

PROGRESSING TOWARD DATA INTIMACY: A REVIEW OF
WITHIN-SESSION DATA ANALYSIS

TARA A. FAHMIE AND GREGORY P. HANLEY

UNIVERSITY OF KANSAS

Visual inspection of data is a common method for understanding, responding to, and communicating important behavior–environment relations in single-subject research. In a field that was once dominated by cumulative, moment-to-moment records of behavior, a number of graphic forms currently exist that aggregate data into larger units. In this paper, we describe the continuum of aggregation that ranges from distant to intimate displays of behavioral data. To aid in an understanding of the conditions under which a more intimate analysis is warranted (i.e., one that provides a richer analysis than that provided by condition or session aggregates), we review a sample of research articles for which within-session data depiction has enhanced the visual analysis of applied behavioral research.

DESCRIPTORS: binned analyses, data analysis, visual inspection, within-session data

The depiction of data in graphic form is a conventional method for describing and analyzing behavior and is a hallmark of our scientific discipline (Cooper, Heron, & Howard, 1987; Johnston & Pennypacker, 1993; Kazdin, 1982). There exists a reciprocal interaction between the experimenter and his or her data; graphic analysis streamlined this interaction by providing an instantaneous means of visually depicting and responding to behavior–environment relations. B. F. Skinner recognized the importance of visual access to behavior patterns and championed the cumulative record as a graphical standard in the early development of the field of behavior analysis (e.g., Ferster & Skinner, 1957). The invention of the cumulative recorder permitted the automatic depiction of the behavior of a single organism at specific and continuous moments in real time. Skinner's interest in the analysis of a molecular display of ongoing behavior patterns was shared by the basic research community. Cumulative response records that depicted responses in real time were the dominant graphical form in the first issue of

the *Journal of the Experimental Analysis of Behavior* (JEAB; 1958). However, only 18 years later, Skinner (1976) anticipated publication of the first cumulative-record-free issue of JEAB in an editorial entitled, "Farewell, My Lovely!" He wrote,

What has happened to experiments where rate changed from moment to moment in interesting ways, where a cumulative record told more at a glance than could be described in a page? Straight lines and steady states are no doubt important, but something is lost when one must reach a steady state before an experiment begins. There was a special kind of orderliness in a smooth curve lasting a few minutes or at most an hour. It suggested a really extraordinary degree of control over an individual organism as it lived its life from moment to moment. (p. 218)

In 1979, Poling published a graph of the number of figures in JEAB that displayed a cumulative record. The sharp declining slope of this graph led him to conclude that "a collective fondness for a direct and graphic depicter of ongoing behavior—the cumulative record—has faded" (p. 126).

A parallel cannot be traced in the graphic history of the *Journal of Applied Behavior Analysis* (JABA). The first issue of JABA (1968) displayed a mélange of graphs, ranging from within-session cumulative records to aggregated line and bar graphs. The collection

Address correspondence to Gregory P. Hanley, Western New England College, Psychology Department, 1215 Wilbraham Road, Springfield, Massachusetts 01119 (e-mail: ghanley@wnec.edu).

doi: 10.1901/jaba.2008.41-319

of moment-to-moment data was likely more difficult in applied settings where, instead of machines, humans observed and recorded instances of behavior. Due to an understanding that particular conditions controlled the likelihood of socially important repertoires (i.e., the notion of stimulus control) and the obvious difficulties of continuous data collection throughout each day, samples of behavior in relevant environments were collected. These sample or observation periods commenced when observers began recording behavior and ceased when observers stopped, thus delimiting the somewhat arbitrary boundaries of the familiar "session." After termination of a session, entering summary figures into a spreadsheet likely became the most convenient way to transduce data.

Beyond the pragmatic reasons just noted, data aggregated into single-session bins are perhaps a popular form of graphic depiction in applied behavior analysis due to a researcher's ability to detect changes in the level, trend, and variability in the data and the reinforcing properties of order that emerge in the data following such aggregation. In addition, the popularity of discontinuous recording procedures in applied behavior analysis (e.g., whole- and partial-interval recording) undoubtedly compromises the depiction of behavior-environment interactions at specific moments in time.

Different types of graphic displays have proven important for organizing, evaluating, and disseminating experimental results in many scientific disciplines, which is evidenced by the high proportion of graph-to-text space in scientific journals (Best, Smith, & Stubbs, 2001). Parameters of appropriate graph structure (e.g., bar graphs, line graphs) have been clearly delineated by the American Psychological Association (2001; see section 3.77) and multiple research-methods textbooks (e.g., Bailey & Burch, 2002; Cooper et al., 1987; Johnston & Pennypacker, 1993; Meyers &

Grossen, 1974). However, the manner in which data are aggregated *before* transforming them into a visual display serves an equally influential role in data analysis and has received considerably less attention. Each variation of data aggregation slightly (or sometimes, grossly) changes a graph's function as a stimulus for visual analysis. These variations can be superimposed on a continuum that ranges from very distant displays to very intimate displays (Figure 1). The terms *distant* and *intimate* are used throughout this paper as metaphors that represent the amount of time over which responding is summarized and refer to the proximity of the display to that of the raw, moment-to-moment data. Displays on the far left of the continuum are distant from the raw data in that data are aggregated into larger units of time such as condition or phase means. Moving right along the continuum, displays are more representative of raw data in that units are aggregated into smaller units of time, such as session means. Displays at the rightmost point on the continuum depict raw data and are thus intimate displays of the smallest measurement unit in applied work (i.e., 1 s).

There are no standard rules governing the selection of any particular graphic display; each point along the continuum offers relative benefits and limitations. Johnston and Pennypacker (1993) refer to this as the "paradox of summarization: It can simultaneously obscure and reveal important features of the data" (p. 303). Although it would be valuable to describe the conditions under which different levels of data depiction enhance a visual analysis, the seamless nature of the continuum makes this task very difficult. Instead, we offer three general rules: (a) Data depiction should occur at the points along the continuum that best support an understanding of important relations, (b) data depiction should not distort important relations, and (c) a researcher should combine units when their division adds superfluous detail.

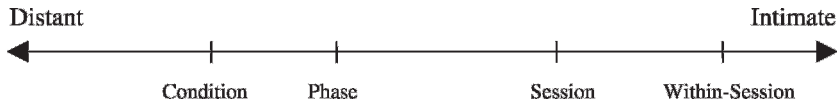


Figure 1. A continuum of data depiction.

Visually depicting behavioral data at the session level continues to be the most popular method of data display in *JABA*. By counting the number of different types of data display in successively larger random samples of single-subject empirical articles published in *JABA* between 1996 and 2006 (Volumes 29 to 39), we determined that session-by-session data far outnumbered any other forms of data display. Although we conducted an analysis of a range of sample sizes (i.e., 5, 10, 15, 20, 25, 30, and 35 articles), we ceased sampling at 35 articles due to the similar conclusions evident across the last three sample sizes. Figure 2 presents the distribution of displays among a sample of 35 articles and shows that 13% of displays included a depiction of data aggregated at the condition level, 3% at the phase level, 77% at the session level, and 8% at the within-session level.

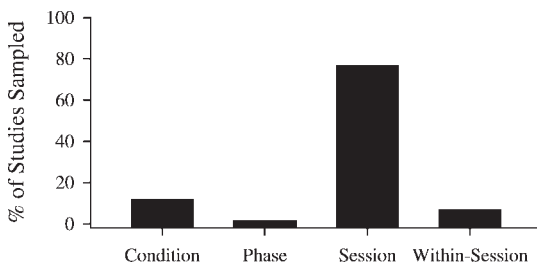


Figure 2. The percentage of displays that aggregated data at the within-session, session, phase, and condition levels in a random sample of 35 single-subject empirical articles published in *JABA* between 1996 and 2006. Condition aggregations were defined as the depiction of data combined across each session of a given condition. Phase aggregations were defined as the depiction of data combined across each session of a given phase. Session aggregations were defined as the depiction of data combined across time units (e.g., minutes, hours) in a single session. Within-session data were defined by the depiction of single time units (e.g., seconds, minutes) within a single session.

Aggregating data into larger units is a common practice in psychology (e.g., condition and group means with standard deviations), and although a discussion of when behavior analysts should consider presenting evidence of important behavior–environment relations in this aggregated form is valuable to some extent, it will not be discussed further in this paper. By contrast, a description of the conditions that have evoked the use of within-session data displays seems warranted because (a) the field of behavior analysis was founded on intimate displays of data, and (b) despite recommendations to explore the conditions under which within-session data depictions are advantageous (Vollmer, Iwata, Zarcone, Smith, & Mazaleski, 1993), such a review has yet to be conducted. In addition, within-session analyses are currently limited in our flagship journal (see Figure 2), so readers may be unaware of this important method of data analysis. Thus, the purpose of this paper is to identify the conditions under which intimate displays of data should be considered by reviewing the successful adoption of within-session data analysis in published applied behavioral research and describing the conditions under which intimate data displays should be considered in future studies.

ARTICLE INCLUSION CRITERIA

The present review includes studies that (a) provide a unique application of a within-session analysis, (b) serve as an illustrative example of a more common application of a within-session analysis, or (c) suggest a novel future application for a within-session analysis. Appropriate articles were preliminarily identified through a search of *PsycInfo*, *Google Scholar*, and the *JABA* Web site using the following keywords: *within*,

within-session, bin, binned analysis, binned analyses, cumulative, moment, moment-to-moment, and moment-by-moment. Only studies that were both applied and behavioral were retained from the search (Baer, Wolf, & Risley, 1968). Additional articles were identified through an examination of the reference section of each obtained article.

CONDITIONS FOR THE WITHIN-SESSION ANALYSIS OF BEHAVIOR

To Describe Naturally Occurring Behavioral Relations

Within-session analysis of data collected via descriptive assessments is common. Descriptive assessments are characterized by the observation and measurement of behavior in relevant contexts without systematic manipulation of features of that context. The common aim of a descriptive assessment is to understand naturally occurring behavior–environment relations. Within-session inspection of data is necessary for identifying these relations, which are often expressed as conditional probabilities.

Vollmer, Borrero, Wright, Van Camp, and Lalli (2001) described a procedure for calculating both conditional and background probabilities with the aim of identifying the strength of social reinforcement contingencies in simulated environments involving children and their caregivers. Data were collected on problem behaviors, potential establishing operations, and potential reinforcers during sessions of varying lengths. Conditional probabilities were then calculated by imposing time windows onto individual data streams, centered on the occurrence of problem behavior. Occurrences and nonoccurrences of a putative reinforcer, as well as proportion of time with a putative reinforcer, were summarized by conditional probability calculations. Background probabilities were calculated by randomly selecting a number of single points in time from the data stream and creating similar windows around these points. These within-session estimations

of contiguous and noncontiguous reinforcer delivery provided information regarding the possible contingencies of reinforcement operating in the simulated environments. Similar types of within-session analyses have been used to determine the prevalence of social consequences delivered by preschool teachers (McKerchar & Thompson, 2004) and to detect potential contingencies between caregiver reprimands and child problem behaviors (Sloman et al., 2005).

To Determine Behavioral Function via Direct Assessment

The benefits of pretreatment functional analyses (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994) for developing interventions for severe forms of problem behavior have been demonstrated (Carr & Durand, 1985; Iwata, Pace, Cowdery, & Miltenberger, 1994). The widespread implementation of functional analyses, which involves repeated measures of problem behavior under multiple conditions in which some environmental event is systematically manipulated, has been established (Hanley, Iwata, & McCord, 2003). Nevertheless, time constraints typical of many clinical settings and consultation models make it difficult to collect sufficient data to determine behavioral function (Kahng & Iwata, 1999; Northup et al., 1991; Vollmer et al., 1993). This concern was first addressed by Northup et al., who developed a set of brief functional analysis and treatment procedures that partly relied on within-session analyses of behavior.

Northup et al. (1991) initially examined session-by-session rates of the aggression of 3 individuals across single sessions of multiple conditions. A treatment condition was then initiated in which prompted and unprompted emission of an appropriate alternative behavior was differentially reinforced and extinction was arranged for aggression. The authors were able to show functional control of the contingency reversal on problem behavior using minute-by-minute analyses that required 20 to 22 min per

participant. These methods allowed direct assessment and treatment evaluations of problem behavior to occur within the 90 min allotted for outpatient visits. The assertions regarding behavioral function were somewhat compromised due to the reliance on single data points per condition for 2 participants (i.e., no replication) and the continued prompting of the alternative response during the contingency reversal phase for all participants. Nevertheless, this study was important because of its description of brief assessments for determining maintaining reinforcers and within-session analyses for identifying effective treatments for problem behavior.

Vollmer et al. (1993) addressed the limitation of using single points of data to infer behavioral function by examining the extent to which minute-by-minute analyses of single sessions matched the results of full functional analyses in which multiple session means per condition were analyzed. In three of the four cases the authors examined, results of a within-session analysis were consistent with the results of the full functional analysis. In a fourth case, within-session patterns suggested an extinction-induced effect across control conditions that could not be detected in the more aggregated measures of the full assessment.

Kahng and Iwata (1999) also examined the correspondence between within-session analyses of brief assessments (Northup et al., 1991), session-by-session analyses of brief assessments, and session-by-session analyses of full assessments (Iwata et al., 1982/1994) across 50 sets of data. Within-session data analyses of brief assessments resulted in the same interpretation of behavioral function as did session-by-session analysis of full assessments in 68% of cases. More relevant to our review was that there was one case in which the within-session analysis of the brief assessment conferred advantage over the session-by-session analysis of the brief assessment data. For another case, however, within-session data analysis obscured the inter-

pretation of the brief assessment. Thus, the authors concluded that within-session data conferred both advantages and disadvantages and recommended the use of both levels of analysis to determine behavioral function when brief assessments are unavoidable.

In an attempt to provide additional solutions for conditions under which time constraints limit the ability to conduct full functional analyses, Vollmer, Marcus, Ringdahl, and Roane (1995) used within-session analyses as a starting point on which to base decisions to conduct more extended analyses. Their model progressed from brief to extended analyses. The initial assessment, a within-session analysis of brief assessment data, resulted in differentiated outcomes in 6 of 20 cases. Their results suggest that decisions about the length of behavioral assessment can be based on data and may begin with the inspection of within-session patterns of responding during brief assessments.

Motivating operations, defined as alterations of the value of specific consequences and probabilities of their related responses (Laraway, Snyckerski, Michael, & Poling, 2003; Michael, 1982), may cause transient changes in responding within a functional analysis session, resulting in undifferentiated patterns of responding (i.e., similar rates of problem behavior across multiple or all conditions). Such undifferentiated patterns can be interpreted in at least two ways: (a) The target behavior is multiply controlled, or (b) the target behavior is maintained by automatic reinforcement. Roane, Lerman, Kelley, and Van Camp (1999) used a unique within-session analysis to clarify undifferentiated functional analysis results. When problem behavior was observed to be elevated across multiple conditions, the authors compared the number (or duration) of responses that occurred during the presence of a specific event programmed as a consequence (e.g., attention delivery in the attention condition) to those that occurred during the absence of the same event. When responding was consistently

observed in the absence of the programmed consequence, it was inferred that deprivation from that consequence was establishing the value of the consequence and thus evoking elevated rates of the problem behavior. When this pattern was observed across multiple conditions, support was demonstrated for the interpretation that behavior was sensitive to multiple forms of reinforcement as opposed to automatic reinforcement. By contrast, when the occurrence of problem behavior was more equally distributed among periods in the presence and in the absence of the target consequences, the behavior was hypothesized to be maintained by automatic reinforcement. The same conclusions regarding behavioral function could have been made in the absence of the within-session data analysis for all 4 participants (i.e., this study did not demonstrate the necessity of the within-session analysis). Nevertheless, the tactic of analyzing within-session data partitioned into time periods marked by the presence or absence of a target motivating operation holds promise as a means for clarifying undifferentiated functional analysis results.

To Detect Within-Session Trends

When environmental changes are not specifically programmed to occur in a session, within-session trends in behavior may be apparent due to behavioral processes that are associated with the passage of time (e.g., habituation, deprivation, fatigue). Understanding how these trends vary across individuals may assist in creating individualized treatment plans. Smith, Iwata, Goh, and Shore (1995) found three general within-session patterns of escape behavior during 15-min observations in which task demands were presented every 30 s. Of the 5 individuals who participated, 2 showed acceleration of escape behavior across time, 2 showed slight deceleration of escape behavior across time, and 1 showed no consistent within-session pattern. The value of these findings was discussed in light of their clinical implications. For instance, frequent but brief skill-training

sessions would be better suited to individuals for whom long durations of task demands increase the probability of escape. By contrast, less frequent but extended sessions of skill training would be better suited to individuals for whom escape is more likely to be observed at the beginning of training sessions. In sum, analysis of within-session patterns allow individualized treatments to be prescribed.

Empirical studies of appropriate behavior (academic responses, social skills, athletic repertoires) often measure parameters of acquisition and maintenance that may be most clearly depicted at the within-session level. Consistent with this notion, Daly, Martens, Hamler, Dool, and Eckert (1999) noted in their study on improving oral reading fluency that the measurement of within-session endurance is rare but may prove germane to the understanding of fluent performances. Martens, Bradley, and Eckert (1997) provided a rare example of a within-session analysis of desirable behavior in their evaluation of the effect of various reinforcement histories on persistence of student engagement in work tasks. The authors arranged three conditions that consisted of 2 min of an experimental manipulation (praise only, praise and redirection, or praise and positive attention) followed by 8 min of extinction (i.e., no interaction). The authors visually depicted the mean level of engagement during each minute of the 8-min extinction period to evaluate persistence. For both participants, engagement persisted following a brief history of praise contingent on engagement and redirection contingent on off-task behavior. That is, the slope of the line depicting engagement was either flat or positively accelerated across the 8 min that followed this experimental manipulation. By contrast, for both participants, engagement did not persist following a brief history of praise contingent on either engagement or off-task behavior. In this condition, the slope of the line depicting engagement was negative.

To Safeguard Clients

Within-session analyses of brief assessments are beneficial if prolonged assessment periods pose harmful risks to the health and safety of an individual. For instance, Van Houten (1993) analyzed minute-by-minute data to determine the impact of wrist weights on reducing an individual's severe self-injurious behavior (face slapping). During a single 35-min observation, Van Houten alternated the presence and absence of wrist weights within 5-min periods. A minute-by-minute data display showed that the presence of wrist weights was consistently associated with lower levels of face slapping. He noted that the intentionally brief assessment attenuated risks associated with the high-frequency face slapping.

To Create Sufficient Data for Analysis

If a session, phase, or condition is unexpectedly truncated, additional data generated by analyzing within-session patterns may be a practical tool to confirm appropriate conclusions and may also be necessary for a scientifically sound analysis. Coleman and Holmes (1998) presented session-by-session data in a multiple baseline design across participants, but added an inset for 1 participant who had an abbreviated number of treatment sessions. The inset contained a minute-by-minute depiction of disruptive behavior during the three treatment sessions and was added to support the assertion that the treatment (in this case, noncontingent escape) reduced disruptive behavior. The inclusion of the within-session analysis did not seem to be essential because the effect of the treatment on disruptive behavior was apparent in the session-by-session data. Nevertheless, the tactic to include a within-session analysis to evaluate whether a socially important behavior was influenced by a treatment when assessment conditions were prematurely terminated is unique to this study and represents a viable strategy of reanalysis for applied researchers. Brief analyses represent a limited sample of

behavior and thus inherently threaten the accuracy of steady-state conclusions (for a more detailed discussion of steady-state identification, see Johnston & Pennypacker, 1993, pp. 201–216). It is important to note that this strategy of including within-session data of truncated analyses is not recommended to replace the more rigorous practice of conducting conditions until stability is detected, but rather may be useful for enhancing an analysis that is unavoidably brief but that otherwise clearly demonstrates experimental control.

To Determine Observation Session Duration

The process of binning data into smaller units necessarily creates more single snapshots of data in a display. Roane, Vollmer, Ringdahl, and Marcus (1998) therefore suggested that within-session data analyses be used to establish optimal session durations by determining the point at which further experience with two or more experimental conditions would yield no additional information regarding the variables that control behavior. Mueller, Piazza, Patel, Kelley, and Pruett (2004) used this strategy to determine the optimal session duration necessary to assess and treat severe food refusal in a pediatric day-treatment program. Although the within-session analysis was not depicted in their study, Mueller et al. reported that it supported a decision to adopt a 1-hr rather than a 2-hr session duration and thus allowed them to conduct a greater number of treatment sessions in a similar time period. Future researchers would benefit from a description and evaluation of a systematic method that could be used to determine optimal session duration (see Mudford, Beale, & Singh, 1990, for some guidance).

To Clarify Counterintuitive Response Patterns

Within-session analyses have also been useful for clarifying counterintuitive changes in socially important behavior. For example, Vollmer, Ringdahl, Roane, and Marcus (1997) attempted to reduce the aggressive behavior of a 13-year-

old girl whose aggression was shown to be sensitive to tangible reinforcement (e.g., access to magazines). The treatment initially involved noncontingent reinforcement (NCR; continuous access to preferred magazines in this case), which eliminated aggression. The schedule of reinforcer delivery was then thinned by introducing increasingly longer intervals between deliveries. During schedule thinning, aggression was observed to increase above baseline levels. A within-session analysis revealed that reinforcer delivery was coinciding with larger and larger bouts of aggression (i.e., adventitious reinforcement was apparent), which most likely accounted for the treatment failure. A brief omission contingency was included with the reinforcement thinning and was effective at quickly reducing the adventitiously reinforced aggression to below baseline levels. Within-session analyses proved to clarify the counterintuitive effects of NCR while also allowing an effective intervention to be identified.

To Understand Behavioral Processes

The underlying processes affecting behavior change are not always apparent in aggregated data displays. Within-session analyses have been used for distinguishing between two or more behavioral processes because different processes are often associated with unique within-session patterns of responding. For example, Kahng, Iwata, Thompson, and Hanley (2000) described a method for identifying the behavioral processes involved in the reduction of a target behavior via NCR schedules. The researchers parceled out satiation and extinction effects by analyzing within-session patterns of responding during an extinction component that followed an NCR component. An elevated level of within-session responding during the brief extinction components would have indicated that NCR reduced behavior via satiation, whereas continued low levels of responding during the extinction component was indicative of an extinction process occurring during NCR. The results of this study revealed that the

operative processes of NCR were idiosyncratic in that it appeared that satiation, extinction, or both were operating during NCR for different participants.

Analyses of differential responding to schedules that alternate within a session (i.e., compound schedules) have also revealed idiosyncratic response patterns that have both conceptual and practical implications. For instance, Hanley, Iwata, and Thompson (2001) analyzed the within-session patterns of alternative responding during reinforcement and extinction components of mixed and multiple schedules. These two particular components alternated as part of a schedule-thinning procedure aimed at maintaining moderate levels of newly acquired responses that were functionally equivalent to severe problem behavior. The reinforcement and extinction components alternated either with (multiple schedule) or without (mixed schedule) schedule-correlated stimuli. Highly variable responding occurred during the mixed schedule, and within-session analyses revealed two undesirable but highly discrepant patterns of responding that produced the variability. One pattern involved responding toward the beginning of an extinction component, leading to a subsequent suppression of that response across the entire session. The second pattern involved a high rate of responding toward the end of an extinction component that culminated in the reinforcement component, thereby maintaining high response rates throughout all subsequent extinction components. Neither of these undesirable patterns was evident within multiple-schedule sessions; instead, responding generally occurred when and only when reinforcement was available. The within-session analyses identified the behavioral processes occurring within different applications of compound schedules (extinction, intermittent reinforcement, and stimulus control) while underscoring the importance of schedule-correlated stimuli in schedule-thinning procedures.

CONCLUSIONS

The analysis of behavior–environment relations relies on the behavior analyst’s ability to observe, record, display, and inspect data. The options for graphing data are numerous, and no standard exists for the manipulation of raw data prior to its inclusion in a visual display. Given that slight variations in levels of data aggregation may evoke or occasion different interpretations and conclusions, it is important to make informed decisions regarding the most appropriate way to analyze and depict data. The present article reviewed several conditions under which intimate data analyses have been beneficial to descriptive and experimental assessment procedures, treatment decisions, and conceptual accounts of applied behavioral research.

To this point, we have not yet highlighted the difference between a within-session analysis and a within-session depiction of data because the terms can be used somewhat interchangeably; nonetheless, the former depends on the latter. We do find it necessary, however, to consider the different conditions under which intimate depiction and analysis should be *conducted* and those under which they should be *shared* (e.g., used to communicate behavior–environment relations in published studies). Within-session analyses can be quite laborious. It is unreasonable to expect all researchers to intimately inspect every stream of raw data, and we do not wish to imply that it is necessary to do so. However, more harm can be done in remaining at one point along the continuum (especially at a point on the distant end) than can be done in thoroughly analyzing data at several points. Thus, the practice of analyzing data at different points along the continuum of aggregation is highly recommended throughout the process of any behavioral analysis (Johnston & Pennypacker, 1993). However, the conditions under which within-session analyses have proved most important are when brief assessments produce inadequate data displays that hide or obscure the important dimensions of

level, trend, and variability (e.g., Kahng & Iwata, 1999).

In regards to data depiction, it is important to acknowledge that the dissemination of research inherently involves the analytic behavior of multiple readers of various backgrounds. Researchers should take steps to ensure inclusion of data displays that most efficiently and effectively lead to valid conclusions. Unless the purpose of an article is to evaluate a particular within-session analysis procedure (e.g., Roane et al., 1999), we recommend that researchers publish or present within-session data only when their inclusion confers a distinct advantage to the reader as a fellow analyst. In other words, an affirmative response to the question “Will the analyst arrive at a different conclusion given the addition of the within-session data?” is necessary for within-session data to be shared. For instance, the within-session data sets produced by Vollmer et al. (1993) brought to light important temporal relations that were hidden in the aggregated data, and the inclusion of the within-session data in Hanley et al. (2001) provided an explanation for variability observed in the aggregated data and allowed stimulus control to be identified. By contrast, if the inclusion of within-session data makes visual inspection less efficient by adding redundant detail to the article, it should be omitted. In summary, the conditions under which a within-session data display should be used to communicate behavioral relations are much more limited than the conditions under which a within-session analysis should occur.

Our review of applied behavioral research studies revealed the following conditions under which intimate data displays are likely to be particularly useful: (a) to describe naturally occurring behavioral relations, (b) to determine behavioral function via direct assessment, (c) to detect within-session trends, (d) to safeguard clients, (e) to create sufficient data for analysis, (f) to determine observation session duration, (g) to clarify counterintuitive response patterns,

and (h) to understand behavioral processes. We will close with some additional and perhaps novel applications of intimate data analysis that may be helpful for future behavioral analyses.

Within-session analyses are promising for understanding behavioral processes in different types of interventions. For example, Kahng et al. (2000) suggested that depicting data in small, within-session bins may aid in the identification of the underlying behavioral processes that operate in differential-reinforcement-of-other-behavior schedules. In addition, a within-session data analysis may be useful for understanding the effects of motivating operations on responding in natural settings. Although the inspection of within-session patterns of responding has evoked speculation about potential fluctuations in establishing operations (Lalli, Mace, Livezey, & Kates, 1998; Lindberg, Iwata, Roscoe, Worsdell, & Hanley, 2003; Smith et al., 1995; Vollmer et al., 2001), actual cases of within-session manipulation and analysis of establishing operations are rare. Future investigation of the utility of within-session analyses to help to identify underlying behavioral processes in common treatment procedures and to individualize treatment plans is necessary, especially in light of multiple demonstrations of the idiosyncratic effects different schedules of consequences may have across individuals (Hanley et al., 2001; Kahng et al.; Smith et al., 1995).

The movement towards more efficient behavioral assessments is prevalent and important due to both practical constraints (e.g., time-limited clinical assessments) and ethical considerations (e.g., to expedite treatment). It seems likely that experimental validity may at times be sacrificed for the sake of efficiency. This need not be the case. Within-session analysis of data yielded from brief assessments, whether used to determine behavioral function (Northup et al., 1991; Vollmer et al., 1993) or treatment efficacy (Van Houten, 1993), may allow efficient *and* valid behavioral assessments to occur.

The multielement design, in which two or more conditions are rapidly alternated, is considered another efficient means to demonstrate functional relations between behavior and environmental conditions (Hains & Baer, 1989; Ulman & Sulzer-Azaroff, 1975), but the sequence of conditions may affect behavior in lawful ways and skew an analysis of the independent effects of each condition. Hains and Baer proposed multiple ways to modify multielement designs to display and analyze these carryover effects, but these modifications necessarily extend the time required for the assessment. Perhaps because requiring more time is antithetical to the goal of a multielement design (i.e., efficiency), the authors end the article with a poignant reservation: "In the world of experimental design, perhaps we should always doubt that a great deal of information can ever be gained in a very short time" (p. 68). However, behavior analysts may be able to capitalize on the efficiency of the multielement design while gaining a great deal of information by making carryover effects visible via within-session analyses. For instance, if the level, trend, and variability of the first few moments or minutes of a session are markedly different than those towards the end of that same session and also are reliably similar to those at the beginning of sessions that have occurred in similar sequence, conclusions about carryover are possible (see Vollmer et al., 1993).

Other assessment designs may also benefit from within-session analyses. For example, concurrent-chains arrangements, in which selections among simultaneously available response options result in access to correlated contexts, have been shown to be a useful measure of individuals preferences for behavioral interventions (Hanley, Piazza, Fisher, Contrucci, & Maglieri, 1997; Hanley, Piazza, Fisher, & Maglieri, 2005), protracted activities (Hanley, Iwata, & Lindberg, 1999), and learning contexts (Tiger, Hanley, & Hernandez, 2006). Between 10 and 15 opportunities to

choose from among the different contexts by selecting context-correlated cues were arranged in each session of these studies. Session-by-session data, showing the percentage of selections towards each of the three options, were used to determine when a preference for a single context emerged. Relying on session-by-session data required between 30 and 180 selection trials to determine individual preferences across these studies. By contrast, when using a similar assessment to determine preschool children's preferences for variations of multiple-schedule arrangements, Tiger, Hanley, and Heal (2006) displayed cumulative trial-by-trial (as opposed to session-by-session) data and were able to determine children's preferences in 6 to 30 selections. Thus, the viability of intimate analyses in enhancing assessment efficiency should be explored across multiple assessment types.

Within-session analyses have been shown to aid in the evaluation of skill endurance (Martens et al., 1997) and the representation of differential patterns of newly acquired behavior (Hanley et al., 2001). Within-session data are also frequently presented when responding occurs within a single, prolonged acquisition session (e.g., Arntzen & Almas, 2002; Rodgers & Iwata, 1991; Yoon & Feliciano, 2007). Yet, the demonstration of within-session analyses seems to be more thorough in the assessment and treatment of problem behavior than in the study of the acquisition of appropriate behavior. It is unclear whether this difference exists because within-session analyses of skill acquisition have not contributed important detail to account for the behavior–environment relations or because the options for such analyses have yet to be fully explored. We believe it is likely to be the latter case, and that more intimate analyses of behavioral acquisition or synthesis are warranted. For instance, multiple types of prompts (e.g., verbal prompt, model prompt) to engage in an appropriate response may be delivered at various times within teaching sessions. Perhaps

an intimate analysis of the temporal distribution of these prompts in relation to responses will prove valuable for understanding the conditions necessary for acquisition of socially important and independent behaviors.

Considering the utility of within-session analyses in the past and the potential areas of within-session application for the future, our general recommendation is a simple one: Do not fear intimacy with your data.

REFERENCES

- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Arntzen, E., & Almas, I. K. (2002). Effects of mand-tact versus tact-only training on the acquisition of tacts. *Journal of Applied Behavior Analysis*, 35, 419–422.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1, 91–97.
- Bailey, J. S., & Burch, M. R. (2002). *Research methods in applied behavior analysis*. Thousand Oaks, CA: Sage.
- Best, L. A., Smith, L. D., & Stubbs, D. A. (2001). Graph use in psychology and other sciences. *Behavioural Processes*, 54, 155–165.
- Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*, 18, 111–126.
- Coleman, C. L., & Holmes, P. A. (1998). The use of noncontingent escape to reduce disruptive behaviors in children with speech delays. *Journal of Applied Behavior Analysis*, 31, 687–690.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis*. Columbus, OH: Merrill.
- Daly, E. J., Martens, B. K., Hamler, K. R., Dool, E. J., & Eckert, T. L. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *Journal of Applied Behavior Analysis*, 32, 83–94.
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.
- Hains, A. H., & Baer, D. M. (1989). Interaction effects in multielement designs: Inevitable, desirable, and ignorable. *Journal of Applied Behavior Analysis*, 22, 57–69.
- Hanley, G. P., Iwata, B. A., & Lindberg, J. S. (1999). Analysis of activity preferences as a function of differential consequences. *Journal of Applied Behavior Analysis*, 32, 419–435.
- 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.

- Hanley, G. P., Iwata, B. A., & Thompson, R. H. (2001). Reinforcement schedule thinning following treatment with functional communication training. *Journal of Applied Behavior Analysis*, 34, 17–38.
- Hanley, G. P., Piazza, C. C., Fisher, W. W., Contrucci, S. A., & Maglieri, K. A. (1997). Evaluation of client preference for function-based treatment packages. *Journal of Applied Behavior Analysis*, 30, 459–473.
- Hanley, G. P., Piazza, C. C., Fisher, W. W., & Maglieri, K. A. (2005). On the effectiveness of and preference for punishment and extinction components of function-based interventions. *Journal of Applied Behavior Analysis*, 38, 51–65.
- Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, 27, 197–209. (Reprinted from *Analysis and Intervention in Developmental Disabilities*, 2, 3–20, 1982)
- Iwata, B. A., Pace, G. M., Cowdery, G. E., & Miltenberger, R. G. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis*, 23, 11–27.
- Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and tactics of behavioral research*. Hillsdale, NJ: Erlbaum.
- Kahng, S., & Iwata, B. A. (1999). Correspondence between outcomes of brief and extended functional analyses. *Journal of Applied Behavior Analysis*, 32, 149–159.
- Kahng, S., Iwata, B. A., Thompson, R. H., & Hanley, G. P. (2000). A method for identifying satiation versus extinction effects under noncontingent reinforcement schedules. *Journal of Applied Behavior Analysis*, 33, 419–432.
- Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.
- Lalli, J. S., Mace, F. C., Livezey, K., & Kates, K. (1998). Assessment of stimulus generalization gradients in the treatment of self-injurious behavior. *Journal of Applied Behavior Analysis*, 31, 479–483.
- Laraway, S., Snyderski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis*, 36, 407–414.
- Lindberg, J. S., Iwata, B. A., Roscoe, E. M., Worsdell, A. S., & Hanley, G. P. (2003). Treatment efficacy of noncontingent reinforcement during brief and extended application. *Journal of Applied Behavior Analysis*, 36, 1–19.
- Martens, B. K., Bradley, T. A., & Eckert, T. L. (1997). Effects of reinforcement history and instructions on the persistence of student engagement. *Journal of Applied Behavior Analysis*, 30, 569–572.
- McKerchar, P. M., & Thompson, R. H. (2004). A descriptive analysis of potential reinforcement contingencies in the preschool classroom. *Journal of Applied Behavior Analysis*, 37, 431–444.
- Meyers, L. S., & Grossen, N. E. (1974). *Behavioral research, theory, procedure and design*. San Francisco: Freeman.
- Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior*, 37, 149–155.
- Mudford, O. C., Beale, I. L., & Singh, N. N. (1990). The representativeness of observational samples of different durations. *Journal of Applied Behavior Analysis*, 23, 323–331.
- Mueller, M. M., Piazza, C. C., Patel, M. R., Kelley, M. E., & Pruett, A. (2004). Increasing variety of foods consumed by blending nonpreferred foods into preferred foods. *Journal of Applied Behavior Analysis*, 37, 159–170.
- Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., et al. (1991). A brief functional analysis of aggressive and alternative behavior in an outclinic setting. *Journal of Applied Behavior Analysis*, 24, 509–522.
- Poling, A. (1979). The ubiquity of the cumulative record: A quote from Skinner and a frequency count. *Journal of the Experimental Analysis of Behavior*, 31, 126.
- Roane, H. S., Lerman, D. C., Kelley, M. E., & Van Camp, C. M. (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.
- Roane, H. S., Vollmer, T. R., Ringdahl, J. E., & Marcus, B. A. (1998). Evaluation of a brief stimulus preference assessment. *Journal of Applied Behavior Analysis*, 31, 605–620.
- Rodgers, T. A., & Iwata, B. A. (1991). An analysis of error-correction procedures during discrimination training. *Journal of Applied Behavior Analysis*, 24, 775–781.
- Skinner, B. F. (1976). Farewell, my lovely! *Journal of the Experimental Analysis of Behavior*, 25, 218.
- Slooman, K. N., Vollmer, T. R., Cotnoir, N. M., Borrero, C. S. W., Borrero, J. C., Samaha, A. L., et al. (2005). Descriptive analyses of caregiver reprimands. *Journal of Applied Behavior Analysis*, 38, 373–383.
- Smith, R. G., Iwata, B. A., Goh, H. L., & Shore, B. A. (1995). Analysis of establishing operations for self-injury maintained by escape. *Journal of Applied Behavior Analysis*, 28, 433–445.
- Tiger, J. H., Hanley, G. P., & Heal, N. A. (2006). The effectiveness of and preschoolers' preferences for variations of multiple-schedule arrangements. *Journal of Applied Behavior Analysis*, 39, 475–488.
- Tiger, J. H., Hanley, G. P., & Hernandez, E. (2006). An evaluation of the value of choice with preschool children. *Journal of Applied Behavior Analysis*, 39, 1–16.
- Ulman, J. D., & Sulzer-Azaroff, B. (1975). Multielement baseline design in educational research. In E. Ramp & G. Semb (Eds.), *Behavior analysis: Areas of research and application* (pp. 377–391). Englewood Cliffs, NJ: Prentice Hall.
- Van Houten, R. (1993). The use of wrist weights to reduce self-injury maintained by sensory reinforcement. *Journal of Applied Behavior Analysis*, 26, 197–203.

- Vollmer, T. R., Borrero, J. C., Wright, C. S., Van Camp, C., & Lalli, J. S. (2001). Identifying possible contingencies during descriptive analyses of severe behavior disorders. *Journal of Applied Behavior Analysis, 34*, 269–287.
- Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). Within-session patterns of self-injury as indicators of behavioral function. *Research in Developmental Disabilities, 14*, 479–492.
- Vollmer, T. R., Marcus, B. A., Ringdahl, J. E., & Roane, H. S. (1995). Progressing from brief assessments to extended experimental analyses in the evaluation of aberrant behavior. *Journal of Applied Behavior Analysis, 28*, 561–576.
- Vollmer, T. R., Ringdahl, J. E., Roane, H. S., & Marcus, B. A. (1997). Negative side effects of noncontingent reinforcement. *Journal of Applied Behavior Analysis, 30*, 161–164.
- Yoon, S., & Feliciano, G. M. (2007). Stimulus-stimulus pairing and subsequent mand acquisition of children with various levels of verbal repertoires. *The Analysis of Verbal Behavior, 23*, 3–16.

Received May 21, 2007

Final acceptance July 7, 2007

Action Editor, James Carr