February 2021

An Analysis of the Effects of Extinction Relative to Baseline Measures Including and Excluding Consumption Time

Kelti Keister
kkeister@rollins.edu

An Analysis of the Effects of Extinction

Relative to Baseline Measures Including and Excluding Consumption Time

A Thesis

Kelti Keister

Rollins College

Applied Behavior Analysis and Clinical Science

Acknowledgements

I would like to start by thanking my partner, Tyler Harmon, for supporting me while working on this project. Without their consistent and unconditional support, I would not have been able to focus on developing and completing my thesis. Next, I would like to thank my father, Bryant Keister, for always being available to talk when this project became overwhelming and helping me focus on the bigger picture when I needed it. I would also like to thank my friends, Gabby Moore and Julia Truman, for loving me even though I was physically and mentally absent during this time. Thank you to my mentors, Nicole Forbes, Erin Lopater, and Luz Cabrera, for encouraging me and believing in me when I did not. And finally, I would like to thank Dr. Stephanie Kincaid and Dr. Rachelle Yankelevitz for their wisdom and guidance. You all made it easier to feel connected and loved during a time where distance and isolation were all too prevalent. Thank you again.

Table of Contents

Abstract

Extinction bursts are a widely discussed phenomenon. The data analysis methods used to analyze the effects of extinction bursts have not been examined thoroughly within the field of applied behavior analysis, however. In this study, we implemented extinction procedures in a reversal design on three typically developing pre-school aged individuals. We then conducted two different data analysis methods. First, extinction effects were evaluated relative to one baseline measure including reinforcer consumption time and one baseline measure excluding reinforcer consumption time. Second, extinction was evaluated at the whole- and within-session (last five minutes of baseline and first ten minutes of extinction) level of analysis, based on procedures implemented by Katz and Lattal (2020). In total 1 extinction burst was observed out of 5 baseline to extinction transitions, when reinforcer consumption time was included. When reinforcer consumption time was excluded, the magnitude of this extinction burst was minimized. No additional extinction bursts were uncovered at the within-session level that were not observed at the whole-session level. The results of both data analysis methods are further discussed with regard to the implications they hold for utilizing rate of responding as a dependent measure when evaluating extinction effects in clinical treatment settings.

An Analysis of the Effects of Extinction Relative to Baseline Measures Including and Excluding Consumption Time

Extinction is a widely discussed procedure within the field of behavior analysis. In their widely disseminated textbook, Cooper and colleagues describe the effects of extinction as follows, "if reinforcement is withheld for all members of a previously reinforced response class... the frequency of the behavior will gradually decrease to its pre-reinforcement level or cease to occur altogether" (Cooper, Heron, & Heward, 2007, p. 37). The effects of operant extinction were first empirically evaluated by Skinner, when the operant chamber he was using to conduct an experiment malfunctioned and reinforcer deliveries ceased, resulting in a decrease in responding (described by Iversen, 1992). Since then, researchers have developed a robust literature evaluating extinction and its effects. Following this progress, applied researchers have employed extinction as a therapeutic intervention in the treatment of problem behavior. Extinction has been implemented to effectively treat a wide variety of topographies of problem behaviors, including self-injurious behavior (e.g., Zarcone et al., 1993), aggression (e.g., Lerman, Iwata, & Wallace, 1999), and object destruction (e.g., Magee & Ellis, 2000).

Although extinction has the ability to eliminate problem behavior rapidly (Repp, Felce, & Barton, 1988), it is also capable of evoking various undesirable effects (Lerman, Iwata, & Wallace, 1999). The potential byproducts related to the implementation of extinction procedures include extinction-induced variability (e.g., Lattal et al., 2013), extinction-induced aggression (e.g., Azrin, Hutchinson, & Hake, 1966), and the extinction burst (e.g., Goh & Iwata, 1994). Perhaps due to its profound implications for treatment success or failure, the extinction burst has received some attention within the applied literature. The extinction burst is commonly and simply defined as, "an increase in responding" following the implementation of an extinction procedure (e.g., Cooper, Heron, & Heward, 2008, p. 340). Other authors have defined specific criteria for extinction bursts such as, “an increase in responding during any of the first three treatment sessions above that observed during all of the last five baseline sessions (or all of baseline if it was briefer than five sessions)” (Lerman & Iwata, 1995, p. 93). In an unpublished review, Huffman (2019) noted that discrepancies in criteria for the extinction burst are prevalent within the literature. There is some disagreement regarding the topography of the burst (i.e., what the burst “looks like”), the temporal locus of the burst (i.e., how soon in the extinction process does this effect occur), and the magnitude of the burst (i.e., how much of an increase in responding needs to occur to qualify as a burst). Such discrepancies in defining the extinction burst display the potential sensitivity of this effect to inconsistent interpretations, which in turn could greatly alter the perceived prevalence of the extinction burst in the literature.

Extinction bursts are often described as a common extinction-related effect (e.g., Kazdin, 2012, p. 360), but just how common is this phenomenon? Two studies have examined the prevalence of extinction bursts in intervention contexts. Lerman and Iwata (1995) assessed 113 data sets in which extinction was implemented as a treatment component to determine how frequently extinction bursts occur. The results indicated that 27 out of the 113 data sets (24%) portrayed an extinction burst, according to the author’s criteria. Subsequently, Lerman, Iwata, and Wallace (1999) investigated multiple side effects of extinction, including extinction-induced aggression and extinction bursts, by conducting an analysis of their own clinical data sets between the years of 1989 and 1997. Similar criteria to the 1995 study for the occurrence of extinction bursts were employed. Of the 41 cases evaluated, 16 (39%) exhibited response bursting.

Despite these reported estimates of prevalence, and the substantial concerns regarding the impact of extinction bursts on treatment outcomes (e.g., as discussed by Lerman, Iwata, &

Wallace, 1999), little research has been carried out to systematically evaluate extinction bursts experimentally (Lattal et al., 2013). One of the few systematic investigations of the extinction burst was recently conducted by Katz and Lattal (2020). The authors implemented repeated cycles of reinforcement and extinction to evaluate the presence of an extinction burst following VR schedules (i.e., Experiment 1 & 2) and continuous reinforcement schedules (i.e., Experiment 3). Interestingly, the results suggested that the method of data analysis substantially impacted their ability to detect extinction bursts. Effects were evaluated at three levels: whole-session, local (first min of extinction cf. baseline), and local-end (the first session of extinction cf. the first min of the last five baseline sessions). In Experiment 2, when the data were analyzed at the whole-session level, extinction bursts (as characterized by the criteria set by Lerman & Iwata, 1995) occurred in 34.8% of datasets. However, when the within-session data were analyzed, extinction bursts occurred approximately 48% of the time in local analyses and 87% of the time in local-end analyses. Katz and Lattal concluded that the type of transition (e.g., whole- vs. within-session transitions) between baseline and extinction has an influence on the frequency of observed extinction bursts. The researchers attributed this conclusion to evidence suggesting that the occurrence of an extinction burst in the first extinction session can be “washed” or averaged out when data are summarized at the level of the whole-session response rate. Specifically, the authors noted that though responding may occur at a high rate during the first few minutes of the session because considerably lower response rates may occur later in the session, the extinction burst effect may be missed. In summation, the authors conclude that extinction bursts may be sensitive to data analysis manipulations; such that the method of analysis may appear to change the magnitude of a burst or make the extinction burst disappear altogether.

Evaluating effects under varying levels of analysis may also be instrumental in understanding efficacious procedures in applied settings. For example, Wunderlich and Vollmer (2015) evaluated the efficacy of response interruption and redirection (RIRD) under two different data presentation methods: percentage of session engaged in stereotypy the entire session, and percentage of session engaged in stereotypy outside of RIRD implementation. The results indicated that RIRD seemed to be an effective treatment for suppressing stereotypy when evaluated outside RIRD implementation, but when evaluated during the entirety of the session, reductions were minimized. In this way, RIRD appeared to be effective because the time spent implementing the procedure placed a "ceiling" on the rate of problem behavior when analyzed at the whole-session level. Thus, the authors suggest that the levels of analysis used in previous research might have contributed to the overestimation of RIRD's efficacy in reducing stereotypy.

A similar ceiling-like effect may affect the evaluation of extinction bursts. As previously discussed, extinction bursts are assessed relative to rate of responding observed in the immediately preceding baseline phase. In applied investigations, baseline typically consists of continuous reinforcement procedures. When the behavior is reinforced, the participant is typically allowed 20 to 30 s to consume the reinforcer (e.g., tangible item, escape from the demand, etc.). This consumption time is typically included in the whole-session response rate, which is reasonable in principle because responding could technically occur during the reinforcement period. However, if efficient discriminated responding is observed in the baseline, responding is unlikely to occur during consumption intervals. Thus, including consumption time in the rate calculation is likely to result in a rate of baseline responding that rarely (if ever) exceeds 2 responses per minute (RPM). While not a true ceiling effect, this ceiling-like constraint on response rate is effectively removed when extinction is subsequently implemented (and consumption time is no longer provided) in the extinction phase. This could result in a pattern of responding resembling what is commonly defined as an extinction burst that is, in fact, an artifact of the method of data analysis. Similar to how Wunderlich and Vollmer (2015) suggested that data analysis conventions might have contributed to the overestimation of RIRD's efficacy in reducing stereotypy, data analysis conventions may have also contributed to the overestimation of the magnitude and prevalence of extinction bursts. Even if the extinction burst is not shown to be entirely an artifact of the data analysis (i.e., does not disappear when consumption time is excluded from baseline), it is possible that the magnitude of the burst may be substantially reduced by analyzing the data in this manner. That is, if baseline is artificially suppressed by consumption time, then removing consumption time will increase the level of baseline responding. Thus, the increase in responding demonstrated in extinction, even if it is above baseline rates, may not be as significant of an increase as previously thought. Thus, excluding consumption time may cause the overall pattern of responding to "look different," resulting in different conclusions when such data are analyzed via visual analysis.

The present study evaluated the effects of extinction under various data presentation methods. Similar to Wunderlich and Vollmer (2015), extinction was evaluated relative to one baseline measure including consumption time in the analysis and one baseline measure that did not. Similar to Katz and Lattal (2020), extinction was also evaluated at the within-session level. We hypothesized that if extinction bursts were observed, this effect would be highly sensitive to data analysis manipulations. Additionally, we investigated how frequently extinction bursts occur when extinction procedures are implemented. We hypothesized that we would find evidence that extinction bursts occur at rates similar to those found by Lerman and Iwata (1995).

Method

Participants, Setting, and Materials

Three typically developing individuals, Jacob, Gabby, and Martha, ages 3- to 5-years-old, were recruited from preschools in the Central Florida region. Two of the participants were female (Gabby & Martha) and one was male (Jacob). Sessions took place in a classroom that contained a one-way mirror, video camera, and microphone. The room also included a child-sized table, two child-sized chairs, and typical classroom features (e.g., whiteboards, markers, tables, and bins of miscellaneous materials; these were cleared from the participant's immediate area).

In the preference assessment sessions, leisure items were present. Examples included but were not limited to coloring books, wooden trains, water color paints, wooden blocks, Play-doh, and wooden figurines.

In baseline and extinction sessions, laminated communication cards, measuring 0.05 m by 0.05 m, displaying screen captures of the video that the card corresponded with, were provided for each session. Images used for each communication card were individualized per participant based on the indirect assessment results and were selected by the researcher based on how well the image represented the video or leisure item. One communication card per video or leisure item per participant was provided for each session.

Response Measurement and Reliability

In the preference assessment phase, the target response was item selection. A selection was recorded when the participant touched an item on the table in front of them.

In baseline and extinction phases, the target response was communication card exchanges. This study utilized translational research methods, thus, this (somewhat arbitrary) target response was selected to be analogous to problem behavior maintained by access to tangible reinforcement. A target response was recorded when the participant handed a communication card to the researcher. The researcher accepted the response when the participant placed the communication card on or in the researcher's hand. The researcher then reset the card to the location it was taken from, 12.7 cm in front of the participant, which was marked by an X made of tape. Observers collected frequency data on communicative responses, and these data were converted to response per minute (RPM) for each session by dividing the total number of responses in that session by the duration of the session.

To permit analysis of response rates including and excluding reinforcement time, observers also recorded duration of reinforcer consumption for each session. Reinforcer consumption time began when the researcher placed the unlocked iPad with the preferred video paused and available on the screen or leisure item on the table in front of the participant and ended when the researcher paused the iPad and removed the iPad or leisure item from the participant.

In baseline and extinction sessions, a moderately preferred item was continuously available. Duration of engagement with the moderately preferred item was measured when the participant touched or attended (i.e., defined as, the participant's body and eye gaze were simultaneously oriented towards the moderately preferred item) to the item until the participant placed the item down or was no longer attending to the item.

Throughout the study, researchers used the app "Countee" to collect data from videotaped recordings of the sessions. For at least 30% of sessions, two observers independently recorded data. Independent observer agreement (IOA) data were used to quantify the reliability of data collected. IOA was calculated through proportional agreement for communication card exchanges and averaged 90.83% (range: 80% - 96.67%), 90% (range: 73.33% - 100%), 95%

(range: 80% - 100%), reinforcer consumption time and averaged 91.23% (range: 83.65% - 100%), 95.39% (range: 81.57% - 100%), 96.38% (range: 85.53% - 100%), and moderately preferred item engagement and averaged 100% (range: 100%-100%), 97.45% (range: 93.33% - 100%), 93.84% (range: 82.97% -100%), for Jacob, Gabby, and Martha, respectively.

To assess the reliability of the independent variable, treatment integrity data were collected for a minimum of 30% of sessions for each participant. An observer collected data on correct implementation of relevant antecedents and consequences for each condition, as detailed below. Treatment integrity was scored in 30-s intervals, in which the observer marked a 1 if the procedures were implemented correctly and a 0 if they were not: communication card placement, moderately preferred item accessibility, instruction delivery, highly preferred item interval ranges (i.e., 25- to 35-s access period), and communication card replacement following an exchange. This was converted into an overall percentage of intervals with 100% correct implementation of procedures. Treatment integrity measures were high across participants (90%, 97.5%, and 97.5% for Jacob, Gabby, and Martha respectively) and across phases (90% and 100% for baseline and extinction phases respectively).

Procedure

Sessions were conducted 1 to 2 days a week and multiple sessions were conducted per day. All baseline and extinction sessions lasted 5 min. An ABAB reversal design was used to demonstrate experimental control for two participants. For one participant, an ABA reversal design was utilized to demonstrate experimental control due to extenuating circumstances (described below). Prior to participation, the caretakers of participants returned an informed consent document, demographic survey, and preference questionnaire. Prior to the experiment proper, an indirect assessment and a preference assessment were conducted as described below.

Each phase of the reversal continued until responding was deemed stable via visual analysis of graphed data.

Indirect Assessment

At least one caretaker (i.e., parent or teacher) of the participant reported in writing, to the lead experimenter, 2-3 leisure activities their child frequently preferred to play with and 2-3 YouTube videos, television shows, or movies that the child frequently preferred to watch.

Preference Assessment

A multiple stimulus without replacement (MSWO) preference assessment was implemented using procedures similar to those as described by DeLeon and Iwata (1996) to assess preference for leisure items for each participant. Conducting a pre-experimental MSWO preference assessment served to identify highly preferred and moderately preferred items. The highly preferred item identified by the preference assessment served as the reinforcer in subsequent baseline phases and was unavailable during extinction phases. The moderately preferred item identified by the preference assessment was present in all subsequent conditions and served as an alternative response during extinction phases (i.e., to decrease the likelihood of problem behavior due to the sparse environment). Leisure items included activities found in the participant's natural classroom environment that the participant's caregiver indicated on the indirect assessment. Every MSWO preference assessment included 2-3 iPads, each with a video suggested by caregivers pre-loaded.

Prior to the assessment, the researcher conducted one exposure trial with each item. During exposure trials, the researcher prompted the participant touch the item, using least-to-most prompting, and then delivered 30-s access to the item. After all exposure trials were conducted, the researcher conducted the MSWO by presenting all leisure items (including the iPad). All items were placed equidistant from other items and the participant (i.e., approximately 0.3 m between each item, each item 0.3 m in front of the participant). The researcher verbally delivered the instruction, “Pick one.” Upon selection of an item, the researcher delivered 30-s access to the selected item, and then removed the item from the participant’s view. The item was not replaced in subsequent trials. All attempts to select more than one item at a time were blocked by the researcher. If no selection was made within 5 s, the researcher delivered the verbal instruction again. If still no selection was made within the next 5-s interval, the researcher removed the remaining activities on the table, and “no selection” was scored for that trial. Trials continued until all items were selected or no selection occurred.

Baseline

In each baseline (BL) session, the communication card identifying the highest-preferred item was placed on the marked spot on the table in front of the participant as described above, a moderately preferred item was available and within the participant’s line of sight, and the highly preferred item was on a side table next to the researcher. Prior to the session, the researcher conducted a brief training trial by stating, “if you give me this card, you can watch [name of preferred video]/play with [name of preferred leisure item]” and then prompted the target response using least-to-most prompting (verbal, gestural, modeled, and then physical prompting). The session proper began with the researcher verbally delivering the instruction, “Choose something to play with.” Each communication card exchange made by the participant resulted in 30-s access to the item represented on the communication card. Once a communication card was selected by the participant, the researcher immediately returned the card to its position on the table. If the target response occurred during reinforcement, the card was replaced but no other programmed consequences occurred. The participant had continuous access to the moderately preferred item. If 30 s elapsed in a session during the second BL phase (following the first extinction phase) and no target responses occurred, then a verbal prompt was issued by the experimenter (i.e., “The iPad is available.”). If a vocal mand was emitted by the participant, rather than engaging in the target response, a verbal prompt was delivered by the therapist stating, “If you want the iPad you can hand me the card.” BL sessions were conducted until responding was deemed stable by the lead researcher via visual analysis of graphed data.

Extinction

During the Extinction (EXT) phase, target responses no longer resulted in iPad or leisure item presentation. The materials presented were identical to BL sessions. Each target response (i.e., communication card exchange) made by the participant resulted in the researcher taking the card from the participant and immediately returning the card to its position on the table. The researcher did not provide any programmed stimulus changes following the emergence of novel problem behaviors (i.e., crying or vocal mands for the preferred item). All attempts to access the highly preferred item were blocked by the researcher. The participant had continuous access to the moderately preferred item in this phase.

Results

The preference assessment identified a wooden train set, Play-doh, and Paw Patrol as highly preferred items (i.e., reinforcers); and Superwings, assorted wooden blocks, and a watercolor painting set as moderately preferred items for Jacob, Gabby, and Martha respectively. Reinforcer consumption time was generally high across BL sessions for all participants (average of 73% of session, 3.65 min per 5-min session for Jacob; average of 92% of session, 4.6 min per 5-min session for Gabby; average of 84% of session, 4.2 min per 5-min session for Martha), relative to engagement with the moderately preferred item. Engagement with the moderately preferred item was 0% of the session across BL sessions for all participants. Engagement with the moderately preferred item was generally low across EXT sessions for all participants (average of 3% of session, 0.1 min per 5-min session for Jacob; average of 29% of session, 1.5 min per 5-min session for Gabby; and average of 21% of session, 1 min per 5-min of session for Martha), relative to reinforcer consumption time.

In this study, experimental control was demonstrated for each participant through a reversal design when target responding was lowered to near-zero levels each time the participant was exposed to the EXT contingency and when responding returned to pre-EXT levels when exposed to the second BL phase. To be consistent with standard practices within ABA, the reversal design was implemented based on the rate of responding calculated including reinforcer consumption time. These data appear in Figures 1, 3, and 5 (top panel). After each participant completed the study, subsequent data analyses were conducted, including calculation of response rate excluding reinforcer consumption time (the bottom panels of Figures 1, 3, and 5), and within-session (i.e., minute-by-minute) analyses (Figures 2, 4 and 6). Individual results for each participant will now be considered.

Figure 1 displays Jacob's target responding in responses per minute under two different analysis methods (i.e., including and excluding consumption time in the top and bottom panels, respectively). As seen in the top panel of Figure 1, BL response rates averaged just under 2 RPM and exhibited low variability. In both transitions from BL to EXT, an immediate decrease in level to near-zero responding was observed. This result was unanticipated for the first transition from BL to EXT, as the participant did not engage in the target response (and therefore responding did not contact the EXT consequence) until minute 5 of the first EXT phase (i.e., the last minute of Session 1 of EXT 1). In this study, all datasets were analyzed for the occurrence of

EXT bursts. An occurrence of an EXT burst in the transition from BL to EXT was concluded based on Lerman and Iwata's (1995) criteria that defined an EXT burst as, "an increase in responding during any of the first three treatment sessions above that observed during all of the last five baseline sessions (or all of baseline if it was briefer than five sessions)" (p. 93). Thus, for Jacob, no EXT bursts were observed at the whole-session level of analysis of two transitions from BL to EXT, when reinforcer consumption time was included (i.e., the top panel of Figure 1).

The bottom panel of Figure 1 shows Jacob's BL response rates in which reinforcer consumption time was excluded from the rate calculation. EXT data presented in this panel are identical to EXT data presented in the top panel because no reinforcer consumption occurred during the EXT condition. Because the majority of Jacob's BL sessions consisted of consumption time (i.e., approximately 3.5 min per 5-min session), excluding consumption time resulted in an increased BL level of responding (range: 4.5 to 6.5 RPM). Additionally, excluding consumption time resulted in a different trend in responding in the first BL (i.e., no trend when consumption was included, an increasing trend when consumption was excluded). Overall, for Jacob, there were no EXT bursts demonstrated in the transitions from BL to EXT.

Figure 2 displays the within-session (i.e., minute-by-minute) data for Jacob. In Figure 2, the frequency of responses for the last five minutes of BL and the first ten minutes of EXT are presented in one-min bins. As in Figure 1, the top panel includes reinforcer consumption time, the bottom panel excludes reinforcer consumption time, and EXT session data are identical across panels. The within-session analysis was included based on the findings of Katz and Lattal (2020), which found evidence to support the claim that analyzing within-session transitions from BL to EXT can increase the frequency of observed EXT bursts, as compared to whole-session analyses. Based on the criteria set by Lerman and Iwata (1995), no EXT bursts were observed at the within-session level in the transition from BL to EXT. While the within-session data did not reveal any additional EXT bursts, this analysis did illuminate a more detailed pattern of responding during the transition from BL to EXT (i.e., immediate elimination of Jacob's responding in the first transition from BL to EXT but not in the second). In summary, no additional EXT bursts were observed by analyzing data at the within-session level of analysis for Jacob, regardless of whether consumption time was included.

Figure 3 displays the results for Gabby, presented identically to Jacob's results as shown in Figure 1. Gabby dissented to participate (i.e., began to cry, verbally stated that she needed a break, and declined to return from that break) after the second session of the second BL phase. Thus, only three phases (i.e., an ABA design in which A = BL and B = EXT) were conducted for this participant. The top panel of Figure 3 shows that Gabby's BL response rates averaged just under 2 RPM and exhibited a slight increasing trend. An EXT burst was observed in the transition from BL to EXT. Responding subsequently reached near-zero rates for three consecutive sessions, and then returned to BL levels for the two sessions conducted in the reversal to BL. As seen in the bottom panel, when consumption time was excluded, rates of responding during the first BL phase demonstrated an overall increasing trend with considerable variability. Furthermore, the EXT burst seen in the top panel was not observed when reinforcer consumption time was excluded. Consistent with the pattern of results observed for Jacob, rather a noteworthy decrease in magnitude of the initial rate of responding under EXT was observed in the transition between BL and EXT when reinforcer consumption time was excluded. In summary, for Gabby, an EXT burst was observed in the analysis that included reinforcer consumption time. Then once reinforcer consumption was excluded from the analysis, the magnitude of initial responding following the implementation of EXT was minimized.

Figure 4 displays the within-session data for Gabby and is displayed identically to the data displayed in Figure 2. As seen in the top panel, an EXT burst was observed in the transition from BL to EXT, at the within-session level of analysis, when reinforcer consumption time was included. This is expected given Gabby's whole-session data (Figure 1), as it is guaranteed that if an EXT burst is observed at the whole-session level, that it will be observed at the within-session level too. However, it is still interesting to note that when reinforcer consumption time was excluded from BL response rates (bottom panel), the apparent EXT burst was no longer observed.

Figure 5 displays the whole-session data for Martha, displayed identically to the data in Figures 1 and 3. As seen in the top panel, Martha's BL response rates averaged just under 2 RPM and exhibited little to no trend or variability. As assessed by the criteria specified by Lerman and Iwata (1995), an EXT burst was not observed in either of the two transitions from BL to EXT at the whole-session level. In the bottom panel, excluding reinforcer consumption time resulted in a somewhat different pattern of results relative to the top panel (i.e., a more substantial decrease in EXT responding relative to baseline when consumption time was excluded).

Figure 6 displays the within-session data for Martha. For Martha, an EXT burst was not revealed in the transition from BL to EXT by analyzing data at the within-session level of analysis. Though, as seen in the top panel, responding does increase from the last minute of BL to the first minute of EXT, this pattern of responding does not meet the EXT burst criteria set by Lerman and Iwata (1995) because this increase is not higher than all of the BL data points. The top panel displays BL rates that averaged just under 2 RPM and exhibited no trend. In the bottom panel, excluding consumption time resulted in a differing trend in responding in the first BL (i.e., no trend when consumption was included, a steep decreasing trend when consumption was excluded).

Discussion

The present study evaluated the effects of EXT under several different methods of data analysis. First, EXT effects were evaluated relative to one BL measure including consumption time in the analysis and one BL measure excluding reinforcer consumption time, based on procedures implemented by Wunderlich and Vollmer (2015). Second, EXT was evaluated at the whole- and within-session (last five minutes of BL and first ten minutes of EXT) level of analysis, based on procedures implemented by Katz and Lattal (2020). Under these analyses, we examined the sensitivity of EXT effects, including EXT bursts, to these data analysis manipulations. In sum, 5 BL to EXT transitions were evaluated for 3 participants, under two levels of analysis (i.e., whole-session and within-session). In total, 1 EXT burst was observed and this effect met EXT burst criteria when examined both at the whole-session level (i.e., under typical data analysis methods when consumption time was included) and at the within-session level. Therefore, 20% of BL to EXT transitions evaluated demonstrated an EXT burst. Though the number of data sets evaluated in this initial study was quite small, the number of EXT bursts observed was not substantially lower than number of bursts expected, based on Lerman and Iwata’s (1995) estimates of the prevalence of the effect (i.e., 24%, p. 93).

We hypothesized that, consistent with Katz and Lattal (2020), additional EXT bursts might be revealed by analyzing data at the within-session level that were not apparent at the whole-session level. However, the single EXT burst observed at the within-session level (i.e., Gabby) was also observed at the whole-session level of analysis. Therefore, unlike Katz and

Lattal, we did not observe more bursts when examining the data at the minute-by-minute level. The difference in results could be attributed to the current study using a smaller amount of data sets. Katz and Lattal analyzed between 8 and 16 EXT transitions per subject for 4 subjects, under 2 transition types ("overall" and "local," synonymous with the whole-session and within-session analyses in the present study, respectively), and under 2 measures per transition type (whole-session: overall vs local response rates; within-session: local vs local-end response rates). The present study evaluated 2 EXT transitions for 3 subjects, under 2 transition types (whole- vs within-session) and under 2 different measures (including vs excluding reinforcer consumption time). In total Katz and Lattal analyzed 74 EXT transitions in Experiment 2 of their 2020 study and observed EXT bursts in 34.8% of whole-session transitions (9 of 26), 48% of within-session transitions (12 of 25), and 87% of within-session cf. BL end transitions (20 of 23). Comparatively, the current study analyzed 20 EXT transitions and observed EXT bursts in 10% of whole-session transitions (1 of 10) and 10% of within-session transitions (1 of 10). Therefore, the difference in rates of additional EXT bursts observed in the within-session analysis that were not observed in the whole-session analysis may be due to the discrepancy in amount of EXT transitions conducted per participant (i.e., Katz & Lattal conducted 8-16 transitions per subject and the current study conducted 2 transitions per subject). Generally speaking, Katz and Lattal's design incorporated more repeated demonstrations of BL to EXT transitions which in turn could have increased the probability of observing additional EXT bursts in the within-session analysis that were not observed in the whole-session analysis. Future research should be conducted to examine the effects of repeated transitions from BL to EXT per participant on the prevalence of EXT bursts under whole- and within-session levels of analysis. Such investigations might evaluate at what number of repeated transitions from BL to EXT per participant do EXT bursts begin to appear in the within-session analysis and not in the whole-session analysis.

Furthermore, the present study differed from Katz and Lattal (2020) with respect to the behavioral history of the target response established during BL. In Katz and Lattal's study, the target response was reinforced on a variable schedule for five sessions, with each session ending once 60 reinforcers were delivered to the participant, for a total of 300 reinforcers delivered per BL phase per participant. In the current study, the target response was reinforced on a continuous schedule for 3-5 sessions, with each session ending after 5 min elapsed, equaling to approximately 21-25 reinforcers delivered per BL phase per participant. This is a substantially shorter behavioral history in which engaging in the target response resulted in delivery of reinforcement. Therefore, Katz and Lattal's procedure may have created a stronger association between the target response resulting in reinforcer presentation by creating a longer behavior history in BL phases.

We hypothesized that if EXT bursts were observed, the effect would be highly sensitive to the exclusion of reinforcer consumption time. When reinforcer consumption time was excluded from the BL rate of responding in datasets that an EXT burst was observed, all EXT bursts observed were minimized. Thus, although bursts were only observed in the datasets of 1 participant (i.e., Gabby), in both the whole-session level of analysis and the within-session level, EXT bursts were observed when reinforcer consumption time was included in BL rates of responding and were not observed when reinforcer consumption time was excluded.

As previously discussed, the rationale for the hypothesis that excluding reinforcer consumption from the rate calculation would mitigate any apparent EXT bursts was based on a hypothesized ceiling effect constraining participants' responding, similar to the hypothesized ceiling effect placed on responding when RIRD implementation time was included in response rate measures (Wunderlich & Vollmer, 2015). In the current study, behavior was reinforced for 30-s intervals. If responding was unlikely to occur during reinforcement intervals, the highest level of responding possible would be a rate of 2 RPM. Taken together, the present data illustrate a constraint on baseline rates of responding that functions similar to a ceiling effect was in effect: this is reflected in the top panel of Figures 1-6 in which every BL level never exceeds the value of 2 RPM when consumption time is included. Additionally, in only 1 session for 1 participant, did the target response ever occur during a reinforcement interval (4 occurrences in Session 7 for Jacob, Figure 1). In all other BL sessions in which reinforcement intervals occurred, 0 target responses occurred during the reinforcement interval. For the one participant that exhibited an EXT burst (i.e., Gabby, see Figure 3), the present results provide initial evidence that this effect was an artifact of efficient, discriminated responding that functioned similar to a ceiling effect when reinforcer consumption time was included. Thus, the present study may have implications for studies that utilize RPM as the primary dependent variable.

Additionally, when reinforcer consumption time was excluded from BL rates of responding, noteworthy differences in the pattern of responding were observed relative to the typical method of calculating response rate (i.e., including consumption time) across whole- and within-session methods of analysis, regardless of whether an EXT burst was observed. Re-analyzing the data to exclude consumption time resulted in a change in trend for 4 out of 10 BL patterns of responding. For 2 out of 10 datasets (Jacob, Fig. 1, BL 1; Martha, Fig 6, BL 1) the difference consisted of a change from no trend when reinforcer consumption time was included in BL rates of responding to an increasing trend when reinforcer consumption time was excluded from BL rates of responding. For the other 2 datasets (Jacob, Figure 2, BL 1; Gabby, Figure 4, BL 1) the difference consisted of a change from no trend to a decreasing trend when reinforcer consumption time was excluded from BL rates of responding. These effects suggest that excluding reinforcer consumption time from rates of responding results in a more detailed picture of the participants' pattern of responding during the session, which may reveal important changes in responding (e.g., increase in response efficiency, satiation, etc.) that are obscured when consumption time is included. Ultimately, response rate excluding reinforcer consumption time may be a more sensitive dependent variable. This is noteworthy because in both applied research and clinical treatment settings, decisions are made based off the visual analysis of patterns of responding. If treatment decisions were made based on the including reinforcer consumption time BL data, then clinicians would likely make the decision to move into EXT procedures based off the observation that there is no trend in BL. However, if treatment decisions were made based on the excluding reinforcer consumption time BL data, then clinicians might make an entirely different treatment decision for 40% of datasets based on the now increasing or decreasing trend.

The degree to which consumption time "skews" the response rate calculation is a function of reinforcement procedures, specifically, reinforcer duration. The effects of including consumption time in response rate are more of a concern if the functional reinforcer has temporal extent (e.g., leisure items and escape functions), relative to reinforcers that are relatively momentary (e.g., reprimands, praise, small edible items, etc.). In the present study, we evaluated a tangible reinforcer with a programmed reinforcer duration of 30 s, which we based on typical parameters used in the ABA literature. Further research is invited to examine if including reinforcer consumption time in response rate also affects the analysis of responding maintained by an escape from demand function. Moreover, future research should examine if excluding reinforcer consumption time from response rate changes the treatment decisions a clinician would make based on visual analysis, as well as the degree to which a clinician would deem a treatment as effective.

The present study successfully implemented a series of experimental conditions which may serve as a useful analogue to the acquisition and elimination of problem behavior in clinical settings. The primary goal of this procedure was to evaluate the EXT burst in a controlled environment, thereby producing results that were analogous to clinically relevant EXT bursts (e.g., bursts in problem behavior when EXT is implemented as a treatment component). However, the results of the present study may also have implications for understanding clinical situations including: using response rate to measure problem behavior while working on goals such as, accepting denied access to highly preferred items, prompting the use of moderately preferred items when highly preferred items are unavailable, and the use of electronics (i.e., videos or games) as a functional reinforcer during EXT procedures. The current study prioritized the internal validity of the results via highly controlled procedures (i.e., an arbitrary target response, a short behavioral history with the target response and reinforcer association, and a controlled setting without typical extraneous stimuli found in the natural environment), this was done so at the expense of the results' external validity. Therefore, all discussion of how the results of this study apply to the clinical treatment setting are largely extrapolatory. Future research is recommended to explore the applicability of the current study's procedures in appropriately assessing the treatment of problem behavior when maintained by a tangible function when in the clinical treatment setting.

EXT bursts have been described as a common extinction-related effect (e.g., Kazdin, 2012, p.360), but what if they are not as common as we thought? Specifically, the present findings provide some preliminary evidence for the suggestion that the standard practice of including reinforcer consumption time into rate of responding may result in effects that appear to be EXT bursts, that are in fact just an artifact of the data analysis method. Furthermore, the present data suggest that this practice could contribute to an underestimation of the effectiveness of EXT in eliminating responding. If ABA researchers and clinicians are overestimating the rate of EXT bursts and are not as likely an effect of EXT procedures as previously thought, then clinicians are likely being misled by the data that they are interpreting via visual analysis. This could result in less-than-ideal treatment changes being made. For example, clinicians may become concerned when hypothesized EXT bursts occur during treatment because the high rates of responding in EXT procedures can indicate the need to implement additional restrictive procedures. It is important to note that the excluding reinforcer consumption time data view does not erase the high rates of responding observed during EXT but rather demonstrates that those high rates may not be as problematic as previously thought, when considered relative to an appropriate BL. Indicating that further research should be conducted on what is truly the appropriate dependent variable and BL conditions to analyze EXT procedures under. Additionally, if EXT bursts simply do not exist at all (because they are entirely an artifact of the method of analysis), then focus may need to shift to the detection and mitigation of other EXT-related effects (e.g., EXT-induced aggression, behavioral variability, increased magnitude of responding, etc.) Ultimately, the way with which we measure the EXT burst has not been critically evaluated enough even though the implications of its presence in treatment are crucial.

References

Azrin, N. H., Hutchinson, R. R., & Hake, D. F. (1966). Extinction-induced aggression. *Journal of the Experimental Analysis of Behavior, 9,* 191-204. https://doi.org/10.1901/jeab.1966.9-191

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis (Second Edition)*. Pearson Education, Inc.

Goh, H. L. & Iwata, B. A. (1994). Behavioral persistence and variability during extinction of self-injury maintained by escape. *Journal of Applied Behavior Analysis*, *27*(1), 173-174. https://doi.org/10.1901/jaba.1994.27-173

Iversen, I. H. (1992). Skinner's early research: From reflexology to operant conditioning. *American Psychologist*, *47*(11), 1318–1328. https://doi.org/10.1037/0003-066X.47.11.1318

Katz, B. & Lattal, K. A. (2020). An experimental analysis of the extinction-induced response burst. *Journal of the Experimental Analysis of Behavior*, *114*(1), 24-46. https://doi-org.ezproxy.rollins.edu/10.1002/jeab.611

Kazdin, A. E. (2012). *Behavior Modification in Applied Settings (Seventh Edition).* Waveland Press, Inc.

Lattal, K. A., St. Peter, C., & Escobar, R. (2013). *Operant extinction: Elimination and generation of behavior*. In G. J. Madden (Ed.), APA handbook of behavior analysis (p. 77-107). American Psychological Association.

Lerman, D. C. & Iwata, B. A. (1995). Prevalence of the extinction burst and its attenuation during treatment. *Journal of Applied Behavior Analysis*, *28*(1), 93-94. https://doi.org/10.1901/jaba.1995.28-93

Lerman, D. C., Iwata, B. A., & Wallace, M. D. (1999). Side effects of extinction: Prevalence of bursting and aggression during the treatment of self-injurious behavior. *Journal of Applied Behavior Analysis, 32*(1), 1-8. https://doi.org/10.1901/jaba.1999.32-1

Magee, S.K., & Ellis, J. (2000). Extinction effects during the assessment of multiple problem behaviors. *Journal of Applied Behavior Analysis, 33*(3), 313-316. https://doi.org/10.1901/jaba.2000.33-313

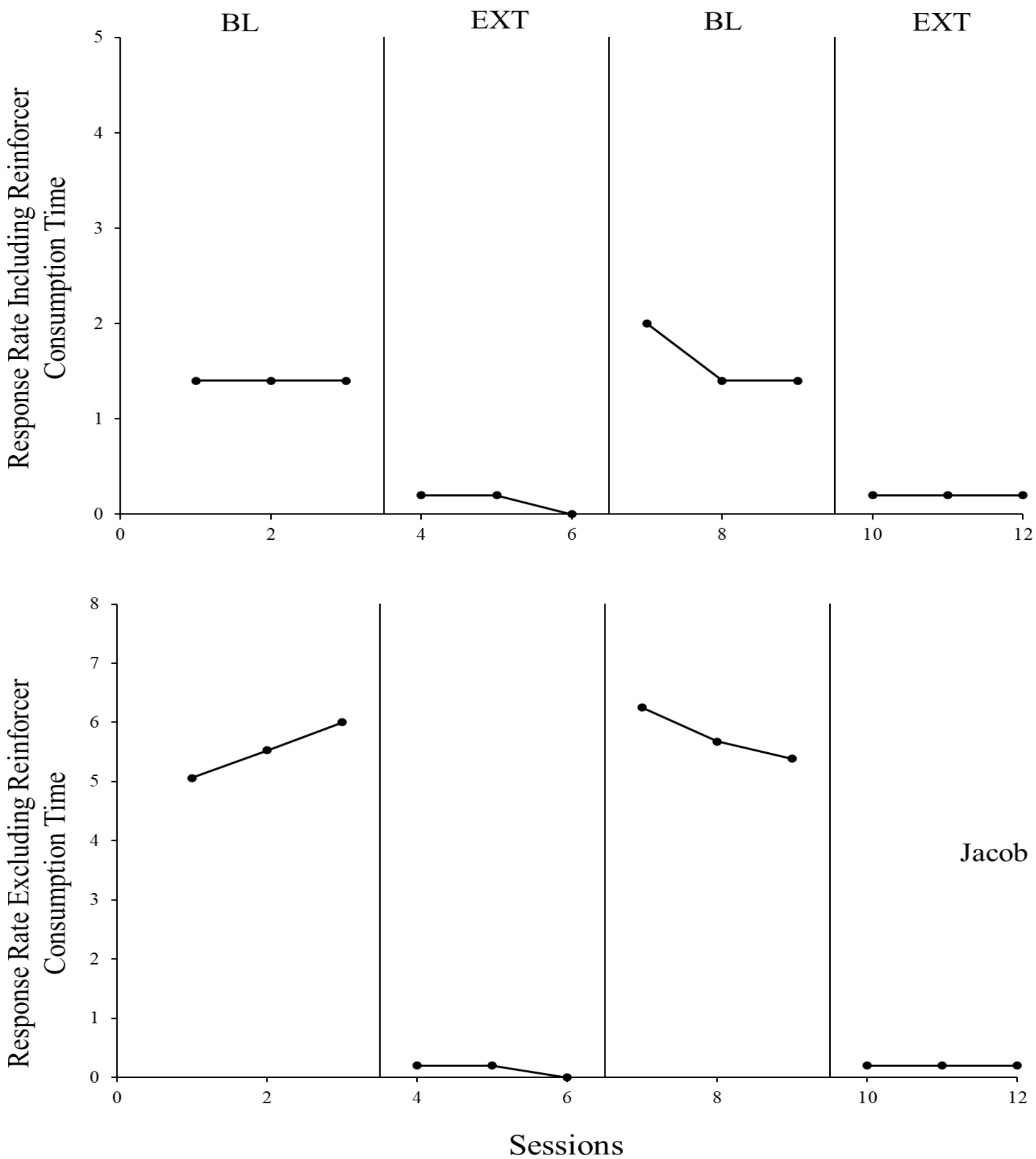
Repp, A. C., Felce, D., & Barton, L. E. (1988). Basing the treatment of stereotypic and self-injurious behaviors on hypotheses of their causes. *Journal of Applied Behavior Analysis, 21*(3), 281-289. https://doi.org/10.1901/jaba.1988.21-281

Wunderlich, K. L. & Vollmer, T. R. (2015), Data analysis of response interruption and redirection as a treatment for vocal stereotypy. *Journal of Applied Behavior Analysis, 48*(4), 749-764. https://doi.org/10.1002/jaba.227

Zarcone, J. R., Iwata, B. A., Vollmer, T. R., Jagtiani, S., Smith, R. G., & Mazaleski, J. L. (1993). Extinction of self-injurious escape behavior with and without instructional fading. *Journal of Applied Behavior Analysis*, *26*(3), 353-360. https://doi.org/10.1901/jaba.1993.26-353

Figure 1

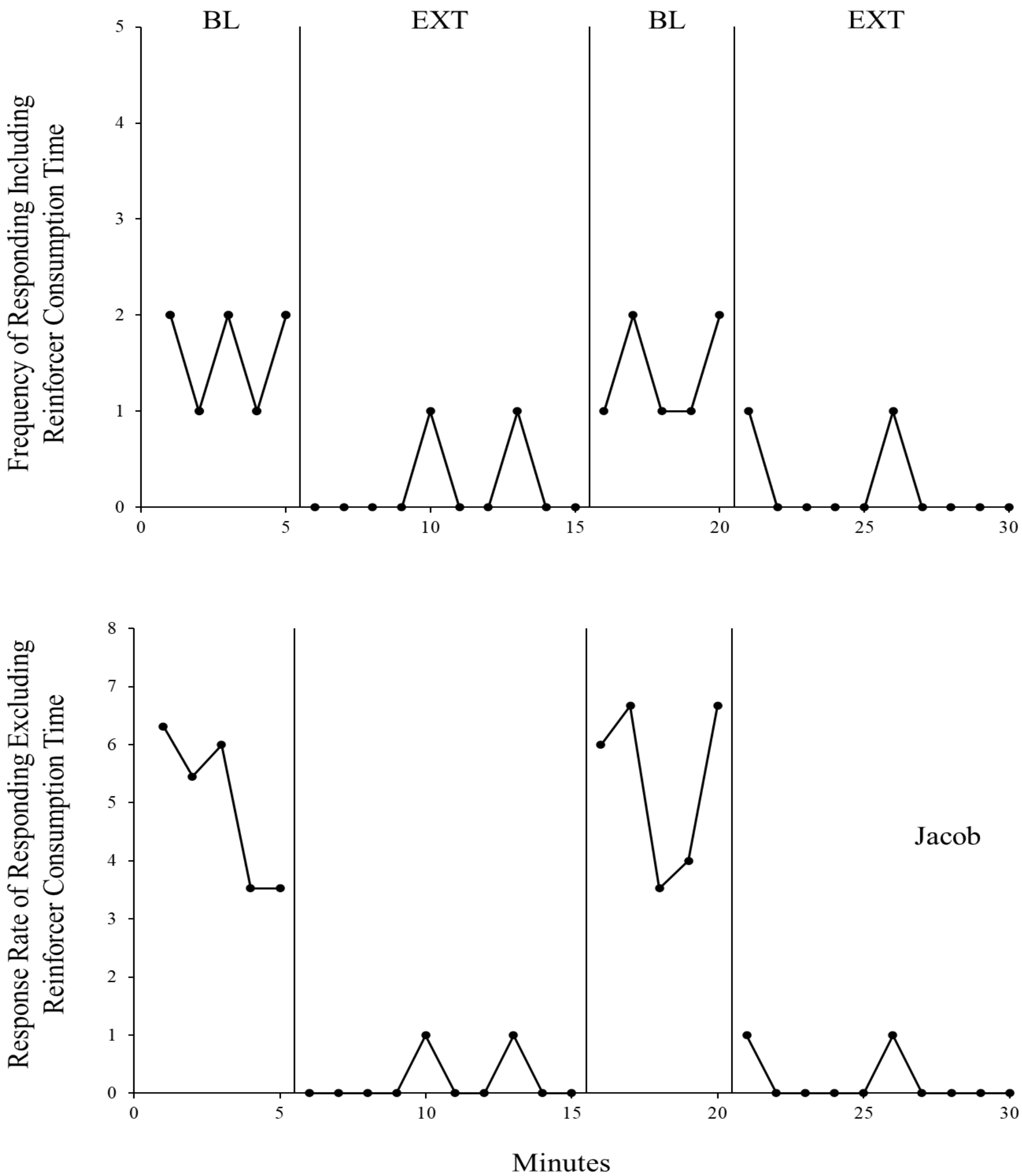
Whole-Session Analysis of Extinction Transitions for Jacob



Note. The whole-session analysis of responses per minute in which baseline includes reinforcer consumption time (top panel) and with baseline excludes reinforcer consumption time (bottom panel) for Jacob. BL and EXT indicate Baseline and Extinction phases, respectively. Please note that the scale for the y-axes differ per panel.

Figure 2

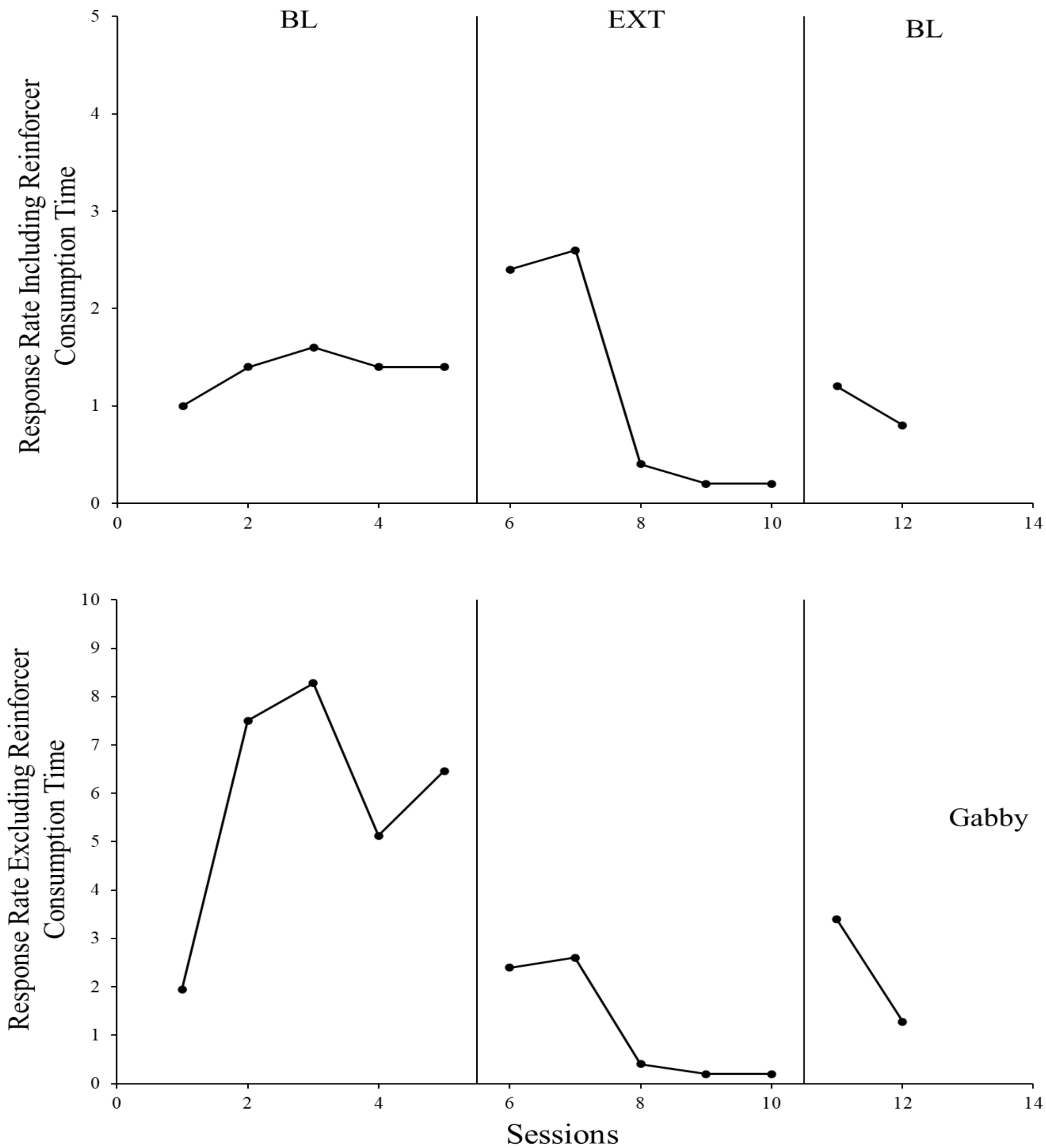
Within-Session Analysis of Extinction Transitions for Jacob



Note. The within-session analysis of frequency of responding emitted for the last five min of each baseline phase and the first ten min of each extinction phase for Jacob. The top panel displays the frequency of responding in which reinforcer consumption time is included in baseline measures and the bottom panel displays the frequency of responding in which reinforcer consumption time is excluded from baseline measures. Last 5 BL and First 10 EXT indicate Baseline and Extinction phases, respectively. Please note that the scale for the y-axes differ per panel.

Figure 3

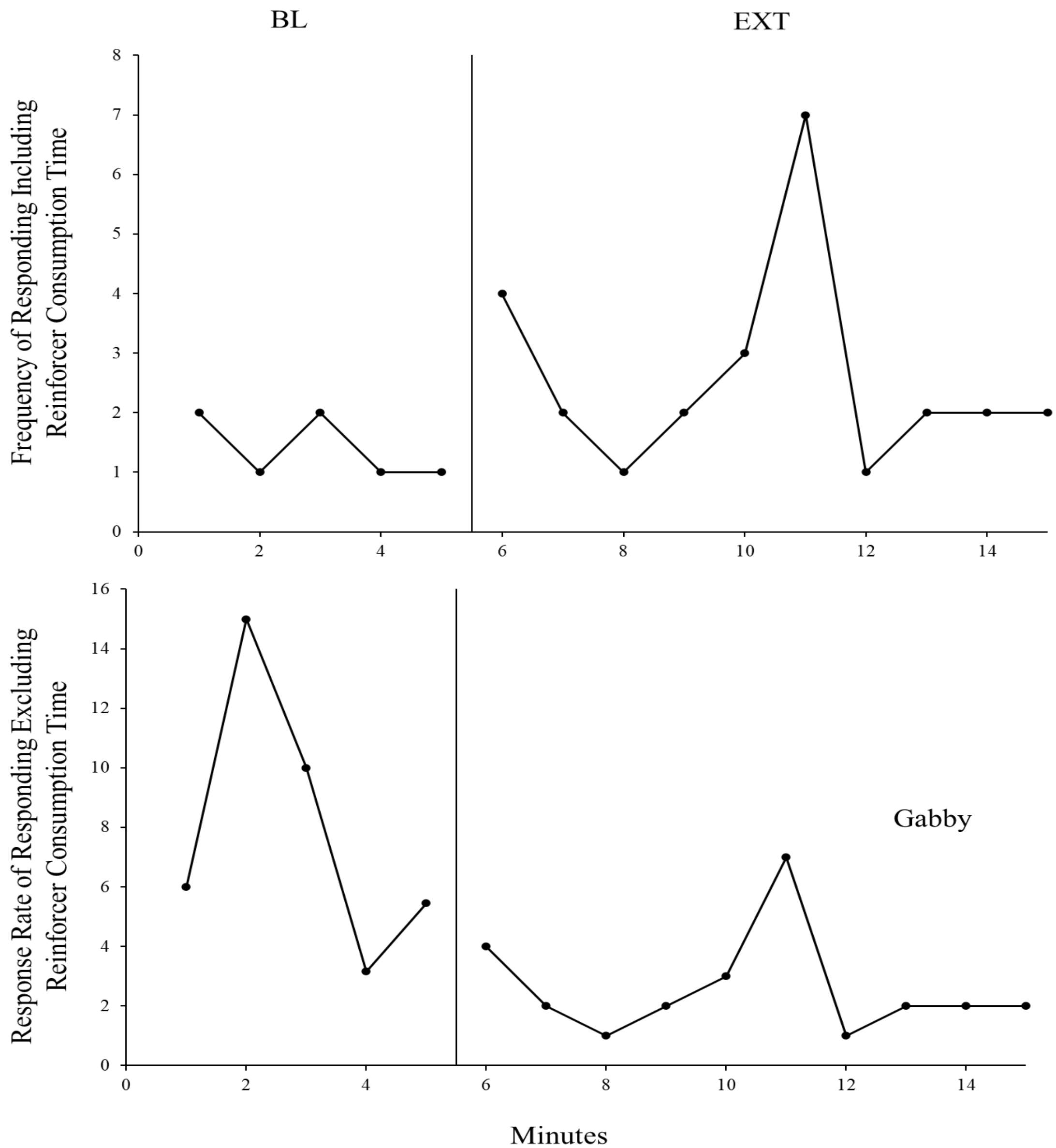
Whole-Session Analysis of Extinction Transitions for Gabby



Note. The whole-session analysis of responses per minute in which baseline includes reinforcer consumption time (top panel) and with baseline excludes reinforcer consumption time (bottom panel) for Gabby. BL and EXT indicate Baseline and Extinction phases, respectively. Please note that the scale for the y-axes differ per panel.

Figure 4

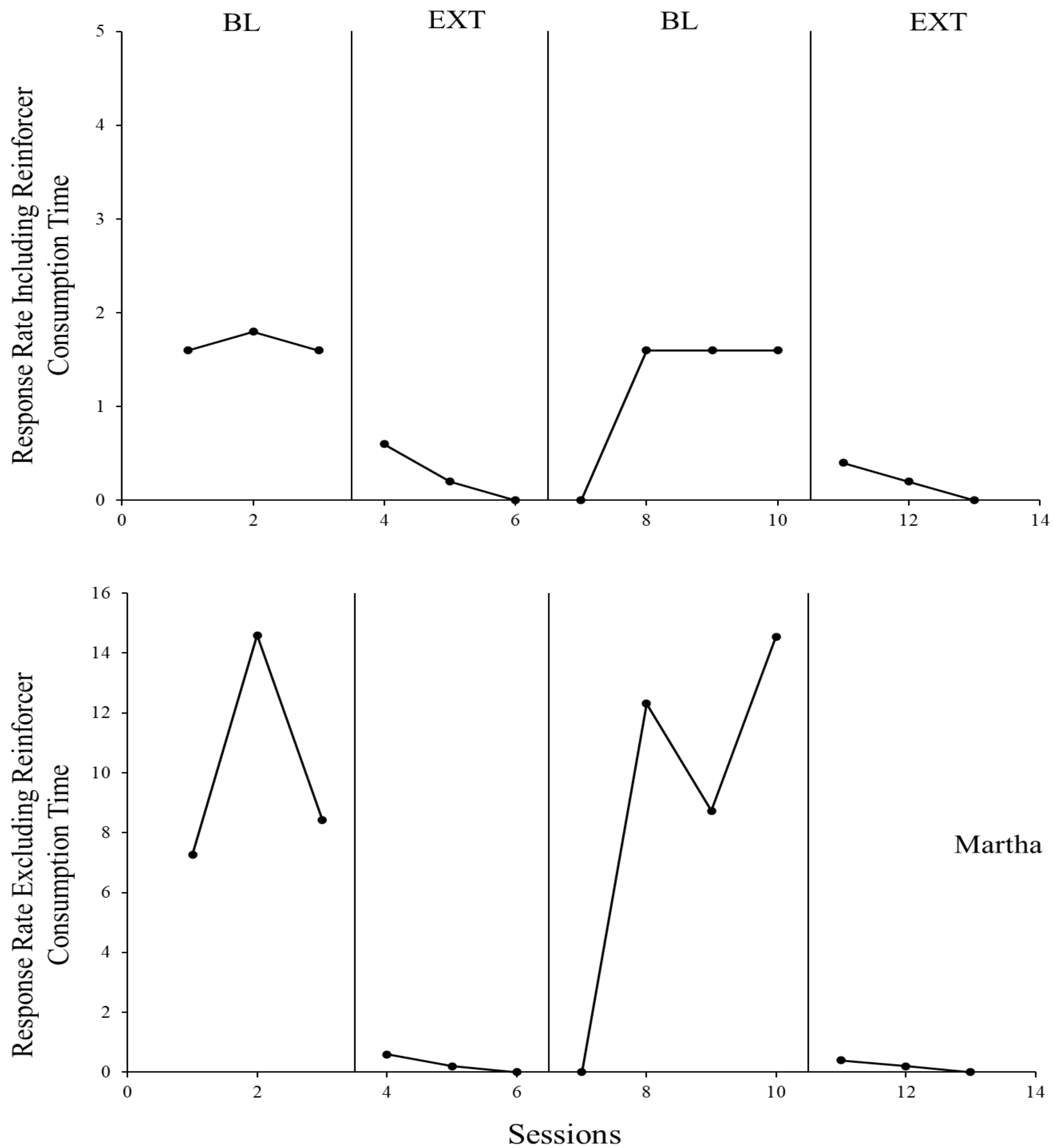
Within-Session Analysis of Extinction Transitions for Gabby



Note. The within-session analysis of frequency of responding emitted for the last five min of each baseline phase and the first ten min of each extinction phase for Gabby. The top panel displays the frequency of responding in which reinforcer consumption time is included in baseline measures and the bottom panel displays the frequency of responding in which reinforcer consumption time is excluded from baseline measures. Last 5 BL and First 10 EXT indicate Baseline and Extinction phases, respectively. Please note that the scale for the y-axes differ per panel.

Figure 5

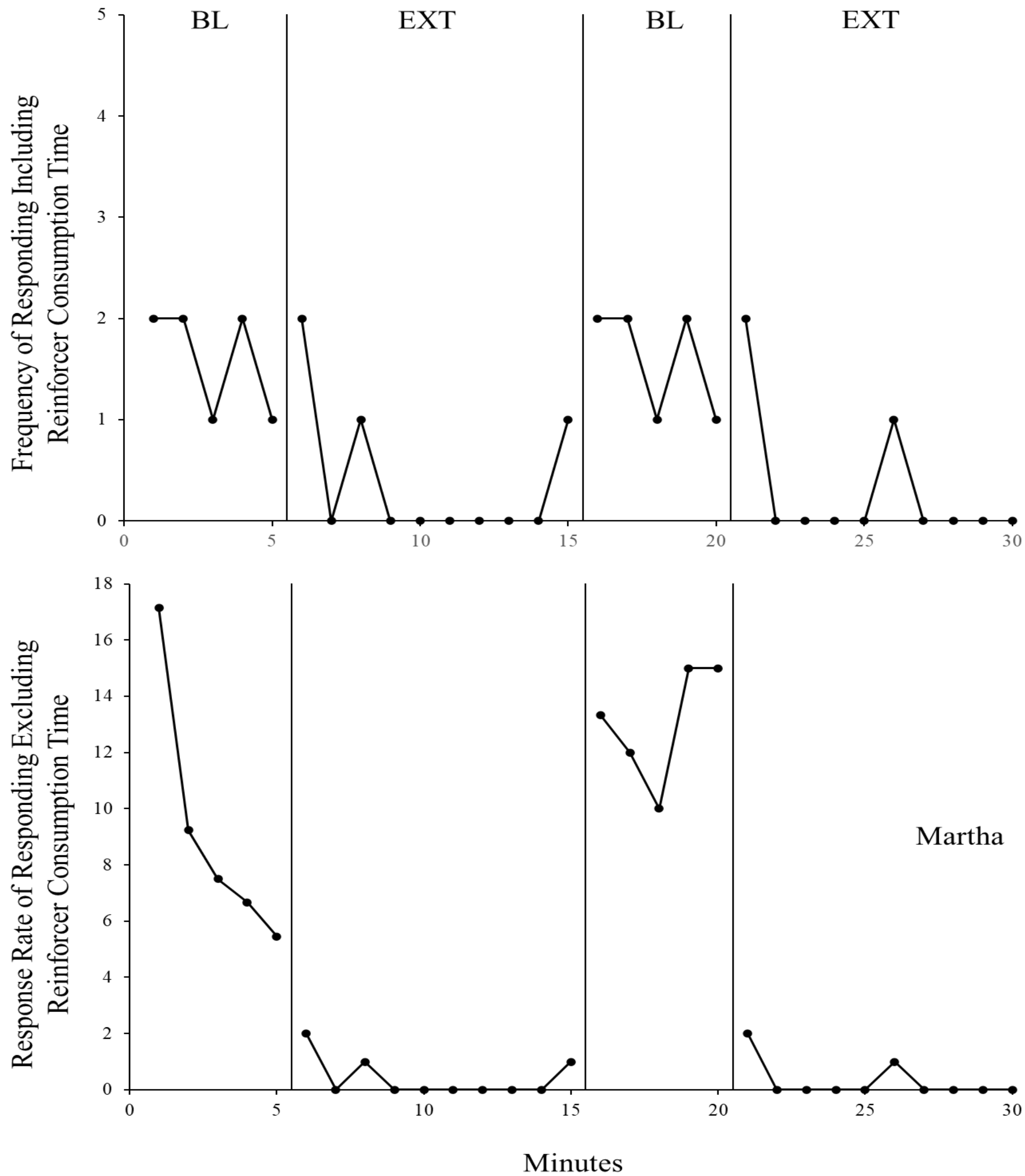
Whole-Session Analysis of Extinction Transitions for Martha



Note. The whole-session analysis of responses per minute in which baseline includes reinforcer consumption time (top panel) and with baseline excludes reinforcer consumption time (bottom panel) for Martha. BL and EXT indicate Baseline and Extinction phases, respectively. Please note that the scale for the y-axes differ per panel.

Figure 6

Within-Session Analysis of Extinction Transitions for Martha



Note. The within-session analysis of frequency of responding emitted for the last five min of each baseline phase and the first ten min of each extinction phase for Martha. The top panel displays the frequency of responding in which reinforcer consumption time is included in baseline measures and the bottom panel displays the frequency of responding in which reinforcer consumption time is excluded from baseline measures. Last 5 BL and First 10 EXT indicate Baseline and Extinction phases, respectively. Please note that the scale for the y-axes differ per panel.