, 1999; Vollmer, Iwata, Zarcone, Smith, & Mazaleski, 1993; Vollmer et al., 1995). For example, Jessel, Hanley, and Ghaemmaghami (2018) conducted a single-session synthesized assessment with three participants and graphed each reinforcer-present interval separate from each reinforcer-absent interval from within this single session. For two participants, problem behavior was exclusively observed during the reinforcer-absent intervals; for the other participant, problem behavior only occurred during one reinforcer-present interval of the five-minute session.

A second variable to consider with respect to data display is response aggregation, which involves the combining multiple topographies of problem behavior in a single FA (i.e., analyzing aggression, property destruction, and self-injurious behavior in a single FA; Derby et al., 1994; Lalli et al., 1995). When all of the response topographies are members of a single operant class, combining and analyzing those topographies in a single FA is obviously more efficient than conducting a separate FA for each individual topography. However, this method should be used with caution because numerous studies have shown that response aggregation during an FA can often lead to erroneous conclusions regarding operant function (Asmus, Franzese, Conroy, & Dozier, 2003; Beavers & Iwata, 2011; Derby et al., 1994; Hagopian, Rooker, Jessel, & DeLeon, 2013; Richman, Wacker, Asmus, Casey, & Andelman, 1999). That is, efficiency in identifying functional relations in an FA should not be at the cost of accuracy. As a result, one might consider minimizing the number of topographies in the contingency class when conducting an FA (e.g., Thompson, Fisher, Piazza, & Kuhn, 1998). At a minimum, when multiple topographies are included in the contingency class, data should be collected and displayed for aggregated responses as well as for each individual topography (cf. Derby et al., 1994), which allows one to determine if response topography systematically varies with FA conditions (e.g., aggression occurs in escape condition and self-injurious behavior occurs in alone condition).

Response measurement.—The manner in which dependent variables are measured also may affect the overall efficiency with certain measures being more efficient under certain circumstances. For example, Thomason-Sassi et al. (2011) examined the correspondence between response rate and response latency during multielement FAs and found that latency was a reliable and accurate index of standard FA outcomes that used response rate (i.e., the two measures were inversely correlated). The use of response latency as a primary measure

should significantly reduce the time needed to complete an FA because sessions are terminated upon the first occurrence of the response, and repeated observations of the behavior within each session are unnecessary.

Content and sequencing of conditions.—The content and order of conditions may affect the efficiency with which an FA reliably identifies a functional relation between a reinforcing consequence and the target behavior. For instance, the addition of stimuli that facilitate discrimination among the various test conditions and the control condition can reduce assessment duration. Connors et al. (2000) compared multielement FAs conducted without additional discriminative stimuli (i.e., all conditions conducted in the same room by the same therapist) to FAs that included additional discriminative stimuli (i.e., conditions correlated with different therapists and different colored rooms). The authors found that for half the participants the inclusion of programmed discriminative stimuli facilitated discriminated responding during the FA. These results suggest that the inclusion of programmed discriminative stimuli could potential increase the efficiency of an FA by more readily producing clear outcomes between test and control conditions.

Hammond, Iwata, Rooker, Fritz, and Bloom (2013) examined the effects of a fixed sequence of conditions designed to capitalize on EO effects (e.g., conducting the attention condition following the ignore condition to increase the value of attention as reinforcement) in comparison with a random sequence of conditions. Hammond et al. found more efficient (or more clearly differentiated outcomes) under the fixed condition sequence relative to a random sequence. This relatively simple modification increased FA efficiency by 57% as compared to a randomly sequenced FA. These investigators showed that they obtained this increased efficiency without sacrificing accuracy. As a result, behavior analysts might consider using the fixed-sequence approach proposed by Hammond et al. in order to increase FA efficiency.

Fahmie, Iwata, Quierm, and Harper (2013) suggested that test-specific control conditions for each potential source of reinforcement may be beneficial under certain circumstances and compared four FA control conditions: alone, ignore, play, and differential reinforcement for other behavior (DRO). Results showed that all of the conditions functioned as effective controls for problem behavior maintained by positive reinforcement, but the DRO condition consistently failed as a control for problem behavior maintained by negative reinforcement. Ultimately, they concluded that the most preferred control condition would be one that provides noncontingent access to the presumed reinforce(s) for problem behavior. However, with respect to efficiency, one also might consider using an alone or ignore condition as the control, which controls for positive and negative reinforcement contingencies in the various test conditions while simultaneously testing for problem behavior maintained by automatic reinforcement.

Accommodation influences.—The inclusion of caregivers and known settings might serve as another avenue to increase FA efficiency. For instance, Kurtz, Fodstad, Huete, and Hagopian (2013) found that caregiver-conducted FAs identified at least one function of problem behavior when staff-conducted FAs produced undifferentiated results. Moreover, when results of the staff-conducted FAs remained unclear, subsequent caregiver-conducted

FAs resulted in agreement between caregiver- and staff-conducted FAs in about 67% of cases. Results of Kurtz et al. and similar studies (e.g., Mueller, Nkosi, & Hine, 2011; Thomason-Sassi et al., 2013) suggest that caregiver-conducted FAs may, in some cases, increase efficiency by more reliably evoking problem behavior across various test conditions.

FA Modifications to Maximize Efficiency

Hanley et al. (2003) outlined several strategies to increase FA efficiency. We updated and revised those recommendations based on more recent findings reviewed in this study. We recommend: (a) using OVI as described in the methods of this review to interpret and identify a reasonable endpoint of FA (b) conducting a series of alone or ignore sessions at the start of an FA to efficiently test for the influence of automatic reinforcement prior to testing for social reinforcement; (c) using trial-based FAs when the goal is to efficiently test for individual functions and synthesized contingency analyses when the goal is to efficiently test for the presence of social reinforcement but not the specific source(s) of reinforcement; (d) that the response class tested in an FA should consist of one or a few behaviors, or if multiple topographies contact reinforcement in an FA, data from individual topographies should be graphed and analyzed both separately and in aggregate; (e) FA conditions should be sequenced to capitalize on EO influences from prior sessions; (f) each FA condition should be correlated with a unique discriminative stimulus to facilitate discriminated responding across conditions; (g) beginning with brief sessions (e.g., 3 min) and extending session duration only if rates of responding vary markedly across sessions and within conditions; (h) considering alternative data displays that allow analysis of within-session patterns of responding; (i) considering alternative measurement strategies (e.g., latency) when warranted; (j) testing for functional relations between problem behavior and tangible reinforcement only when preliminary information suggests a relation might exist; and (k) using a control condition that will control for all potential positive and negative reinforcement contingencies arranged in the test conditions.

Considerations for Future Research

Despite the breadth of research conducted in the area of FA generally and improvements in FA methods designed to increase FA efficiency specifically, we identified areas of future research that warrant consideration. Additional methods to improve efficiency might include enhancements in methods to eliminate problem behavior during the FA control condition, or eliminate problem behavior during the reinforcement interval of test conditions, where problem behavior should typically cease (Roane et al., 1999). Researchers also might consider procedural modifications such as conducting more test and fewer control sessions during a pairwise FA. For example, conducting three of the same test condition sequentially prior to conducting the control condition may capitalize on EOs for a given reinforcer, similar to that described by Hammond et al. (2013). Future researchers also might consider how additional stimulus control refinements (similar to Connors et al. 2000) could facilitate clearer differences between test and control conditions. Another area that requires further investigation related to efficiency is the extent to which milder forms of problem behavior could be assessed in an FA to preclude observation of more dangerous behaviors that require more time and resources to assess (e.g., Fritz, Iwata, Hammond, & Bloom, 2013). Finally,

future researchers might consider examining other variables that are likely to affect efficiency (e.g., incorporating caregivers or known environments into the analysis). As the assessment of severe problem behavior becomes increasingly sophisticated, methods to increase FA efficiency are likely to continue into the foreseeable future.

Acknowledgments

Grants 5R01HD079113 and 5R01HD083214 from the National Institute of Child Health and Human Development provided partial support for this research.

References

- American Psychological Association. (2007). Record keeping guidelines. *The American Psychologist*, 62, 993–1004. doi: 10.1037/0003-066X.62.9.993 [PubMed: 18085845]
- Arron K, Oliver C, Hall S, Sloneem J, Forman D, & McClintock K (2006). Effects of social context on social interaction and self-injurious behavior in Cornelia de Lange syndrome. *American Journal on Mental Retardation*, 111, 184–192. doi: 10.1352/0895-8017(2006)111[184.EOSCOS]2.0.CO;2 [PubMed: 16597185]
- Asmus JM, Franzese JC, Conroy MA, & Dozier CL (2003). Clarifying functional analysis outcomes for disruptive behaviors by controlling consequence delivery for stereotypy. *School Psychology Review*, 32, 624–630. <https://web.a.ebscohost.com/abstract?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=02796015&AN=11924952&h=PN AKIUG8Ih955lhACC9hY1yZAlf303mjTzhMEmlz1bh3nqhahdXfcNkCmgikcpMSq6pxaP%2bBoTdkJDGVitcfLg%3d%3d&crl=c&resultNs=AdminWebAuth&resultLocal=ErrCrlNotAuth&crlhashurl=login.aspx%3fdirect%3dtrue%26profile%3dehost%26scope%3dsite%26authtype%3dcrawler%26jrnl%3d02796015%26AN%3d11924952>
- Asmus JM, Ringdahl JE, Sellers JA, Call NA, Andelman MS, & Wacker DP (2004). Use of short-term inpatient model to evaluate aberrant behavior: Outcome data summaries from 1996 to 2001. *Journal of Applied Behavior Analysis*, 37, 283–304. doi: 10.1901/jaba.2004.37-283 [PubMed: 15529887]
- Austin JL, Groves EA, Reynish LC, & Francis LL (2015). Validating trial-based functional analyses in mainstream primary school classrooms. *Journal of Applied Behavior Analysis*, 48, 274–288. doi: 10.1002/jaba.208 [PubMed: 25916982]
- Baer DM (1977). Perhaps it would be better not to know everything. *Journal of Applied Behavior Analysis*, 10, 167–172. doi: 10.1901/jaba/1977.10-167 [PubMed: 16795545]
- Baer DM, Wolf MM, & Risley TR (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1, 91–97. doi: 10.1901/jaba.1987.20-313 [PubMed: 16795165]
- Beavers GA, & Iwata BA (2011). Prevalence of multiply controlled problem behavior. *Journal of Applied Behavior Analysis*, 44, 593–597. doi: 10.1901/jaba.2011.44-593 [PubMed: 21941388]
- Beavers GA, Iwata BA, & Lerman DC (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis*, 46, 1–21. doi: 10.1002/jaba.30 [PubMed: 24114081]
- Bachmeyer MH, Piazza CC, Frederick LD, Reed GK, Rivas KD, & Kadey HJ (2009). Functional analysis and treatment of multiply controlled inappropriate mealtime behavior. *Journal of Applied Behavior Analysis*, 42, 641–658. doi: 10.1901/jaba.2009.42-641. [PubMed: 20190924]
- Berkson G, & Mason WA (1963). Stereotyped movements of mental defectives: III. Situation effects. *American Journal of Mental Deficiency*, 68, 409–412. [PubMed: 14055764]
- Betz AM, & Fisher WW (2011). History and Methods In Fisher WW, Piazza CC, & Roane HS (Eds.), *Handbook of applied behavior analysis* (pp. 206–225). New York: The Guilford Press.
- Bloom SE, Iwata BA, Fritz JN, Roscoe EM, & Carreau AB (2011). Classroom application of trial-based functional analysis. *Journal of Applied Behavior Analysis*, 44, 19–31. doi: 10.1901/jaba.2011.44-19 [PubMed: 21541140]

- Carr EG, Newsom CD, & Binkoff JA (1976). Stimulus control of self-destructive behavior in a psychotic child. *Journal of Abnormal Child Psychology*, 4, 139–153. doi: 10.1007/BG00916518 [PubMed: 945811]
- Carr EG, Newsom CD, & Binkoff JA (1980). Escape as a factor in the aggressive behavior of two retarded children. *Journal of Applied Behavior Analysis*, 13, 101–117. doi: 10.1901/jaba.1980.13-101 [PubMed: 7189192]
- Carr EG, Yarbrough SC, & Langdon NA (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 [PubMed: 9433791]
- Conners J, Iwata BA, Kahng S, Hanley GP, Worsdell AS, & Thompson RH (2000). Differential responding in the presence and absence of discriminative stimuli during multielement functional analyses. *Journal of Applied Behavior Analysis*, 33, 299–308. doi: 10.1901/jaba.2000.33-299 [PubMed: 11051570]
- Cooper LJ, Wacker DP, Thursby D, Plagmann LA, Harding J... & Derby M (1992). Analysis of the effects of task preferences, task demands, and adult attention on child behavior in outpatient and classroom settings. *Journal of Applied Behavior Analysis*, 25, 823–840. doi: 10.1901/jaba.1992.25-823 [PubMed: 1478905]
- Day HM, Horner RH, & O'Neill RE (1994). Multiple functions of problem behaviors: Assessment and intervention. *Journal of Applied Behavior Analysis*, 27, 279–289. doi: 10.1901/jaba.1994.27-279 [PubMed: 8063626]
- Derby KM, Wacker DP, Peck S, SAsso GARY, DeRaad A, Berg W, ... & Ulrich S. (1994). Functional analysis of separate topographies of aberrant behavior. *Journal of Applied Behavior Analysis*, 27, 267–278. doi: 10.1901/jaba.1994.27-267 [PubMed: 8063625]
- Didden R, Duker PC, & Korzilius H (1997). Meta-analytic study on treatment effectiveness for problem behaviors with individuals who have mental retardation. *AJMR-American Journal on Mental Retardation*, 101, 387–399.
- Fahmie TA, Iwata BA, Querim AC, & Harper JM (2013). Test-specific control conditions for functional analyses. *Journal of Applied Behavior Analysis*, 46, 61–70. doi: 10.1002/jaba.9 [PubMed: 24114085]
- Ferron JM, Joo SH, Levin JR (2017). A Monte-Carlo evaluation of masked-visual analysis in response-guided versus fixed-criteria multiple-baseline designs. *Journal of Applied Behavior Analysis*, 50, 701–716. doi: 10.1002/jaba.410 [PubMed: 28887866]
- Fisher WW, Greer BD, Romani PW, Zangrillo AN, & Owen TM (2016). Comparisons of synthesized and individual reinforcement contingencies during functional analysis. *Journal of Applied Behavior Analysis*, 49, 596–616. doi: 10.1002/jaba.314 [PubMed: 27174563]
- Fisher WW, Kelley ME, & Lomas JE (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis*, 36, 387–406. doi: 10.1901/jaba.2003.36-387 [PubMed: 14596583]
- Fisher WW, Lindauer SE, Alterson CJ, & Thompson RH (1998). Assessment and treatment of destructive behavior maintained by stereotypic object manipulation. *Journal of Applied Behavior Analysis*, 31, 513–527. doi: 10.1901/jaba.1998.31-513 [PubMed: 9891391]
- Fritz JN, Iwata BA, Hammond JL, & Bloom SE (2013). Experimental analysis of precursors to severe problem behavior. *Journal of Applied Behavior Analysis*, 46, 101–129. doi: 10.1002/jaba.27 [PubMed: 24114089]
- Gerow S, Rispoli M, Gregori E, & Sanchez L (2019). Parent-implemented trial-based functional analysis for young children with ASD. *Focus on Autism and Other Developmental Disabilities*, 34, 29–40. doi: 10.1177/1088357618755695
- Greer BD, Mitteer DR, Briggs AM, Fisher WW, & Sodawasser AJ (provisionally accepted). On the accuracy of synthesized contingency analysis for determining behavioral function and the utility of the assessments informing synthesis. *Journal of Applied Behavior Analysis*.
- Hagopian LP, Fisher WW, Thompson RH, Owen-DeSchryver J, Iwata BA, & Wacker DP (1997). Toward the development of structured criteria for interpretation of functional analysis data. *Journal of Applied Behavior Analysis*, 30, 313–326. doi: 10.1901/jaba.1997.30-313 [PubMed: 9210309]

- Hagopian LP, Rooker GW, Jessel J, & DeLeon IG (2013). Initial functional analysis outcomes and modifications in pursuit of differentiation: A summary of 176 inpatient cases. *Journal of Applied Behavior Analysis*, 46, 88–100. doi: 10.1002/jaba.25 [PubMed: 24114088]
- Hagopian LP, Toole LM, Long ES, Bowman LG, & Lieving GA (2004). A comparison of dense-to-lean and fixed lean schedules of alternative reinforcement and extinction. *Journal of Applied Behavior Analysis*, 37, 323–337. doi: 10.1901/jaba.2004.37-323 [PubMed: 15529889]
- Hammond JL, Iwata BA, Rooker GW, Fritz JN, & Bloom SE (2013). Effects of fixed versus random condition sequencing during multielement functional analyses. *Journal of Applied Behavior Analysis*, 46, 22–30. doi: 10.1002/jaba.7 [PubMed: 24114082]
- Hanley GP (2012). Functional assessment of problem behavior: Dispelling myths, overcoming implementation obstacles, and developing new lore. *Behavior Analysis in Practice*, 5, 54–72. doi: 10.1007/BF03391818 [PubMed: 23326630]
- Hanley GP, Iwata BA, & McCord BE (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis*, 36, 147–185. doi: 10.1901/jaba.2003.36-147 [PubMed: 12858983]
- Hanley GP, Jin CS, Vanselow NR, & Hanratty LA (2014). Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments. *Journal of Applied Behavior Analysis*, 47, 16–36. doi: 10.1002/jaba.106 [PubMed: 24615474]
- Hays WL (1988). *Statistics* (4th ed). Fort Worth, TX: Holt, Rinehart and Winston.
- Huwaldt JA (2016, January 9). Plot Digitizer. Retrieved from <https://sourceforge.net/projects/plotdigitizer/>.
- Iwata BA, Dorsey MF, Slifer KJ, Bauman KE, & Richman GS (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 [PubMed: 8063622]
- Jessel J, Hanley GP, & Ghaemmaghami M (2016a). Interview-informed synthesized contingency analyses: Thirty replications and reanalysis. *Journal of Applied Behavior Analysis*, 49, 576–595. doi: 10.1002/jaba.316 [PubMed: 27174653]
- Jessel J, Hanley GP, & Ghaemmaghami M (2016b). On the efficiency of and control shown by different functional analysis formats: a literature review. Paper presented at the 42nd Annual Convention of the Association for Behavior Analysis International Chicago, IL May, 2016.
- Jessel J, Hanley GP, Ghaemmaghami M, & Metras R (2018). An evaluation of the single-session interview-informed synthesized contingency analysis. *Behavioral Interventions*, 34, 62–78. doi: 10.1002/bin.1650
- Johnston JM (1988). Strategic and tactical limits of comparison studies. *The Behavior Analyst*, 11, 1–9. doi: 10.1007/BF03392448 [PubMed: 22477989]
- Kodak T, Fisher WW, Paden A, & Dickes N (2013). Evaluation of the utility of a discrete-trial functional analysis in early intervention classrooms. *Journal of Applied Behavior Analysis*, 46, 301–306. doi: 10.1002/jaba.2 [PubMed: 24114103]
- Kuhn DE, DeLeon IG, Fisher WW, & Wilke AE (1999). Clarifying an ambiguous functional analysis with matched and mismatched extinction procedures. *Journal of Applied Behavior Analysis*, 32, 99–102. doi: 10.1901/jaba.1999.32-99 [PubMed: 10201106]
- Kurtz PF, Fodstad JC, Huete JM, & Hagopian LP (2013). Caregiver-and staff-conducted functional analysis outcomes: A summary of 52 cases. *Journal of Applied Behavior Analysis*, 46, 738–749. doi: 10.1002/jaba.87 [PubMed: 24114788]
- Lalli JS, Mace FC, Wohn T, & Livezey K (1995). Identification and modification of a response-class hierarchy. *Journal of Applied Behavior Analysis*, 28, 551–559. doi: 10.1901/jaba.1995.28-551 [PubMed: 16795881]
- Lee YY, Sugai G, & Horner RH (1999). Using an instructional intervention to reduce problem and off-task behaviors. *Journal of Positive Behavior Interventions*, 1, 195–204. doi: 10.1177/109830079900100402
- Lovaas OI, Freitag G, Gold VJ, & Kassorla IC (1965). Experimental studies in childhood schizophrenia: Analysis of self-destructive behavior. *Journal of Experimental Child Psychology*, 2, 67–84. doi: 10.1016/9922-0965(65)90016-0

- Lovaas OI, & Simmons JQ (1969). Manipulation of self-destruction in three retarded children. *Journal of Applied Behavior Analysis*, 2, 143–157. doi: 10.1901/jaba.1969.2-143 [PubMed: 16795215]
- Mace FC, & Lalli JS (1991). Linking descriptive and experimental analyses in the treatment of bizarre speech. *Journal of Applied Behavior Analysis*, 24, 553–562. doi: 10.1901/jaba.1991.24-553 [PubMed: 1752842]
- Mueller MM, Nkosi A, & Hine JF (2011). Functional analysis in public schools: A summary of 90 functional analyses. *Journal of Applied Behavior Analysis*, 44, 807–818. doi: 10.1901/jaba.2011.44-807 [PubMed: 22219531]
- Northup J, Wacker D, Sasso G, Steege M, Cigrand K, Cook J, & DeRaad A (1991). A brief functional analysis of aggressive and alternative behavior in an outclinic setting. *Journal of Applied Behavior Analysis*, 24, 509–522. doi: 10.1901/jaba.1991.24-509 [PubMed: 1752840]
- Piazza CC, Fisher WW, Brown KA, Shore BA, Patel MR, Katz RM, ... & Blakely-Smith A. (2003). Functional analysis of inappropriate mealtime behaviors. *Journal of Applied Behavior Analysis*, 36, 187–204. doi: 10.1901/jaba.2003.36-187 [PubMed: 12858984]
- Querim AC, Iwata BA, Roscoe EM, Schlichenmeyer KJ, Ortega JV, & Hurl KE (2013). Functional analysis screening for problem behavior maintained by automatic reinforcement. *Journal of Applied Behavior Analysis*, 46, 47–60. doi: 10.1002/jaba.26 [PubMed: 24114084]
- Repp A (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 [PubMed: 8063637]
- Retzlaff BJ, Fisher WW, Akers JA, & Greer BD (in press). A translational evaluation of potential iatrogenic effects of single and combined contingencies during functional analysis. *Journal of Applied Behavior Analysis*.
- Richman DM, Wacker DP, Asmus JM, Casey SD, & Andelman M (1999). Further analysis of problem behavior in response class hierarchies. *Journal of Applied Behavior Analysis*, 32, 269–283. doi: 10.1901/jaba.1999.32-269 [PubMed: 10513024]
- Roane HS, Fisher WW, Kelley ME, Mevers HL, & Boussein KJ (2013). Using modified visual-inspection criteria to interpret functional analysis outcomes. *Journal of Applied Behavior Analysis*, 46, 130–146. doi: 10.1002/jaba.13 [PubMed: 24114090]
- Roane HS, Lerman DC, Kelley ME, & Van Camp CM (1999). Within-session patterns of responding during functional analyses: The role of establishing operations in clarifying behavioral function. *Research in Developmental Disabilities*, 20, 73–89. doi: 10.1016/S0891-4222(98)00033-X [PubMed: 9987812]
- Rooker GW, Iwata BA, Harper JM, Fahmie TA, & Camp EM (2011). False-positive tangible outcomes of functional analyses. *Journal of Applied Behavior Analysis*, 44, 737–745. doi: 10.1901/jaba.2011.44-737 [PubMed: 22219526]
- Roscoe EM, Schlichenmeyer KJ, & Dube WV (2015). Functional analysis of problem behavior: A systematic approach for identifying idiosyncratic variables. *Journal of Applied Behavior Analysis*, 48, 289–314. doi: 10.1002/jaba.201 [PubMed: 25930176]
- Saini V, Fisher WW, & Retzlaff BJ, (2018). Predictive validity and efficiency of ongoing visual-inspection criteria for interpreting functional analyses. *Journal of Applied Behavior Analysis*, 51, 303–320. doi: 10.1002/jaba.450 [PubMed: 29527741]
- Shirley MJ, Iwata BA, & Kahng S (1999). False-positive maintenance of self-injurious behavior by access to tangible reinforcers. *Journal of Applied Behavior Analysis*, 32, 201–204. doi: 10.1901/jaba.1999.32-201 [PubMed: 10396773]
- Skinner BF (1953). *Science and human behavior*. New York, NY: Simon and Schuster.
- Smith RG, Iwata BA, Vollmer TR, & Zarcone JR (1992). Experimental analysis and treatment of multiply controlled self-injury. *Journal of Applied Behavior Analysis*, 26, 183–196. doi: 10.1901/jaba.1993.26-183
- Thomason-Sassi JL, Iwata BA, & Fritz JN (2013). Therapist and setting influences on functional analysis outcomes. *Journal of Applied Behavior Analysis*, 46, 79–87. doi: 10.1002/jaba.28 [PubMed: 24114087]

- Thomason-Sassi JL, Iwata BA, Neidert PL, & Roscoe EM (2011). Response latency as an index of response strength during functional analyses of problem behavior. *Journal of Applied Behavior Analysis*, 44, 51–67. doi: 10.1901/jaba.2011.44-51 [PubMed: 21541141]
- Thompson RH, Fisher WW, Piazza CC, & Kuhn DE (1998). The evaluation and treatment of aggression maintained by attention and automatic reinforcement. *Journal of Applied Behavior Analysis*, 31, 103–116. doi: 10.1901/jaba.1998.31-103 [PubMed: 9532754]
- Thompson RH, & Iwata BA (2005). A review of reinforcement control procedures. *Journal of Applied Behavior Analysis*, 38, 257–278. doi: 10.1901/jaba.2005.176-03 [PubMed: 16033174]
- Tiger JH, Hanley GP, & Bruzek J (2008). Functional communication training: A review and practical guide. *Behavior Analysis in Practice*, 1, 16–23. doi: 10.1007/BF03391716 [PubMed: 22477675]
- Van Houten R (1987). Comparing treatment techniques: A cautionary note. *Journal of Applied Behavior Analysis*, 20, 109–110. doi: 10.1901/jaba.1987.20-109 [PubMed: 3583961]
- Vollmer TR, Iwata BA, Zarcone JR, Smith RG, & Mazaleski JL (1993). Within-session patterns of self-injury as indicators of behavioral function. *Research in Developmental Disabilities*, 14, 479–492. doi: 10.1016/0891-4222(93)90039-M [PubMed: 8296027]
- Vollmer TR, Marcus BA, Ringdahl JE, & Roane HS (1995). Progressing from brief assessments to extended experimental analyses in the evaluation of aberrant behavior. *Journal of Applied Behavior Analysis*, 28, 561–576. doi: 10.1901/jaba.1995.28-561 [PubMed: 16795882]
- Wacker DP, Lee JF, Dalmau YCP, Kopelman TG, Lindgren SD, Kuhle J, ... & Waldron DB. (2013). Conducting functional analyses of problem behavior via telehealth. *Journal of Applied Behavior Analysis*, 46, 31–46. doi: 10.1002/jaba.29 [PubMed: 24114083]

Figure 1.

Hypothetical datasets for a brief functional analysis (FA; top panel), trial-based FA (second panel), reversal design (third panel), and multielement FA (bottom panel).

Figure 2.

The mean duration in minutes required for the assessment (black bars) and per function tested (white bars) across functional analysis (FA) types. Error bars represent one standard deviation above the mean.

Table 1.

The number of data points in the functional analysis (FA) and the corresponding number of data points that needed to be consistent with the presence of a function in order to consider the FA differentiated using the binomial distribution.

Total Number of Sessions	Number of Sessions Needed to be Consistent with a Function for an FA to be Considered Differentiated
5	5 (p =.03)
6	6 (p =.02)
7	6 (p =.05)
8	7 (p =.03)
9	8 (p =.02)
10	8 (p =.04)
11	9 (p = .03)
12	9 (p =.05)
13	10 (p =.03)
14	11 (p =.02)
15	11 (p =.04)
16	12 (p =.03)
17	12 (p =.05)
18	13 (p =.03)
19	13 (p =.05)
20	14 (p =.04)
21	14 (p =.05)

Table 2.

Means (*M*) and standard deviations (*SD*) for the number and duration of sessions or trials required to complete the ongoing-visual inspection for various functional analysis (FA) types reported on a per-FA and per-function basis.

FA Type	<i>M</i> Sessions or Trials	<i>M</i> duration (min)	<i>M</i> Sessions or Trials per Function	<i>M</i> Duration (min) per Function
Brief FA	7.2 (<i>SD</i> = 0.6)	72.0 (<i>SD</i> = 5.6)	7.2 (<i>SD</i> = 0.6)	72.0 (<i>SD</i> = 5.6)
Trial-Based FA	41.6 (<i>SD</i> = 23.5)	59.6 (<i>SD</i> = 23.3)	13.0 (<i>SD</i> = 4.3)	19.5 (<i>SD</i> = 5.5)
Reversal	27.6 (<i>SD</i> = 12.3)	281.7 (<i>SD</i> = 171.6)	15.4 (<i>SD</i> = 5.6)	163.3 (<i>SD</i> = 140.9)
Multielement	16.6 (<i>SD</i> = 10.2)	150.6 (<i>SD</i> = 165.4)	5.3 (<i>SD</i> = 3.4)	47.9 (<i>SD</i> = 55.5)
Pairwise	8.9 (<i>SD</i> = 6.0)	82.2 (<i>SD</i> = 70.7)	8.9 (<i>SD</i> = 6.0)	82.2 (<i>SD</i> = 70.7)
Synthesized Contingency Analysis	5.2 (<i>SD</i> = 0.7)	25.1 (<i>SD</i> = 11.7)	5.2 (<i>SD</i> = 0.7)	25.1 (<i>SD</i> = 11.7)

Table 3.

Exact agreement between ongoing-visual inspection (OVI) and (a) the post-hoc interpretation made using the Roane et al. (2013) criteria on each full functional analysis (FA) and (b) the post-hoc interpretations provided by the authors of the original article.

FA Type	Exact Agreement Between OVI and Structured Criteria Applied to the Entire FA Dataset	Exact Agreement Between OVI and Authors' Interpretation from the Entire FA Dataset
Brief FA	100%	86.7%
Trial-Based FA	90.0%	90.0%
Reversal	100%	93.8%
Multielement	94.3%	84.8%
Pairwise	98.6%	91.7%
Synthesized Contingency Analysis	100%	92.5%