Spring 5-5-2022

Trial-based Measurements as an Index of Response Strength

Vasily Belichenko
bacileus86@gmail.com

Trial-based Measurements as an Index of Response Strength

A Thesis
by
Vasily V Belichenko

Submitted to the Faculty of the Department of Health Professions
at Rollins College in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF ARTS IN APPLIED BEHAVIOR ANALYSIS AND CLINICAL SCIENCE

April 2022
Winter Park, FL

Table of Contents

Abstract

Identifying and evaluating dependent variables within functional analyses are essential steps that affect the assessment's overall acceptability. Although most of the studies on functional analyses suggest that measuring repeated occurrences of behavior with rate measurements generates the most accurate data, several researchers suggest using alternative measurement methods to avoid the heavy burden of the procedure. In the present study, I partially replicated a study by Thomason-Sassi and colleagues (2011) that evaluated the correspondence between latency measures of responding and rate. Additionally, trial-based measurements, such as the occurrence of the behavior, represented in the percentage of trials, were compared with rate measurements. The research outcomes have shown that trial-based measurements could be useful measures of responding when latency and rate are impractical to arrange or undesirable due to the heavier observer burden these measurements represent. The results indicate a correspondence between the three measurement methods. Implications for trial-based measurements and future directions for research are discussed.

Keywords: functional analysis, trial-based measurements of responding

Trial-based Measurements as an Index of Response Strength

Improving and assessing the quality of behavioral measurements is essential for researchers and clinicians. No responsible applied behavior specialist would consider using invalid or inaccurate data in their practices. Johnston and Pennypacker (2008) described indicators of trustworthy measurement and threats to measurement validity, accuracy, and reliability. The authors mentioned that, without these characteristics, scientific measurements would be useless. One of the vulnerable areas for biased measurements (i.e., measurements which are likely to overestimate or underestimate the actual values) is behavioral assessment, especially experimental functional analysis (FA). Compromised outcomes of behavioral assessments directly threaten the efficacy of subsequent behavioral interventions. On the other hand, valid and accurate data lead to effective treatments for clients. Moreover, if valid and reliable data could be collected with less effort, it could enhance the assessment procedures.

The gold standard of functional behavior assessments is the FA. These assessments involve manipulating environmental events, including consequences that potentially maintain maladaptive behavior (Iwata et al., 1994). Despite the popularity of FAs in the research field, the assessment's applicability and social acceptance can be improved. Moreover, low social validity of FAs is reported by practitioners, caregivers, and teachers (Iwata & Dozier, 2008). Obstacles in the implementation of FA include mainly the severity of the behavior, time constraints, and the difficulty of conducting the assessment (Hanley, 2012).

The standard functional analysis includes three test conditions: Attention, Escape, No Interaction, and a control condition to which behavior under each of the three test conditions is compared. During a standard FA, conditions are systematically presented to the participant, typically in successive 10-min sessions via a multielement design until the behavior's function is identified. For example, during the test condition of escape from demands, a therapist presents the client tasks. As soon as target behavior occurs, a break is provided. The obtained results are then graphed and visually analyzed. The functional reinforcer is typically characterized by differentiation from the control condition. This approach has been extended and replicated in many studies focused on assessing and treating a spectrum of behaviors (Cooper et al., 2007; Iwata & Dozier, 2008).

In addition to the standard FA, several other methods for conducting FA have been identified, including *latency-based functional analysis* (LBFA; e.g., Thomason-Sassi et al., 2011) and *trial-based functional analyses* (TBFA; e.g., Sigafoos & Saggers, 1995; Bloom et al., 2011). In LBFA, the sessions are terminated after the first response, and response latency is used as the primary dependent variable (rather than rate, as in a standard FA). Latency is the interval between the beginning of the session to the first occurrence of the target behavior. Several researchers have proposed that latency can be a valid index of responding that may have utility in cases dealing with behaviors that are hard to "reset" or need immediate intervention (Piazza & Fisher, 1991; Zarcone et al., 1993). For example, Thomason-Sassi and colleagues (2011) compared latency to rate measures of responding during the FA of problem behavior. During LBFA, sessions were terminated contingent on the target behavior's first occurrence or after 5 min elapsed, whichever came first. For example, during the escape condition, the therapist provided tasks to the participants. Escape from aversive stimuli was provided contingent on the first occurrence of the target behavior (i.e., elopement). The first occurrence of problem behavior resulted in a 30-s break and, following the break session, session termination. Because the results of the LBFA showed high correspondence (in terms of identifying the function of behavior) with the standard FA, the authors concluded that latency to emit the target behavior could be an effective, alternative measure when dealing with high-risk behaviors or when there is limited time to conduct the analysis. Identical FA conditions are typically included in LBFA as in a standard FA to demonstrate experimental control in both assessments (via multielement designs).

Though LBFA is considered a safe and feasible assessment procedure, there are potential limitations. One limitation is a single exposure to the contingencies during the procedure. Thus, the LBFA relies heavily on antecedents instead of consequences to determine the function of behavior. Additionally, researchers have suggested that the LBFA is potentially more prone to false-positive outcomes than the standard FA (Thomasson-Sassi et al., 2011). For example, the LBFA may identify a function of behavior that is not identified as a maintaining variable under extended conditions of the standard FA.

Another variation of FA is the trial-based functional analysis (TBFA). TBFA uses a trial-based rather than a time-based format, so each condition is examined in 10-20 trials. Each trial typically consists of two segments, one of which is the control condition. The reinforcer is continuously available during the 1-2 min control segment. In the following 1-2 min test segment, the contingency for problem behavior is present (i.e., reinforcement is provided contingent upon problem behavior). Each segment is terminated contingent on problem behavior occurrence. For example, during the control segment of the escape condition, the experimenter is seated next to the participant. No demands are placed, and no task materials are provided to the participant. The test segment of the escape condition begins contingent on the occurrence of the target behavior or after 2 min elapses. The test segment starts with delivering instruction for tasks. Compliant responses result in verbal praise.

During the test segment, the occurrence of problem behavior results in a 30-s break and, subsequently, the termination of the session. Consequently, the results of each condition are compared via data analysis. TBFA data are analyzed by comparing the percentage of control and test segments with problem behavior within and between conditions. Unlike LBFA and the standard FA, TBFA data are typically presented and visually analyzed via a bar graph. Sigafoos and Saggers (1995) were the first experimenters who implemented TBFA in a classroom setting. During TBFA, experimenters presented tasks in the test segment and then terminated them if problem behavior occurred. In the second segment (control), the tasks were absent. The experimenters identified reinforcement contingencies that maintained the problem behavior for both of their subjects.

Several advantages of trial-based measurements have been identified to address the limitations of continuous measurements for FAs, including reduced observer burden in a busy environment (TBFA; e.g., Bloom et al., 2011; Sigafoos & Saggers, 1995). For example, during trial-based data collection, the occurrence of the behavior would be recorded in each trial, and the percentage of correct responses would then be calculated. This procedure provides enough information for many situations, including examining data collected in the classrooms, thus increasing the ease of implementation (Sigafoos & Saggers, 1995). Moreover, trial-based measurements can track high-frequency behavior such as pencil-tapping and behaviors occurring for extended periods, such as tantrums lasting more than an hour (Fiske & Delmolino, 2012). Finally, another significant advantage of trial-based measurements is that data can be collected within the ongoing school or home activities with minimal disruption, resulting in higher ecological validity of assessment (Hanley et al., 2003).

The current study examined three data collection methods: rate, response latency, and percentage of trials in which the behavior occurred. Rate, often reported as responses per minute, is considered the best indicator of response strength (Killeen & Hall, 2001; Kubina, 2005; Lovitt, 1968; Skinner, 1966). Rate measurements provide more information than frequency data alone and typically represent the target behavior more accurately by accounting for session duration. In addition, rate measurements are flexible and could be collected during varying session durations, which makes it the preferred measure for free operant behavior (LeBlanc et al., 2015). Finally, rate measurements are direct measures of behavior as it occurs and could be used for measuring repeated events. Thus, rate is preferred over other measurement methods. Despite these considerable strengths, rate measurements do suffer from some limitations. For example, LeBlanc et al. (2015) suggested avoiding rate data collection during high-frequency or non-discrete behavior. Also, rate measurements are not recommended during discrete trials because behaviors have limited opportunities to occur. Additionally, the understanding of some topographies of responses (e.g., vomiting) could be constrained when summarized by rate (Thomasson-Sassi et al., 2011).

Response latency (i.e., the time between an event and a response) is an alternative to the traditional measures based on response repetition, such as rate (Lerman, 2011; Thomasson-Sassi et al., 2011). This involves evaluating the speed of responding to a particular stimulus. For example, decreasing latency data shows the rate of responding during discrete trial training (DTT) is increasing, and increasing latency indicates the rate of responding is decreasing. Response latency measurements are recommended when responses occur too quickly or too slowly after stimulus presentation. In addition, latency measurements can reduce session duration because these require only the first occurrence of a behavior (Thomasson-Sassi et al., 2011). Also, latency measures are helpful when the length of time between a specific cue, event, or verbal prompt and the occurrence of behavior is essential. Lastly, response latency data collection can help identify the hierarchy of response strength in cases with multiple functions. For example, one of the participant's raw data in the Thomasson-Sassi et al. (2011) study indicates self-injurious behavior (SIB) occurred in the demand conditions but was maintained by automatic reinforcement instead of negative reinforcement. However, given the potential advantages of latency as a primary dependent measure, several limitations should be considered. One limitation is a lengthy latency to the first response, which may create patterns of responding and eventually not reflect overall response strength (Thomasson-Sassi et al., 2011). Another significant limitation of latency measure is that data collection could be challenging to collect while teaching or in-home settings. Thus, comparison with trial-based measurements (occurrence/nonoccurrence) might be beneficial in further evaluating these methods.

Trial-based measurements involve reporting the presence or absence of target behavior within each trial and are reflected in a percentage of trials during which behavior occurred (Lerman, 2011; Bloom et al., 2011). The percentage of trials with behavior that occurred could be quickly calculated based on other implemented measures. Properly collected trial-based data directly and accurately reflects the number of times a behavior occurs during discrete trials. The most significant strength of trial-based measurements is the ease of use, which is advantageous when relying on data collection by parents or teachers. Because trial-based measurements involve a lower observer burden than rate, caregivers and practitioners may benefit by using trial-based measurements as the primary dependent variable for collecting the data.

Though these collection methods have been used in separate assessments, all three have not been compared in the same FA within-subjects. Thus, graphs representing the rate, latency, and percentage of trials might not always accurately represent the function(s) of behavior. Several researchers have evaluated the correspondence between outcomes of LBFA and the standard FA or TBFA and standard FA (Bloom et al., 2011; Dayton, 2011; Curtis, 2017; Thomasson-Sassi et al., 2011). However, to my knowledge, no study has yet directly compared trial-based measurements, such as latency and time sampling, with traditional rate measurements using a single assessment. Therefore, the present study has shown that trial-based measurements are comparable in identifying the FA function. In the present study, the data were produced by a single FA and graphed according to rate and latency as in

Thomasson-Sassi et al., 2011 to reaffirm the correspondence between rate and latency. Additionally, rate and trial-based measurements are evaluated.

Method

Participants, Materials, and Setting

Two males, ages 9 and 13 years and diagnosed with developmental disabilities were recruited for this study. The participants received behavioral services for maladaptive behaviors (e.g., elopement, aggression, property destruction, etc.) at a local ABA clinic. Preference assessments and FAs were conducted in the homes where the participants presently receive behavior-analytic treatment. The session rooms contained a table, two chairs, and the materials needed to conduct the sessions during each FA condition. Data were collected via a smartphone with the downloaded application "Countee." At least 80% of sessions were video recorded.

Operational definitions varied across subjects (Jamal and William). Jamal's property destruction behavior was operationally defined as any instance or attempted instance to rip up clothes or other items. Any instance of Jamal's touching and holding the clothes or zip locks for more than 3 s would count as the target behavior. William's elopement was operationally defined as any actual or attempted instance to leave an area without permission. During Escape, Attention and No interaction condition, William's attempts to leave the table would count as target behavior. During Play condition, any instance or attempt to leave the room would count as target behavior.

Response Measurement, Interobserver Agreement and Treatment Integrity

During each session, the leading observer recorded each instance of the target behavior via Countee (Design, 2021). Countee was set up to score the frequency of the target behavior. One of the Countee features is leaving time stamps when the behavior occurs during event recording (i.e., the temporal locus for each instance of the target behavior was recorded). This procedure allowed the calculation of three dependent variable measures in the present study: the rate and latency of the target behavior and the percentage of sessions with target behavior (i.e., trial-based measurement). The outcomes of the three different measures were graphed separately using line graphs for rate and latency and a bar graph for trial-based measurement. Trial-based measurements and response latency were each compared with response rate (which is the best measure of strength).

Rate (i.e., responses per min) was calculated by dividing the total number of responses by the session duration in minutes. For example, if problem behavior occurred five times during a 5-min session, one response per minute was recorded as the outcome for rate measurement. Interobserver agreement (IOA) for selected sessions was assessed using proportional agreement with 15-s intervals. During each 15-s interval, the smaller scored number of instances of the target behavior was divided by the larger number of instances recorded and multiplied by 100. Latency was measured by retrieving data from the *Countee* application. No response was recorded if the target behavior did not occur during the session. For trial-based measurements, latencies were recorded using a simple "YES" or "NO" measure in which latencies shorter than 120 s were scored as "YES" and latencies longer than 120 s were scored as "NO." The percentage of sessions with target responses that occurred within the first 2 min was calculated across all sessions within each condition.

To assess the reliability and treatment integrity of the independent variables, a second observer simultaneously but independently recorded Jamal’s data and assessed via video recordings the researcher’s implementation of the data collection methods for William. IOA and treatment integrity were calculated in 32% and 30% of randomly selected sessions for Jamal and William, respectively. IOA was 72% and 87.5% for Jamal and William, respectively. Treatment Integrity data indicated mean treatment integrity at 88% for Jamal and 100% for William across sessions where treatment integrity checks were conducted.

Procedure

Preference Assessment

A multiple-stimulus-without-replacement preference assessment (MSWO; DeLeon & Iwata, 1996) was conducted. The preference assessment procedures targeted identifying preferred items to include during the FA. Prior to each assessment, the researcher provided the subject with 30 s of access to each stimulus. During the MSWO preference assessments, items were arranged in an array in front of the participant. The participants were instructed to select one item. After the item was chosen, the participant was allowed to play with the item for 15-20 s and the remaining items were removed. The chosen item was then removed, and the remaining items were rearranged in a different order and represented. The procedure was repeated until all items were selected or the participant did not make a selection. The preference hierarchy was calculated by dividing the number of choices for each item by the number of opportunities (i.e., presentations of each stimulus).

Functional Analysis

A minimum of 3 series of conditions were conducted during the FA described by Iwata et al. (1982) until the target behavior’s function is identified via the multielement design. The decision to complete the FA was based on rate graph outcomes or after 10 trials of each condition. The sessions lasted 5 min each. The conditions included in the FA (e.g., Attention, Escape, No Interaction, or Tangible) were selected based on the hypothesized functions of behavior identified via caregiver reports. The play condition served as the control condition.

The participant’s attempts to access items unrelated to the specific condition being conducted were blocked. The same experimenter conducted all FA sessions in the same room for William. The same experimenter conducted the specific condition sessions in the different rooms for Jamal. For example, play and demand conditions were conducted in the garage room, where the client spent most of his ABA sessions. No interaction and attention conditions were conducted in the client's bedroom. At least 5 min breaks occurred between sessions. The correspondence between trial-based measurements and rates was evaluated by comparing the function of behavior concluded by each data collection method.

Attention. A researcher was present in the room and diverted his attention from the participant. Free access to moderately preferred items (identified during preference assessment) was available to the participant. The experimenter said, "I have to do some work," and faced away from the subject. Contingent on the target behavior, the therapist provided reprimand (e.g., "Stop, don't do that") and provided physical contact in the form of blocking if needed for safety concerns. All other responses were ignored.

Escape. A therapist was present in the room. The participant was presented with tasks identified as challenging through indirect assessments (e.g., caregiver interview). No access to leisure materials was provided. During this condition, instructions were provided based on three-step prompting sequences (verbal, model, and physical) for several tasks. For example, if the subject did not respond within 5 s, the experimenter repeated the instruction and demonstrated the correct response. Immediately following displays of the target behavior, instructions were terminated for 30 s. Appropriate responses were confirmed that the task is correct such as "This is touching your nose."

No interaction. The participant was instructed to "wait" in the room without interactions. No attention was provided, and no demands were presented during this condition. A therapist was in the room at least 3 m away from the subject and not engaged with the client or hiding behind the door. Target behavior did not produce any consequences.

Play. The participant has been given continuous access to preferred stimuli with no instructional demands. The researcher noncontingently provided attention (e.g., delivering toys, vocal praise). The researcher engaged in any social iteration initiated by the participant. All target behavior was ignored.

Results

Preference assessment results for Jamal and William are displayed in Figure 1. Jamal's highest preferred items were bubbles and playdoh. His moderately preferred toy was a squishy ball. Jamal's least preferred item were puzzles. William's highest preferred item was iPad. His moderately and least preferred items were chips and toys with sounds and lights, respectively.

Figure 2 shows data for three measurement procedures across conditions for Jamal. The top panel of the figure displays the target responses per minute (rate). A visual analysis of the rate graph indicates that the behavior is potentially automatically maintained. Accordingly, the middle panel of the figure shows that the latency to problem behavior was shorter in the no interaction condition compared to other conditions. The bottom panel of the figure shows that the participant engaged in more problem behavior during the no interaction condition sessions (57% of trials) than other conditions (attention: 29%, escape: 0%, and play: 14% of trials) during the FA.

Figure 3 shows data for three measurement procedures across conditions for William. The top panel of the figure displays the target responses per minute (rate). A visual analysis of the rate graph indicates that the behavior is potentially attention maintained. However, the middle panel of the figure does not show clear differentiation between condition's outcomes, the latency to problem behavior which occurred during all trials was shorter in the attention condition compared to other conditions. The bottom panel of the figure shows that the participant engaged in more problem behavior during the attention condition sessions (100% of trials) than other conditions (no interaction: 67%, escape: 67%, and play: 67% of trials) during the FA.

Discussion

Although the rate and latency graphs provide more information, the trial-based graph clearly shows the behavior has the same function as indicated by the other two graphs. For example, the rate and latency graphs have information such as level, trend, and variability. However, the bar graph of trial-based measurements is more user-friendly in terms of reading it by caregivers and teachers.

The key difference between previous studies (Bloom et al., 2011; Sigafoos & Saggers, 1995) is that the authors conducted individual FAs and compared the corresponding results between TBFA and standard FA. However, the current study compared rate and trial-based measurements retrospectively when both rate and latency data were already available based on conducting a single FA. This methodology allowed a direct comparison of measurement procedures within and across participants. The conclusion about function based on the trial-based graph is similar to the rate graph for each subject. Thus, the trial-based data can arguably suffice as the only determination of function. Also, we found that is the case for both subjects in the study.

During FA with Jamal, different rooms were used for FA conditions, such as garage for Play and Escape condition. Client's bedroom was used once during No Interaction condition in order to prevent accidental attention provided by parents or siblings during the session. Also, Escape and few Play conditions were run by subject's current registered behavior technician (RBT) in order to increase naturalistic environment.

It is essential to note some potential limitations of current research, however. For example, the rate and latency graphs provide more information than the trial-based graphs, such as the hierarchy of the potential functions of behavior (which might be necessary if there are multiple functions of behavior). There is also the potential for an over/underestimation of behavior because the trial-based measures are less sensitive than rate and latency measures. For example, William's data show that behavior occurs at the rate of approximately 4 times per minute based on the rate graph. However, there is no way to determine how many times behavior occurred in the trial-based graph. Therefore, you could have 100% of trials with problem behavior, but it might be occurring at high or low frequency. Another potential limitation of the present study is that the study was conducted in the participant's homes. In-home settings are less controlled, so there is a greater tendency for behavior to be affected by extraneous variables. Therefore, treatment integrity can potentially be compromised.

Although the trial-based measures provide less information about behavior, collecting trial-based data might be easier to train teachers or caregivers to implement. Therefore, because the results of the trial-based measure align with the rate measure, trial-based measures could be preferable to use in home settings. Also, trial-based measures involve less observer burden than rate or latency measurements, meaning they are more practical for non-BCBAs to conduct and analyze, making it possible to move to intervention sooner. Also, although home settings can be less well-controlled, they are more externally valid. Therefore, although the rate measure is the gold standard for FAs, the trial-based measure could be a reasonable substitution for the home setting, meaning we are less reliant on descriptive measures typically used in the home or school settings to hypothesize the function of behavior.

One direction of future studies could include comparing the social validity of measurement methods by different constituents (e.g., parents, teachers) across different settings (e.g., home, school). This is important because more socially acceptable methods are more likely to be used. If trial-based measures are an appropriate substitute for rate measures, more valid interventions are likely to follow (as opposed to interventions resulting from more commonly used descriptive measures). This will hopefully lead to better treatment outcomes for our clients.

References

Bloom, S. E., Iwata, B. A., Fritz, J. N., Roscoe, E. M., & Carreau, A. B. (2011). Classroom application of a trial-based functional analysis. *Journal of Applied Behavior Analysis, 44*(1), 19-31. https://doi.org/10.1901/jaba.2011.44-19

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Pearson Education.

Curtis, Kaitlin S., "A comparison of traditional, trial-based, and synthesized contingency trial-based functional analyses" (2017). *MSU Graduate Theses*. 3214. https://bearworks.missouristate.edu/theses/3214

Dayton, Elizabeth, "Latency as a dependent variable in trial-based functional analysis." (2011). *All Graduate Theses and Dissertations*. 1079. https://digitalcommons.usu.edu/etd/1079.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*(4), 519-533. https://doi.org/10.1901/jaba.1996.29-519

Design, K., 2021. Countee - Data collection system. [online] Counteeapp.com. Available at: <https://www.counteeapp.com/> [Accessed 18 October 2021].

Fiske, K., & Delmolino, L. (2012). Use of discontinuous methods of data collection in behavioral intervention: guidelines for practitioners. *Behavior Analysis in Practice*, *5*(2), 77–81. https://doi.org/10.1007/BF03391826

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

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*(2), 147-185. https://doi.org/10.1901/jaba.2003.36-147

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*, 1982, 2, 3-20. https://doi.org/10.1901/jaba.1994.27-197

Iwata, B. A., & Dozier, C. L. (2008). Clinical application of functional analysis methodology. *Behavior Analysis in Practice*, *1*, 3–9. https://doi.org/10.1007/BF03391714

Johnston, J.M., Pennypacker, H.S., & Green, G. (2008). *Strategies and Tactics of Behavioral Research* (3rd ed.). Routledge. https://doi.org/10.4324/9780203837900

Killeen, Peter & Hall, Scott. (2001). The principal components of response strength. *Journal of the Experimental Analysis of Behavior*. *75*. https://doi.org/10.1901/jeab.2001.75-111

Kubina R. M., Jr (2005). The relations among fluency, rate building, and practice: A response to Doughty, Chase, and O'Shields (2004). *The Behavior Analyst*, *28*(1), 73–76. https://doi.org/10.1007/BF03392105

LeBlanc, L. A., Raetz, P. B., Sellers, T. P., & Carr, J. E. (2015). A proposed model for selecting measurement procedures for the assessment and treatment of problem behavior. *Behavior Analysis in Practice*, *9*(1), 77–83. https://doi.org/10.1007/s40617-015-0063-2

Lerman, D. C., Dittlinger, L. H., Fentress, G., & Lanagan, T. (2011). A comparison of methods for collecting data on performance during discrete trial teaching. *Behavior Analysis in Practice*, *4*(1), 53–62. https://doi.org/10.1007/BF03391775

Lovitt, T. C. (1968). Operant conditioning techniques for children with learning disabilities. *The Journal of Special Education*, *2*(3), 283-289. https://doi.org/10.1177/002246696800200306

Piazza, C. C., & Fisher, W. (1991). A faded bedtime with response cost protocol for treatment of multiple sleep problems in children. *Journal of Applied Behavior Analysis*, *24*, 129–140. https://doi.org/10.1901/jaba.1991.24-129

Sigafoos, J., & Saggers, E. (1995) A discrete-trial approach to the functional analysis of aggressive behaviour in two boys with autism, *Australia and New Zealand Journal of Developmental Disabilities, 20*(4), 287-297. https://doi.org/10.1080/07263869500035621

Skinner, B.F. (1966), What is the experimental analysis of behavior? *Journal of the Experimental Analysis of Behavior*, *9*: 213-218. https://doi.org/10.1901/jeab.1966.9-213

Thomason-Sassi, J. L., Iwata, B. A., Neidert, P. L., & Roscoe, E. M. (2011). Response latency as an index of response strength during functional analyses of problem behaviors. *Journal of Applied Behavior Analysis*, *44*(1), 51–67. https://doi.org/10.1901/jaba.2011.44-51

Zarcone, J., Iwata, B., Hughes, C., & Vollmer, T. (1993). Momentum versus extinction effects in the treatment of self-injurious escape behavior. *Journal of Applied Behavior Analysis*, *26*(1), 135-136. https://doi.org/10.1901/jaba.1993.26-135

Figure 1

Preference assessment results for Jamal (top) and William (bottom). Stimuli are listed in order of highest preferred to lowest preferred.

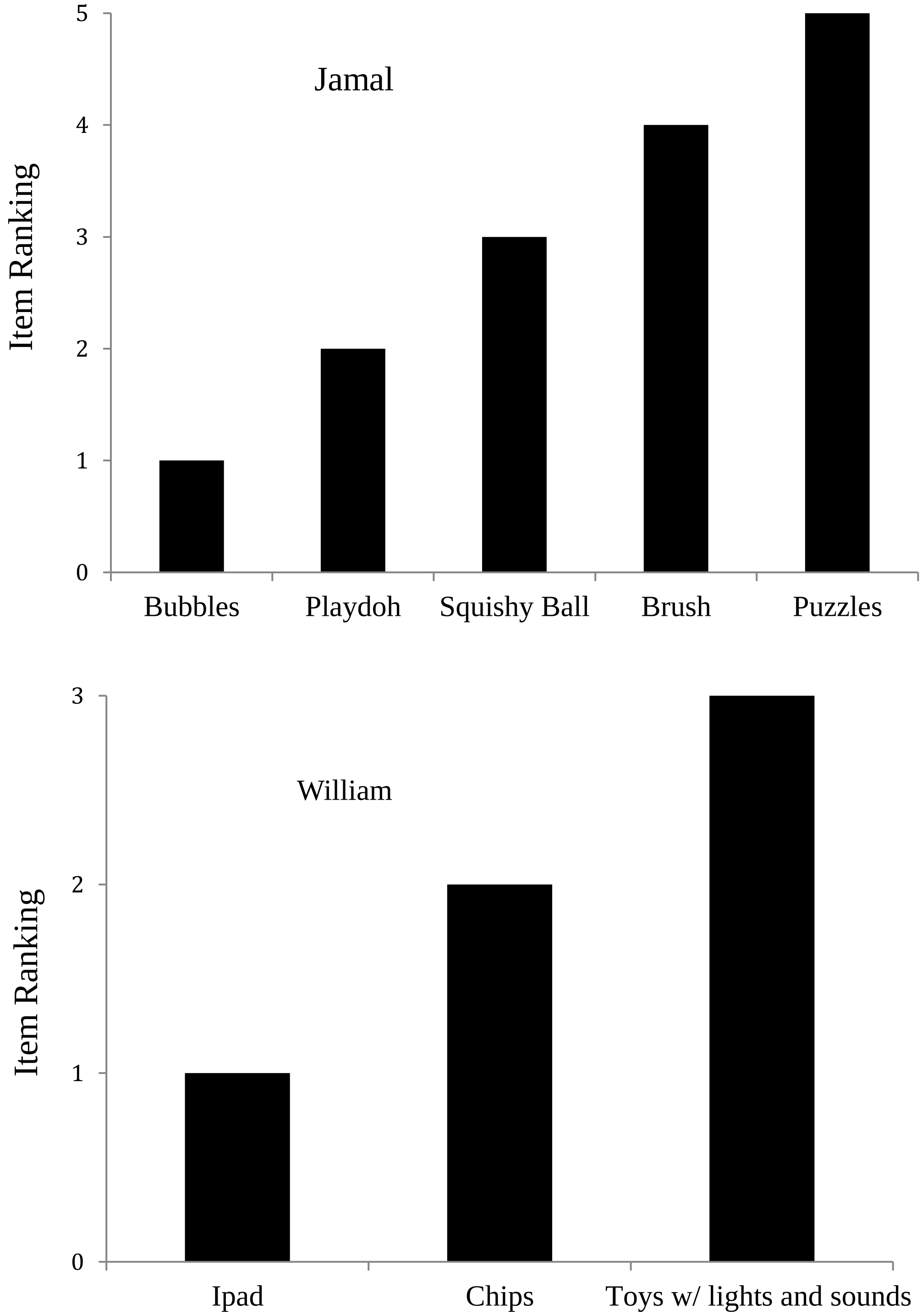

Figure 2

Results of Measurements Across Evaluation Conditions for Jamal

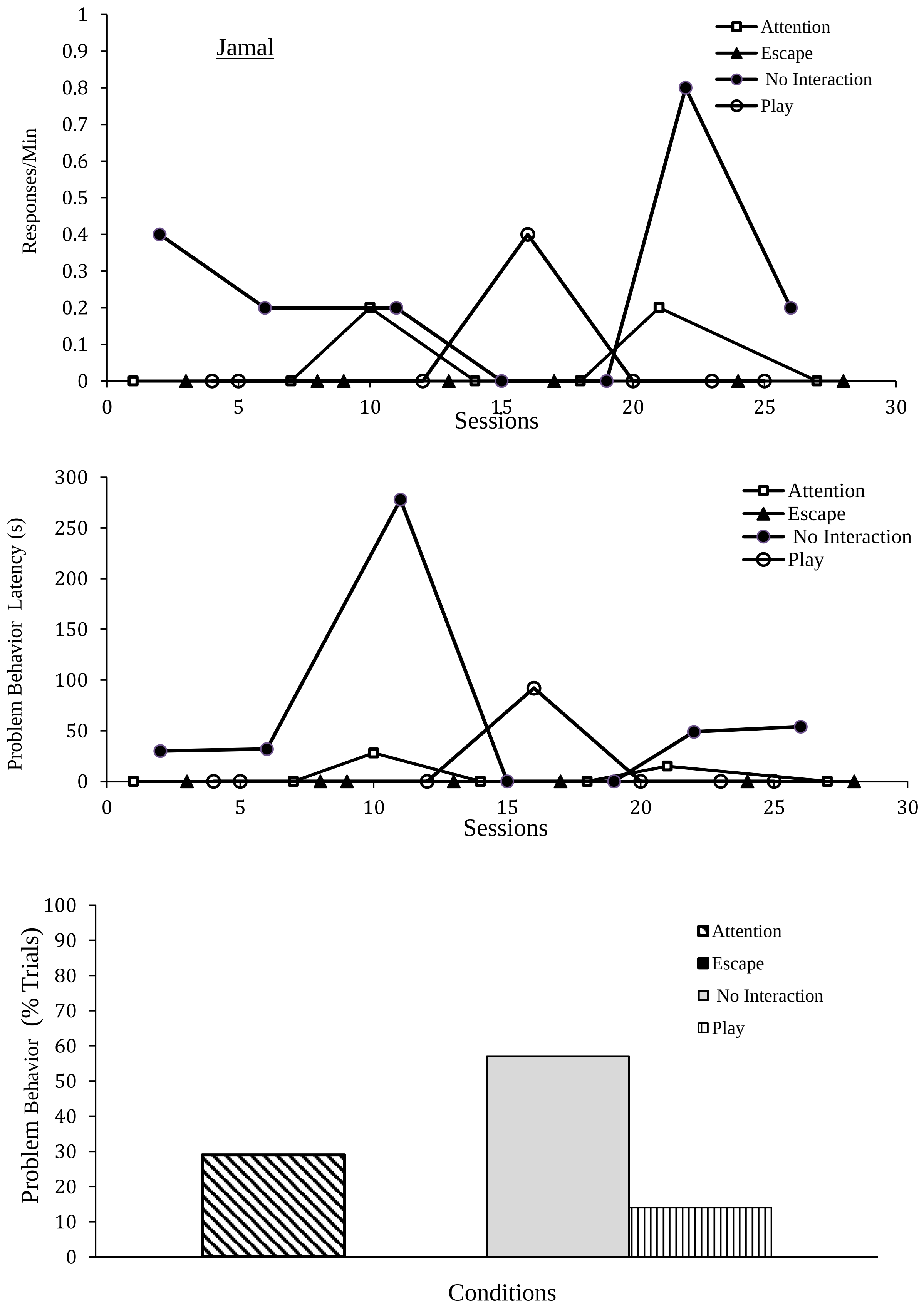

Figure 3

Results of Measurements Across Evaluation Conditions for William