Frequency of Learning Trials Presented During Trial-by-Trial Versus First-Trial Data Recording

Amin D. Lotfizadeh[1], Meghan A. Herron[1], and Alan Poling[2]

[1]Easterseals Southern California

and

[2]Western Michigan University

Abstract

Applied behavior analysts use a variety of strategies to gather data. Continuous recording strategies, such as trial-by-trial recording, are more effortful than discontinuous strategies, such as first-trial recording, and the results of a recent study suggest they may thereby reduce the number of learning opportunities arranged for clients. We compared the number of learning opportunities (trials) presented by behavior therapists when they gathered data using first-trial recording and trial-by-trial recording. An alternating-treatments design was used with six participants as they provided in-home behavioral interventions for individuals diagnosed with autism. The mode of data collection did not consistently affect the number of learning opportunities arranged for clients when data for individual participants were examined. However, trial-by-trial recording resulted in more learning opportunities for most participants, and for all participants combined when mean values were compared. Specific aspects of data recording systems are likely to determine whether they make undue demands on therapists' time, and hence limit learning opportunities, and the accuracy of the obtained data. Further comparison of continuous and discontinuous data collection strategies are warranted.

Keywords: trial-by-trial, first-trial, data collection, learning trials, treatment intensity

Frequency of Learning Trials Presented

During Trial-by-Trial versus First-Trial Data Recording

Using accurate and meaningful data to select target behaviors for clients and to evaluate interventions intended to engender those behaviors are foundational aspects of applied behavior analysis (ABA; Baer, Wolf, & Risley, 1987). A variety of data collection methods are available. For example, both trial-by-trial data and single-trial data collection can be used to track skill acquisition by people with autism (e.g., Leaf, Cihon, Leaf, McEachin, & Taubman, 2017; Schreibman et al., 2015).

Trial-by-trial data collection typically is more accurate than single-trial recording (e.g., Taubman, Leaf, McEachin, Papovich, & Leaf, 2013). It is more sensitive at detecting when the first independent response has occurred (Carey & Bourret, 2014), and more conservative at indicating when a skill has been mastered (Cummings & Carr, 2009; Lerman, Dittlinger, Fentress, & Lanagan, 2011). The differences are often small, however, and trial-by-trial recording and other continuous recording methods may require more time and effort, and thereby reduce the number of learning opportunities arranged in each training session (e.g., Giunta-Fede, Reeve, DeBar, Vladescu, & Reeve, 2016; Najdowski et al., 2013). This is an undesirable outcome, because the rate of skill acquisition usually is directly related to the number of learning opportunities that are arranged (e.g., Bahadourian, Tam, Greer, & Rousseau, 2006; Greer, McCorkle, & Williams, 1989).

Little is known about the relative time required for continuous and discontinuous data recording. Carey and Bourret (2014) compared the time that therapists conducting discrete-trial training for children with autism spent in gathering and recording data when data were collected

for the first trial, first three trials, first five trials, and all trials in a session. Participants gathered

data throughout the entire session. The researchers concluded that a particular data collection

method did not save a meaningful amount of time when compared to the others. In fact, they

suggested that because discontinuous data collection methods led to premature indications of

skill mastery, additional training time may have been dedicated to retraining skills that were

falsely identified as mastered as a result of first-trial recording. Unfortunately, these researchers

did not assess the number of learning trials that were presented as a function of the recording

method.

To our knowledge, only one study has directly compared the number of learning trials

presented as a function of the kind of data collection procedure in effect. Taubman et al. (2013)

compared the number of learning trials presented when trial-by-trial, time-sample (observation of

a random trial during a session, not necessarily the first-trial), and estimations after the session

were used to collect data. The participants used each of the three data collection methods during

separate 3-min discrete-trial training sessions. A researcher also observed the same session and

obtained trial-by-trial data by tallying the number of learning trials that the participant presented

during each 3-min block. The most trials were presented with estimation (12.2 trials per 3 min),

followed by single-trial recording (9.3 trials), and trial-by-trial data collection (8.7 trials).

Taubman et al. only presented group data (for three participants), and it is impossible to

determine if their results were consistent across participants or if they generalize to other forms

of discontinuous data collection, such as first-trial recording. In order to provide further

information about how data collection strategies affect the number of learning trials arranged for

participants, the purpose of this study was to compare the number of learning trials presented by

behavior therapists doing their everyday work when they used trial-by-trial and first-trial recording.

Method

Participants, Setting, and Materials

Participants. Two male and four female interventionists who provided 15-30 weekly hours of in-home ABA interventions to children diagnosed with autism participated. All of the children with whom the participants worked with also received speech and or occupational therapy by qualified professionals. All of the participants had completed some college-level coursework in psychology or a related field. Lily, Robin, and Tracy were enrolled in an applied behavior analysis (ABA) master's program and the others had obtained a bachelor's degree or were enrolled in university courses working towards a bachelor's degree. Theodore, Barney, Lily, and Stella had less than four months of experience in providing ABA interventions and Robin and Tracy had more than one year of experience at the start of the study. For Theodore, we compared the two experimental conditions when he was hired initially, but due to case transitions, we were not able to replicate the comparison until a year later when he had accrued over one year of experience.

All of the participants were trained in the application of behavioral principles when they joined the organization. The training consisted of an initial two-week classroom-based training, followed by at least five days of in-field training. The classroom-based training involved verbal instruction, in-person role play, feedback, and assessment of skills. The participants were trained on reinforcement procedures, extinction procedures, prompting and prompt fading, data recording using continuous and discontinuous methods, instructional techniques (e.g., discrete-trials training), and related components of ABA until they performed specific skills at 80%

accuracy or higher during role-play scenarios. Upon meeting that performance criterion, the

participants shadowed an experienced therapist in the field for at least one week and received in-

situ feedback. The participants also attended monthly refresher trainings online or in-person

thereafter.

       Setting and materials. All sessions took place in the clients' homes during regularly

scheduled therapy sessions Monday-Saturday between 8 a.m. and 7:30 p.m. Research assistants

gathered data on a paper data sheet and videotaped sessions using a secure smartphone video

camera. The participants gathered data using a trial-by-trial data sheet and a first-trial data sheet

(in different phases of the study). Therapy sessions involved the use of naturalistic teaching

strategies during play and daily routines. During all experimental sessions, the participants also

worked with their clients on specific lessons that were prescribed by the Board Certified

Behavior Analyst® (BCBA®) overseeing the case, including medically-necessary therapeutic

interactions provided in the client's home or daily environment.

       The data sheets used by participants were regularly used at the organization where they

worked, and were not designed for this study. The trial-by-trial data sheet comprised five

columns. The first column listed the target behavior to track, the second stated the type of data

collection to use ("trial-by-trial"), the third was used to tally the number of trials that were

presented, the fourth was used to tally the number of correct responses, and the fifth was used to

calculate and record the percent of trials with a correct, independent response.

       The first-trial data sheet comprised four columns. The first column listed the target

behavior to track and the second stated the type of data collection to use ("first-trial"). The third

column contained a series of "+" and "-" symbols; for each trial, the participant circled the

appropriate symbol to indicate whether the client had (+) or had not (-) emitted a correct,

unprompted response. The fourth column was used to calculate the overall performance of the client.

Design and Procedures

This study was approved by an appropriate Human Subjects Institutional Review Board. An alternating-treatments design was used and the comparisons were replicated for a second client behavior for five of the six participants, referred to as different phases hereafter. Each experimental day consisted of at least one trial-by-trial session and one first-trial session, each lasting 10 min, with the sessions alternating thereafter during that particular experimental day. The first condition of each experimental day was selected randomly by holding up the blank sides of both data sheets, so that the data collection type was not visible, and asking the participant to pick one. This procedure resulted in the first-trial condition being implemented as the first condition during 18 of 33 (54%) experimental days sampled. Four research assistants were trained in the experimental protocols and conducted the experimental protocols with 100% fidelity as measured during role-play and the first experimental session.

The participants provided regularly scheduled ABA interventions using naturalistic or structured teaching approaches based on the needs of the clients who they served. The lessons were interspersed so that the participants presented the learning trials for the target behavior with other lessons. At the start of each session, the research assistant gave the proper data sheet to the participant and said, "As you run this lesson, please take (trial-by-trial/first-trial) data; otherwise, please run your session as you regularly would." The participant then gathered data on a specific client target behavior (Table 1) using the assigned data sheet. When trial-by-trial recording was in effect, the participant recorded the client's performance on each individual trial. When first-trial recording was used, the participant recorded only the first trial.

At the start of each condition, the researcher started a timer and the video recorder and began tallying the number of learning trials that the participant presented during the session. If the participant asked any questions about the data collection methods, the assistant stated that no additional information could be provided and instructed the participant to continue the session. The researchers did not provide any feedback during or after the sessions. The research assistant only paused the timer if there was a pause in the session for a legitimate reasons (e.g., parents asking questions to participants, bathroom breaks, etc.). Each session lasted 10 min and each experimental day comprised no more than two sessions per condition (i.e., no more than two first-trial and two trial-by-trial sessions per day). The study was conducted over a 14 month period and data were gathered across seven days for Theodore, five days for Barney, seven days for Lilly, three days for Stella, 10 days for Robin, and 12 total days for Tracy.

Dependent variable. The measure of interest in this study was the number of learning trials presented per 10-min session. A learning trial started when the participant provided an instruction or arranged the learning opportunity and a trial ended when the client engaged in the response correctly or with a prompt. Two independent observers tallied the number of trials presented by the participants and calculated total IOA during 38% of sessions by dividing the larger number of trials recorded by the smaller frequency observed and multiplying by 100, which yielded a score of 98.1% (ranging from 83-100% across sessions).

As a secondary measure, we determined the percentage of responses that a client emitted independently each experimental session when trial-by-trial recording and when first-trial recording was utilized.

Analysis. We utilized visual inspection as the primary method of data analysis and to assess stability. In addition to visual analysis, we compared the total mean responses that

occurred during each of the two experimental conditions (i.e., trial-by-trial vs. first-trial recording) for all participants by conducting a paired-samples *t*-test. Finally, we compared each participant's mean trial-by-trial and mean first-trial responses during each replication phase separately.

Results

Data for all participants appear in Figure 1. There was no consistent difference in the number of trials arranged for clients when participants used trial-by-trial versus first-trial data recording, and no individual participant arranged more trials when first-trial recording was used than when trial-by-trial recording was in effect. One participant, Barney, presented substantially more trials with trial-by-trial recording during the final four sessions during the first exposure to these conditions, but not during the second. Another participant, Robin, presented slightly more trials on some of the trial-by-trial sessions. With these exceptions, data paths for the two conditions were overlapping and generally similar. It is possible that a more clear difference would have emerged if the experimental sessions continued, but due to practical barriers in continuing the study, we were not able to continue these sessions in the setting where the study was conducted.

A comparison of aggregate data across all participants indicated that trial-by-trial recording resulted in 12.7 (SD = 7.9) learning trials per 10-min bin and first-trial resulted in 10.3 (*SD* = 7.7) learning trials. This difference was statistically significant *t*(71) = -3.703, p < .001. Within-subject comparisons of means revealed a similar pattern; individual trial-by-trial means were higher than first-trial means in 10 of 13 comparisons (Table 2).

The client data are presented in Figure 2. In the majority of cases, data obtained with the two recording methods did not support similar conclusions, indicating skill mastery at different

points. For example, when applying a skill mastery criterion of two consecutive sessions at 80%

or higher accuracy on trial-by-trial recording, versus two consecutive sessions with the first-trial

being correct during discontinuous recording, first-trial recording indicated skill mastery sooner

for some participants (e.g., Theodore and Robin's client data during the first lesson) but the

opposite was true for some other participants (e.g., Tracy and Barney's client data during the first

lesson). In addition, due to the binary nature of first-trial data, the data were more variable with

first-trial recording, thereby, providing less sensitive information about the progression of

learning. This point is most clearly illustrated by Robin's and Tracy's clients' data.

Discussion

These findings differ from those of a prior study, which reported that more trials were

completed when data from only a single trial were recorded (Taubman et al., 2013). Noteworthy

differences between that study and ours may explain the discrepant findings, one obvious

difference being the specific data collection method used. We compared first-trial and trial-by-

trial data collection methods, but Taubman et al. compared trial-by-trial, time-sampling

(examining performance during a single trial that occurred at a particular time), and estimation.

A second difference is that their sessions were 3 min in length, whereas ours were 10 min.

Although both time periods are relatively brief – perhaps too brief to provide a solid estimate of

performance during ordinary therapy sessions (which ordinarily lasted 3 hours for our

participants) – the difference in length may have influenced the obtained results. The same is true

of the specific kind of trials that participants arranged, which differed in the two studies.

Taubman et al. had all participants use a consistent discrete-trials training procedure, whereas we

had participants arrange a variety of naturalistic trials, as necessitated by the needs of the clients

with whom they worked. Perhaps because of this, participants in the Taubman et al. study completed substantially more trials per unit time than our participants.

Another difference is how, exactly, data were collected. The data sheets and data collection procedures used in the present study were designed by members of the organization where the participants worked to yield accurate data that could easily be collected, although whether they did so had not been formally evaluated. Taubman et al. (2013) did not report how, exactly, participants recorded data, or whether the procedures they used were intended to be easy to use.

A number of factors may have contributed to the variability in our data, one being the amount of experience that each participant had as a behavior therapist. Tracy had relatively more experience as a therapist compared to the other participants. The experimental effects were slightly less pronounced for Tracy than they were for the other less experienced participants. Similarly, by the time we replicated the treatment effects on a second client behavior for Theodore, he had accumulated over one year of experience, which yielded a much smaller difference between conditions than the first comparison did a year prior. These patterns indicate that therapist experience may have contributed to the results. Other factors may have contributed as well, such as the client lessons recorded, which differed substantially in the amount of preparation time they each needed. In addition, the participants taught lessons for which data were not recorded during experimental sessions (Table 1), including at least one mand lesson. In some sessions, establishing operations relevant to particular mands were in effect, and a substantial amount of time was allocated towards teaching and reinforcing the mand instead of proceeding with the target lesson, which may have increased variability in the data. Although no clear and consistent differences emerged visually because of the variable data, the aggregate data

revealed a different picture. Specifically, trial-by-trial recording resulted in more learning opportunities being presented than first-trial recording. Moreover, most participants arranged more trials when trial-by-trial data were collected.

It is noteworthy that Taubman et al. (2013) determined the accuracy of each data collection strategy they examined and found that trial-by-trial recording was most accurate. This outcome is important, and consistent with the general finding that continuous recording is more accurate than discontinuous recording (e.g., Poling, Methot, and LeSage, 1995). Although we did not compare which method was more accurate in depicting learner performance, we compared client performance data across the two data collection methods and the data did not correspond with one another. In some cases first-trial recording resulted in premature indices of skill mastery, as was also the case for Carey and Bourret (2014), but in some cases the opposite was true. It should be noted that the client data in this study represented multiple target responses, not just the same response within a lesson, which could have accounted for why we did not obtain the same results as Carey and Bourret; therefore, our student performance results should be viewed with caution.

When continuous data recording, such as trial-by-trial recording can be arranged, it is generally preferred by behavior analysts (e.g., Poling al., 1995). There is good reason for this preference: Trial-by-trial data collection typically is more accurate and more sensitive (Carey & Bourret, 2014; Lerman, Dittlinger, Fentress, & Lanagan, 2011) than first-trial or other discontinuous methods. Carey and Bourret found that using discontinuous recording did not save significant time, and we found that doing so did not reduce the number of trials presented by therapists. Given these findings, it appears that continuous data recording is both possible and preferable.

Interestingly, our participants typically arranged more, not fewer, trials when using trial-by-trial recording. There is no obvious explanation for this finding, but it is possible that recording successful student responses reinforced the chain of participant responses that led to client success, and hence increased the speed at which those responses were subsequently performed, and therefore the number of trials arranged. This would be a highly desirable outcome, and the possibility that it occurs merits investigation.

There are noteworthy limitations in this study. For one, we conducted the comparison when recording only one client behavior with the target data sheets instead of using each method across all target client lessons at once. Due to clinical needs, the latter was not possible because some lessons required continuous recording and discontinuous recording would not have provided the necessary information for the clinical teams. Relatedly, a second limitation was that each session lasted only 10 minutes, which was in contrast with the typically scheduled 2-3-hr ABA therapy sessions provided at the organization. However, it should be noted that although a 10-min observation window may not provide a full picture of behavior in the natural setting, it is roughly proportionate to the total amount of time that a therapist uses to teach a particular lesson in a 2-3-hr therapy session comprising many lessons. For example, a 2-hr therapy session that consists of 12 lessons allows 10 minutes for each lesson. A third limitation is that we did not measure whether or not the participant's behavior influenced the rate of client skill acquisition. A fourth limitation is that practical considerations prevented us from continuing sessions until performance stabilized. Finally, although we hoped to do so, we were unable to collect social validity data, thus it is not clear whether participants found either data collection strategy to be acceptable, and hence likely to be used consistently. Future research should examine this possibility.

In summary, unless there is good reason to do otherwise, behavior analysts should consistently use continuous data collection strategies. A substantial reduction in learning opportunities due to the use of continuous data recording would constitute a good reason, but no such reduction was evident in the present study. Instead, there was evidence that more trials often were presented when data were recorded continuously. Further research is needed to ascertain the conditions under which this relation obtains, the costs and benefits of different data collection strategies in diverse situations, and the conditions under which particular strategies are warranted.

References

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still-current dimensions of applied

behavior analysis. *Journal of Applied Behavior Analysis*, *20*, 313-327.

https://doi.org/10.1901/jaba.1987.20-313

Bahadourian, A. J., Tam, K. Y. B., Greer, R. D., & Rousseau, M. K. (2006). The effects of learn

units on student performance in two college courses. *International Journal of Behavioral

Consultation and Therapy*, *2*, 246. http://dx.doi.org/10.1037/h0100780

Carey, M. K., & Bourret, J. C. (2014). Effects of data sampling on graphical depictions of

learning. *Journal of Applied Behavior Analysis*, *47*, 749-764.

https://doi.org/10.1002/jaba.153

Cummings, A. R., & Carr, J. E. (2009). Evaluating progress in behavioral programs for children

with autism spectrum disorders via continuous and discontinuous measurement. *Journal

of Applied Behavior Analysis*, *42*, 57-71. doi: 10.1901/jaba.2009.42-57

Giunta-Fede, T., Reeve, S. A., DeBar, R. M., Vladescu, J. C., & Reeve, K. F. (2016).

Comparing continuous and discontinuous data collection during discrete trial teaching of

tacting by children with autism. *Behavioral Interventions*, *31*, 311-331.

doi:10.1002/bin.1446

Greer, R. D., McCorkle, N., & Williams, G. (1989). A sustained analysis of the behaviors of

schooling. *Behavioral Interventions*, *4*, 113-141. https://doi.org/10.1002/bin.2360040205

Leaf, J. B., Cihon, J. H., Leaf, R., McEachin, J., & Taubman, M. (2017). A progressive

approach to discrete trial teaching: Some current guidelines. *International Electronic

Journal of Elementary Education*, *9*(2), 361-372.

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*, 53-62. doi:10.1007/BF03391775

Najdowski, A. C., Chilingaryan, V., Bergstrom, R., Granpeesheh, D., Balasanyan, S., Aguilar, B.,…& Roane, H. (2009). Comparison of data-collection methods in a behavioral intervention program for children with pervasive developmental disorders: A replication. *Journal of Applied Behavior Analysis*, *42*, 827-832. doi: 10.1901/jaba.2009.42-827

Poling, A., Methot, L. L., & LeSage, M. G. (1995). *Fundamentals of behavior analytic research*. New York: Plenum Press.

Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., ... & McNerney, E. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*, 2411-2428. doi: 10.1007/s1080

Taubman, M. T., Leaf, R. B., McEachin, J. J., Papovich, S., & Leaf, J. B. (2013). A comparison of data collection techniques used with discrete trial teaching. *Research in Autism Spectrum Disorders*, *7*, 1026-1034. doi: 10.1016/j.rasd.2013.05.002

Figure Captions

Figure 1. Number of learning trials presented during ten minute trial-by-trial and first-trial data recording sessions for Theodore, Barney, Lily, Stella, Robin, and Tracy. Note, the ordinate scales vary across participants.

 Figure 2. Percentage of trials that the participant's client engaged in a correct response during trial-by-trial recording and correct (100%) or incorrect (0%) responses on the first-trial recorded during first-trial recording.
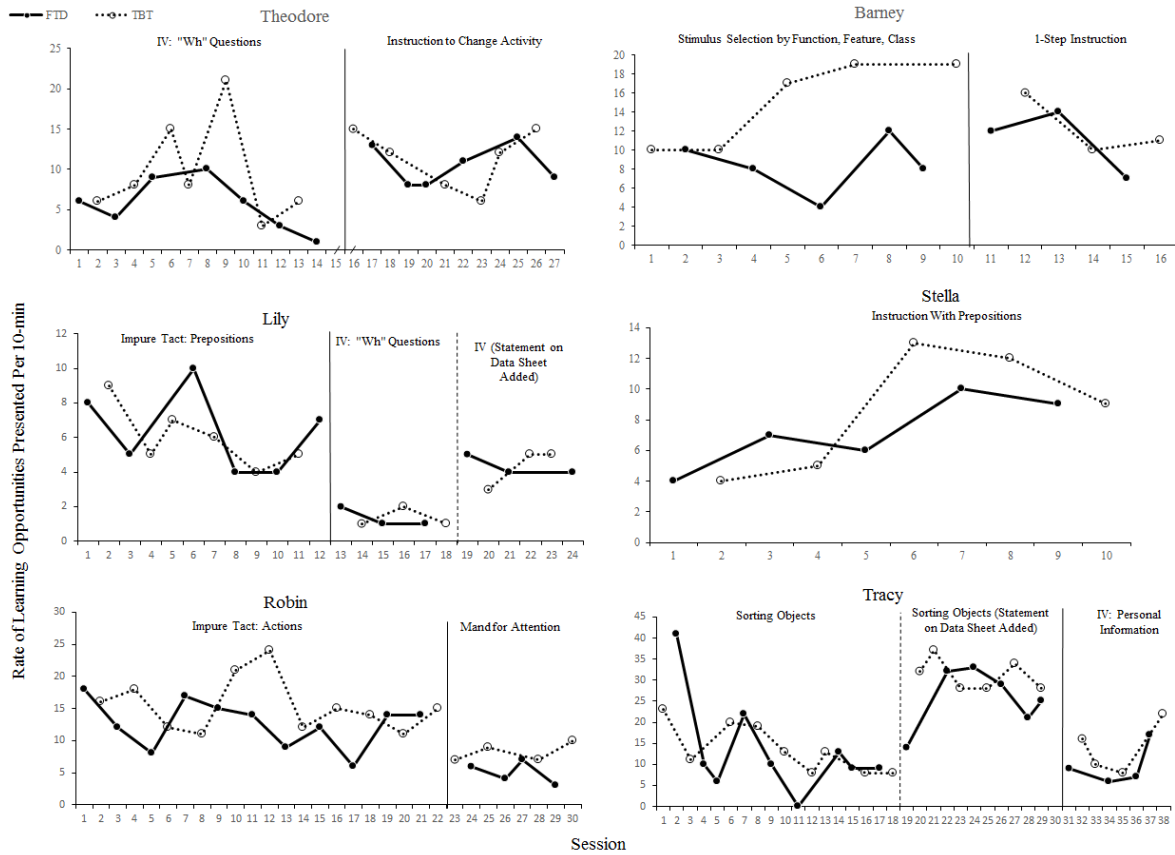
Figure 1 (Top)

Figure 2 (Top)
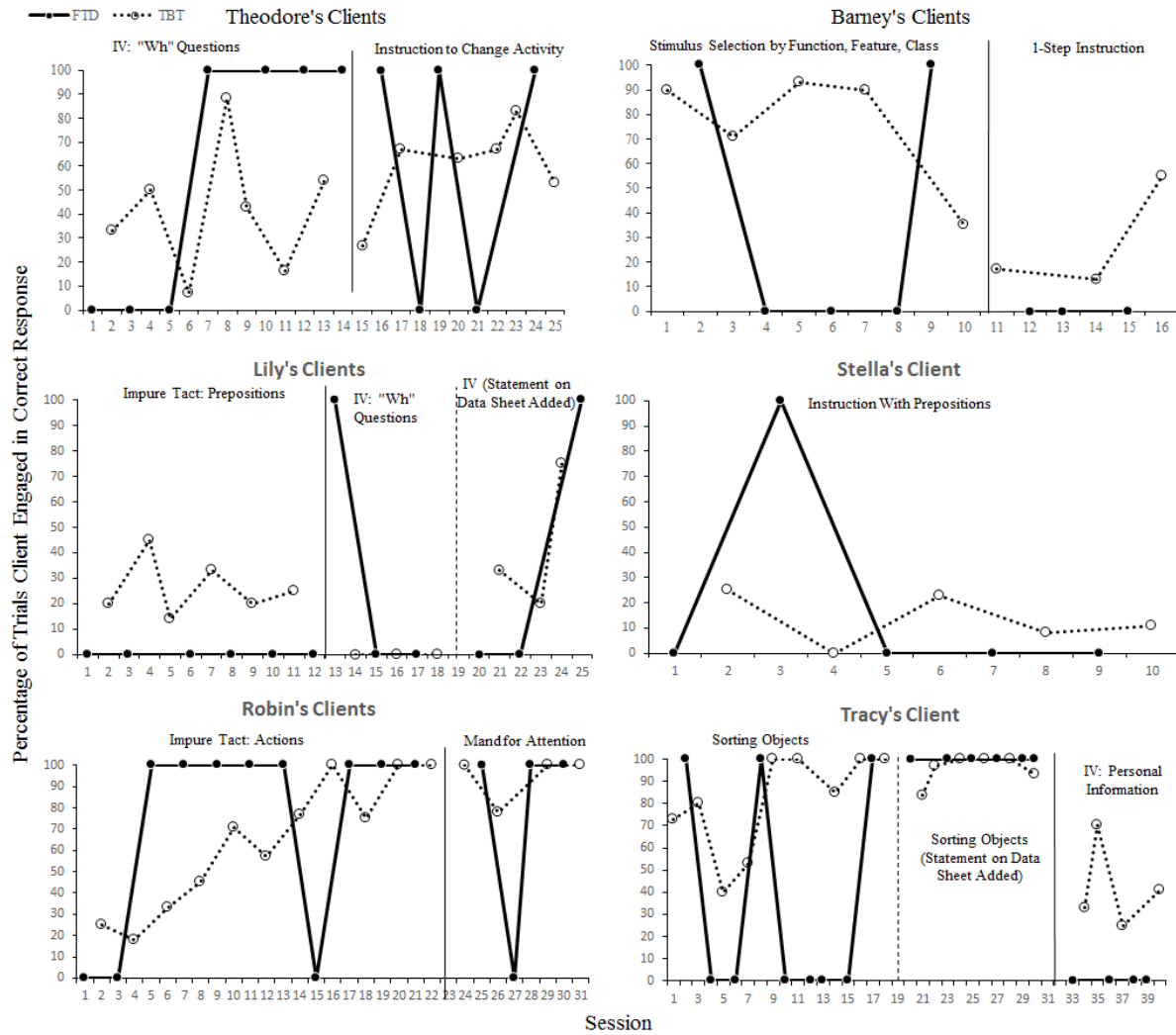
Table 1

List of Client Target Lessons that Each Participant Gathered Data for, the Operational Definitions for the Start and End of Each Learning Trial, Client Characteristics, and Other Common Lessons Targeted

Participant	Client Target Lesson	Trial Start: Participant's Response	Trial End: Child's Response	Client Characteristics at Start	Other Common Lessons Targeted During Sessions
Theodore					
Client 1	Intraverbal: "wh" questions	Asks a questions starting with "who," "what," or "where"	Answers the question within 5-s using a complete sentence	5 years; Level 3 VB-MAPP; 3 years in ABA	Manding for information, attention, tacting directions, and intraverbal (answering questions)
Client 2	Instruction to change activity	States a change in routine will take place (e.g., "We are going to play puzzle next")	Complies without engaging in problem behaviors	4.5 years; Level 3 VB-MAPP; 0.5 years in ABA	Manding for cessation, tacting pronouns, tacting shapes, 3-step instructions, fine-motor imitation, and play with less-preferred items
Barney					
Client 1	Stimulus selection by function, feature, class (clock)	Places array of items and asks, "which one tells time?"	Points to the picture of a clock	4 years; Level 3 VB-MAPP; 1 year in ABA	Manding for items, manding for attention, manding for cessation, responding to name, matching-to-sample, 1-step gross motor imitation, and responding to greetings

Client 2	1-step instruction (gross-motor)	Provides 1-step instruction (e.g., "clap")	Engages in 1-step motor movement within 5-s	2 years; Level 1 VB-MAPP; 0.5 years in ABA	Manding for objects, responding to name, single syllable echoic, and gross motor imitation
Lily					
Client 1	Impure Tact: prepositions	Asks, "Where is the boy" with visual stimulus present	Vocally responds with correct preposition ("on top of," "under," ...)	3.5 years; Level 2 VB-MAPP; 1 year in ABA	Manding for objects, tacting noun-verb combinations, responding to name, matching-to-sample (block stacking), 3-syllable echoic, fine motor imitation, and intraverbal (answering questions)
Client 2	Intraverbal: "wh" questions	Asks a question starting with "who," "what," or "where"	Answers the question within 5-s	7 years; Level 3 VB-MAPP; 3 years in ABA	Manding for help and intraverbal (safety questions)
Stella					
Client 1	Instruction with prepositions	Says, "Put the block next to/under/on top of the house"	Physically places item in indicated position	3 years; Level 2; 1 year in ABA	Manding for attention, manding for cessation, responding to name, matching-to-sample (non-identical), and fine motor imitation

Robin

Client 1	Impure Tact: Actions	Asks, "What is the boy doing?" with visual stimulus present	Vocally indicates the action (running, walking, ...)	2.5 years; Level 2 VB-MAPP; 1 year in ABA	Manding for items, manding for attention, manding for cessation, tacting items, listener responding to identify body parts and objects, 1-syllable echoic, and responding to 1-step instructions
Client 2	Mand for attention: politely interrupting conversation	Turns away and begins speaking to another adult	Says, "excuse me."	7 years; Level 3 VB-MAPP; 2 years in ABA	Manding for cessation, perspective taking (tacting other people's emotions), and petting cat gently

Tracy

Client 1	Sorting objects	Places items in front of child and instructs to sort	Sorts items into correct pile	4 years; Level 2 VB-MAPP; 2 years in ABA	Manding for items, manding for cessation, tacting objects, tacting actions, responding to name, listener responding to identify people and body parts, 1-step gross-motor imitation, and intraverbal (song fill-ins)
Client 1	Intraverbal: personal information	Asks, "What's your address?" or "What's your mom's	Answers question vocally within 5-s		

| phone number?" | (address, phone number, etc.) |