Fall 12-15-2023

Feasibility of a Laboratory Model for Evaluating the Use of an Alternative Response Option in Mitigating Caregiver Treatment Integrity Errors

Jessie K. Weber
University of Nebraska Medical Center

FEASIBILITY OF A LABORATORY MODEL FOR EVALUATING THE USE OF AN ALTERNATIVE RESPONSE OPTION IN MITIGATING CAREGIVER TREATMENT INTEGRITY ERRORS

by

Jessie K. Weber

A DISSERTATION

Presented to the Faculty of the University of Nebraska Medical Center in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Medical Sciences Interdepartmental Area Graduate Program

(Applied Behavior Analysis)

Under the Supervision of Professor Tara A. Fahmie

University of Nebraska Medical Center
Omaha, Nebraska

December 2023

Supervisory Committee

Brenda Bassingthwaite, Ph.D	William Higgins, Ph.D.
Jamie Gehringer, Ph.D.	Catalina N. Rey, Ph.D.

ACKNOWLEDGMENTS

Completing a PhD is not a solitary endeavor, and the support I have received along the way has been immeasurable. I want to express my deep gratitude to Dr. Fahmie whose encouragement and guidance were instrumental throughout my academic journey. Her support bolstered my confidence and reinforced my commitment to my scholarly pursuits. The impact of her mentorship extends far beyond my time as a doctoral candidate and will continue to resonate in my professional endeavors. Additionally, I could not have undertaken this project without the continued advisement from my dissertation committee. Thank you to Drs. Brenda Bassingthewaite, Billy Higgins, Jamie Gehringer, and Catalina Rey for providing your knowledge and expertise.

None of this would have been possible if it weren't for my early mentors in the field. Dr. Billie Retzlaff and Dr. Dan Mitteer welcomed me into the field of behavior analysis with open arms. Dr. Retzlaff and Dr. Mitteer not only guided me academically but also provided unwavering encouragement, sound advice, and steadfast advocacy. I am inspired by their, and Dr. Fahmie's commitment to nurturing talent and strive to emulate their support for my own mentees.

To my classmates and cohort members, Dr. Jordan DeBrine, Brinea Charles, Dr. Maya Fallon, Kyle Dawson, and Isaac Melanson, I am forever grateful for each of you for providing unwavering support throughout this journey. It's a privilege shared by few to have completed a PhD, and I couldn't have imagined embarking on this adventure with any other group of people.

To my husband Cody, my dog Howie, my parents Kaye and Darrell, my family, and friends, you were my pillars of strength, wisdom, and encouragement that supported me through the challenges and triumphs of earning this PhD. I am profoundly thankful for all the people listed above's influence on my life and remain committed to providing the support and guidance I received.

FEASIBILITY OF A LABORATORY MODEL FOR EVALUATING THE USE OF AN ALTERNATIVE RESPONSE OPTION IN MITIGATING CAREGIVER TREATMENT INTEGRITY ERRORS

Jessie K. Weber, Ph.D.

University of Nebraska, 2023

Supervisor: Tara A. Fahmie, Ph.D.

The current study assessed the feasibility of a laboratory model designed to evaluate options for mitigating common caregiver errors in the implementation of destructive behavior treatment. We developed a computer-based analogue to caregiver implementation of extinction for attention-maintained destructive behavior; the model included contingencies on both caregiver errors of commission and adherence to the treatment protocol. We also conducted a preliminary investigation of the effects of participant access to alternative activities as a potential strategy for mitigating integrity errors. Participants included 14 MTurk workers or staff in a severe behavior clinic, and individual response patterns revealed distinct sensitivities to the positive and negative reinforcers programmed in the model. These data support further use of the laboratory model to explore strategies to improve the implementation of behavioral treatment.

Keywords: Treatment integrity, functional communication training, extinction, competing response

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ASD	autism spectrum disorder
DRA	differential reinforcement of alternative behavior
FCR	functional communication response
FCT	functional communication training
ID	intellectual disorder
Mturk	Amazon mechanical turk
N/A	not available

INTRODUCTION

Destructive behavior (e.g., self-injury, aggression, property destruction) is common among individuals with autism spectrum disorders (ASD) and intellectual disabilities (IDs; Emerson et al., 2001; Holden & Gitlesen, 2006; Lowe et al., 2007). Individuals who engage in destructive behavior and their caregivers often have limited access to school and community resources due to their destructive behavior (Carr & Carlson, 1993). These families can be referred to behavior analysts to identify the environmental causes of destructive behavior and evaluate treatment approaches that aid in the success of the individual in the home and community settings. By identifying reinforcers for destructive behavior, practitioners can develop and implement individualized function-based treatments that suppress destructive behavior and increase prosocial skills.

Relevant to the current study, destructive behavior is reported to be reinforced by attention in 14.5% of published cases of destructive behavior having a single function and in 55.6% of cases having multiple functions (Melanson & Fahmie, 2023). Attention as a function of destructive behavior is commonly treated using functional communication training (FCT), which involves reinforcing an alternative functional communication response (i.e., FCR; e.g., "Play, please"). The goal of FCT is to reduce rates of destructive behavior, increase rates of FCRs, thin the schedule of reinforcement, and transfer treatment effects to relevant caregivers and settings. FCT combined with extinction (i.e., ignoring the destructive behavior, when its function is attention) has been shown to be more effective than FCT without extinction when schedules of reinforcement are equal (Athens & Vollmer 2010, Hagopian et al., 1998, Tiger et al., 2008, Trump et al., 2020). In a review of FCT research spanning from 1985 to 2019, Ghaemmaghami et al. (2021) found that a majority of FCT studies relied on applications under dense schedules of reinforcement and conducted in highly controlled settings. Under the conditions described, published literature has shown positive effects of FCT with extinction in suppressing destructive

behavior. That is, FCT has been shown to reduce destructive behavior by 80% or more in 90% of applications (Ghaemmaghami et al.).

Despite the efficacy of FCT with extinction in highly controlled settings and under dense schedules of reinforcement, very little is known about the efficacy of FCT with extinction in natural settings. Additionally, initial effects of FCT with extinction rarely maintain and complete elimination of destructive behavior is reported to occur in only 53% of cases reviewed by Ghaemmaghami et al. (2021). In other words, the efficacy of FCT is high but its effectiveness is less promising (Ghaemmaghami et al.). Common procedures used to prevent treatment failures include exposing the child with destructive behavior to extended periods of extinction (Fisher et al., 2018), using discriminative stimuli to signal the availability of reinforcement (e.g., red and green card; Saini et al., 2016), and implementing generalization training (e.g., home and school; Tiger et al., 2008). Additionally, researchers have recognized the critical role of treatment integrity when transitioning a child from a highly controlled setting with trained staff to the home, school, and community settings with caregivers whose training is likely to vary widely (e.g., St. Peter Pipkin et al., 2010).

Treatment integrity is defined as the correct or incorrect implementation of prescribed treatment components. Regarding FCT, the key components of treatment integrity include reinforcing FCRs and ignoring destructive behavior. St. Peter Pipkin et al. (2010) assessed the impact of treatment integrity errors during FCT in both a human operant and an applied experiment. In their study, the independent variable was treatment integrity errors, defined as *errors of commission* (providing reinforcement for destructive behavior) and *errors of omission* (withholding reinforcement following an FCR). Both types of errors were systematically manipulated to evaluate the impact on the target behavior (i.e., key pressing in the human operant experiment and destructive behavior in the follow up applied experiments). Participants were undergraduate students (human operant experiment) and two individuals with disabilities (applied

experiments). The authors assessed the effects commission and omission errors in isolation and in varied combinations across conditions (e.g., baseline with 100% integrity to DRA with 100% integrity versus baseline with 100% integrity to DRA with 50% integrity). Results showed that child alternative behavior (akin to the FCR) reduced when omission errors occurred in isolation with little to no increase in child target behavior (akin to destructive behavior). However, target behavior was more likely to increase when commission errors were tested in isolation and integrity diminished below 50%. Combining errors (i.e., integrity at or below approximately 50% with both commission and omission errors occurring), magnified these detrimental effects. While a combination of errors may be likely to occur in the natural setting, there were clear differences in the detrimental impact of the type of error occurring on treatment outcomes in this study. Specifically, when integrity was reduced and commission errors occurred (i.e., reinforced a target behavior), participants were more likely to engage in target behaviors. One way to offset commission errors is to train caregivers to withhold reinforcers following target behavior more effectively.

Many researchers have evaluated training packages (e.g., behavior skills training; Brookman-Frazee et al., 2009; Moore & Amado, 2021) to improve caregiver treatment integrity during FCT given that destructive behavior is likely to return following errors in treatment implementation (Gregori et al., 2021). However, researchers have found implementation to decrease as soon as 10 days following training in the absence of ongoing implementation support (Sanetti & Kratochwill, 2009). One potential reason for this decrement in performance is that training has often involved teaching caregivers to engage in a non-behavior (i.e., ignoring) following instances of destructive behavior. Ignoring may be difficult to teach and reinforce because the natural reinforcer for attending to destructive behavior (i.e., temporary suppression of destructive behavior) remains available in the typical environment. For example, if a child engages in self-hitting maintained by attention, and their caregiver responds by providing a

reprimand (i.e., attention), the child is likely to temporarily stop hitting themselves following the attention. Thus, the child's hitting is positively reinforced by attention and the adult's delivery of attention (i.e., commission error) is negatively reinforced by brief cessation of the hitting (e.g., Miller et al., 2010). Moreover, the caregiver's accurate response of ignoring destructive behavior may be punished by a temporary increase in destructive behavior (i.e., extinction burst). The function of caregiver behavior has been minimally discussed in FCT research on caregiver treatment integrity (Stocco & Thompson, 2015).

We hypothesized that providing the caregiver with an alternative response that competes with the potential commission error of attention delivery for destructive behavior may reduce commission errors. This approach most resembles a differential reinforcement of alternative behavior (DRA) treatment for the caregiver, in that a designated response is strengthened; in this case, the designated response could be any activity that competes with the delivery of attention for destructive behavior, such as completing a checklist or covertly counting down. There are several reasons why this approach may be more desirable than simply teaching caregivers to ignore destructive behavior. First, the concept of "ignoring" has recently been scrutinized by the behavioral community as not being an accurate technological description of common actions taken by caregivers (Lloveras et al., 2023). Moreover, the use of DRA has been shown to be preferred by clients receiving treatment (Luczynski and Hanley, 2009) and by caregivers delivering treatment (Gabor et al., 2016). By extension, it may be deemed more socially valid for a caregiver to be taught to emit a response when destructive behavior occurs rather than ignoring it (e.g., Tarbox et al., 2023). Although the negative reinforcer (temporary escape from destructive behavior) for delivery of attention following destructive behavior may remain intact, a qualitatively different reinforcer (e.g., praise, tangible incentives) for a distinct response may effectively compete during the initial stages of training (Vollmer et al., 2020). Finally, destructive behavior research has shown that providing an alternative activity, generally, mitigates delays to

reinforcement (e.g., Fisher et al., 2000, Drifke et al., 2023). In caregiver implementation of FCT, delays to reinforcement of caregiver behavior may come in the form of a gradual reduction in destructive behavior, perhaps after a burst.

Thus, it seems reasonable to hypothesize that explicitly teaching a competing response (as opposed to a non-response or ignoring) when child destructive behavior occurs may be a more successful training strategy than attempting to strengthen any behavior other than providing attention for destructive behavior. The general approach of strengthening alternative behavior is considered evidence-based practice for individuals with destructive behavior, it has yet to be applied to caregiver behavior. Two challenges to conducting studies on this topic are (a) the inherent safety risk in allowing caregivers to practice errors during treatment implementation, and (b) keeping opportunities for error steady across conditions by controlling destructive behavior rates. A laboratory model of destructive behavior, therefore, is warranted.

The purpose of the current study was to assess the feasibility of a laboratory model designed to approximate conditions of caregiver training on destructive behavior treatment. Our paradigm modeled critical features of both positive and negative reinforcement processes, as well as extinction processes, that naturally occur or are programmed during caregiver training and implementation of treatment. We evaluate the feasibility of the model using a preliminary study on the effects of reinforcing a competing response option on the occurrence of errors of commission. Specifically, this study evaluated participant inhibition of responding (analogous to extinction or "planned ignoring" of destructive behavior) and the effect of participant engagement with an alternative activity on their treatment adherence. Thus, this study built on existing translational and applied research on treatment integrity (Mitteer et al., 2018; Wathen & Podlesnik, 2018) and basic research on response inhibition (Romano & St. Peter, 2017)

CHAPTER 1: METHOD

Participants

Sixteen participants were recruited from either Amazon Mechanical Turk (Mturk), a crowdsourcing platform that virtually distributes research projects (*n* = 9) or were staff from a university-based clinic specializing in the treatment of severe destructive behavior (*n* = 7). Two participants met our exclusion criterion, one prior to the start of the study and one following data analysis (see below). One participant did not complete the demographic and preference portion of the study, so their demographic and preference data are listed as not available (N/A) but all other data were retained. For the 13 participants who provided demographic information, all identified as hite with age ranges as follows: 19-30 years old (*n* = 8), 31-40 years old (*n* = 2), 41-50 years old (*n* = 2), and 50+ year old (*n* = 1). Eleven of the thirteen participants reported being married and nine reported having children. Participants with children reported child ages between 0-2 years old (*n* = 4), 3-5 years old (*n* = 4), and 6-10 years old (*n* = 1). See Table 1 for a full list of participant demographics including gender, race/ethnicity, marital status, children, and child age range.

Table 1: Participant Demographics

Variables	*N*	Percent of Sample
Gender		
Female	9	64.3%
Male	4	28.6%
Not listed (Fill in)	-	-
Prefer not to say or N/A	1	7.1
Age		
19-30	8	57.1%
31-40	2	14.3%
41-50	2	14.3%
50+	1	7.1%
Prefer not to say or N/A	1	7.1%
Race/Ethnicity		
White	13	92.9%
African American	-	-
Latino or Hispanic	-	-
Asian	-	-
Native American	-	-
Native Hawaiian or Pacific Islander	-	-
Two or more	-	-
Other/Unknown	-	-
Prefer not to say or N/A	-	7.1%
Marital Status		
Yes	11	78.6%
No	2	14.3%
Prefer not to say or N/A	1	7.1%
Children		
Yes	9	64.3%
No	4	28.6%
Prefer not to say of N/A	1	7.1%
Child Age Range[a]		
0-2 years old	4	44.4%
3-5 years old	4	44.4%
6-10 years old	1	11.1%
10+ years old	-	-
Prefer not to say	-	-

Note. N/A= Not available, [a] Percentage is out of the number of cases who reported having children. One participant did not complete the demographic questionnaire; their data are included in the remainder of the analyses.

All 14 participants consented to participate (see Appendix A), completed a consent quiz (see Appendix B), and passed a visual and audio check (see below) prior to participation. At the end of the study, all participants were invited to complete a brief demographic questionnaire (e.g., age, race, gender, number of children) and an intervention usability and preference survey (see Appendix C). All MTurk participants received a standard rate of compensation and were told that accumulation of points (see below) resulted in bonus earnings. We used MTurk's qualifications feature so only MTurk workers with >99% HIT approval rates could participate. We collected MTurk worker identification numbers to exclude workers trying to complete the experiment more than once. Participants were informed that they could receive up to $5 for participating based on their performance during the experimental task (Smith & Greer, 2022). However, at the end of the study, all MTurk participants were given the highest level of compensation regardless of their point accrual (see Debriefing Form in Appendix D). University staff members' participation was voluntary, and those participants were not compensated for their time due to departmental regulations but were relieved of clinical duties while they participated.

Computer Program and General Experimental Stimuli

The study was completed remotely on a computer program running all experimental tasks and automatically collecting data on all dependent measures on a virtual platform (Pavlovia) that was approved by the security team at the primary researcher's university. Participants were required to have access to a computer, keyboard, and headphones.

Throughout all phases of the study, the computer program ran an ambient task that was designed to be analogous to a caregiver's engagement in a variety of everyday routines (e.g., attending to spouse, driving, talking on the phone) outside of the management of destructive behavior. The ambient task involved the presentation of several visual and auditory matching trials (see Figure 1 for a visual representation; the ratio of visual to auditory presentation was 50/50). During the ambient task, participants were presented with a screen depicting a 3x3 square

grid, with a "+" in the middle square. This screen appeared between matching trial and was presented for 2,000 milliseconds (ms). Matching trials were displayed on this same screen and included the visual depiction of a target letter or a dash in one of the grid spots adjacent to the "+" square. When a dash was present, a target letter sounded through the participants' headphones. Letters (all except k and o for both visual and auditory trials) were rotated across trials and across positions on the grid using a randomization algorithm in the computer program. The visual or auditory letter were present for up to for 500 ms.

The ambient task required participants to press the key on their keyboard that corresponded to the letter they saw or heard. For example, if an "f" appeared in the lower right square of the grid, the participant pressed the letter "f" on their individual keyboard before the next matching trial appeared. Throughout all phases of the study, participants received one point for each correct matching response; a score box was visible at all times in the upper right-hand corner to promote engagement with the ambient task. If a participant scored two standard deviations below their own baseline on all data points in any phase subsequent to baseline, they were excluded from the study; one participant was excluded based on this criterion. This exclusion criterion was in place to ensure participants were attending to stimulus changes in the computer program throughout the study.

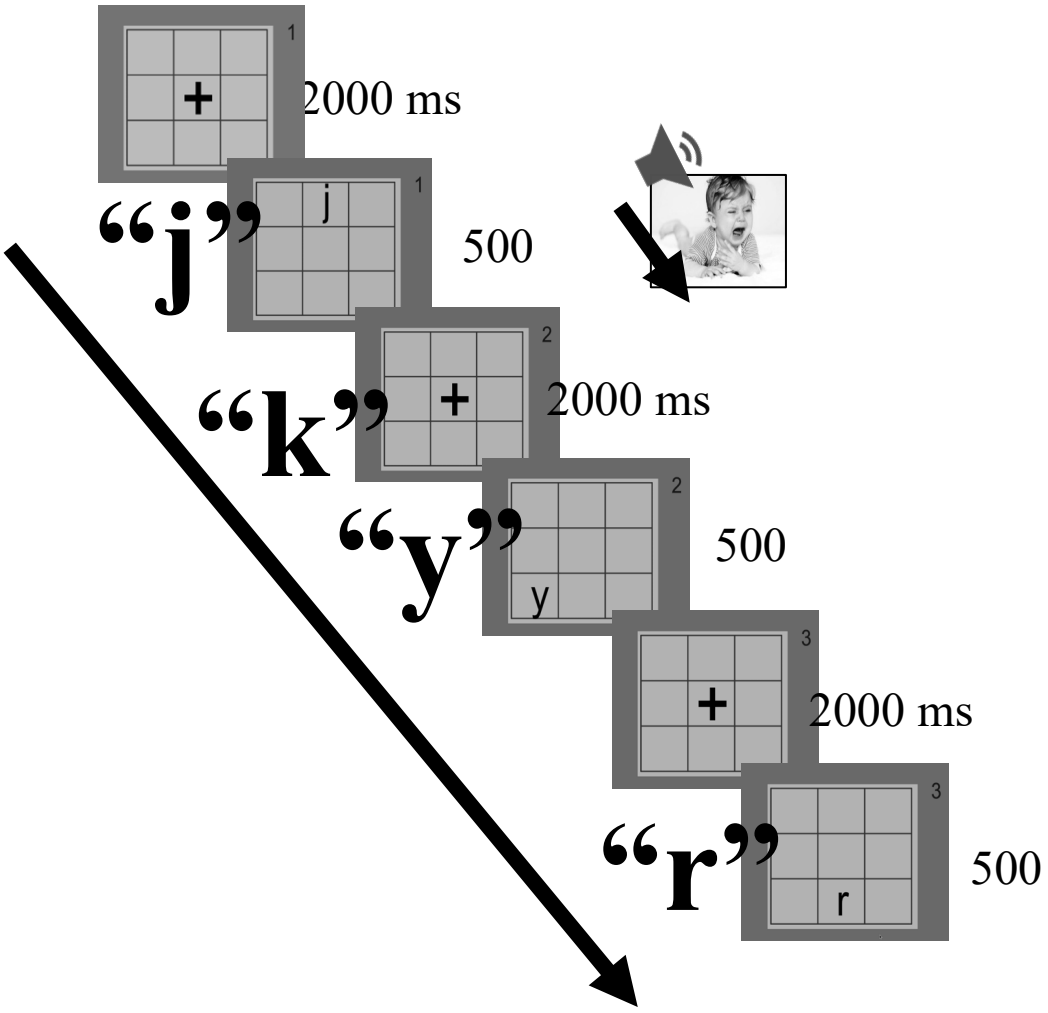

Figure 1: Visual Depiction of Ambient Task. Ms = milliseconds. Arrows represent the passage of time across the course of a session. Individual images within the grid depict matching trials (ambient task), with time displayed appears to the right of the image. The letters in quotation marks indicate participant responding on their computer keyboard. The image of a child crying with a speaker indicates the sound of a child crying occurring through participant's headphones.

Additionally, throughout all phases of the study, a 5-s audio clip of child crying at 80 decibels (dB) (Nuemann & Waters, 2006) intermittently sounded through the headphones (see below), analogous to the presence of destructive behavior in applied clinical research (e.g., Bowman et al., 2013). The specific audio clip was selected based on data collected prior to the study. Seven audio clips were sent to two departments within in the primary researcher's institution who are likely to encounter some level of destructive behavior during their day-to-day work. Three 5-s audio clips, taken from a longer file of a child crying, were compared to other potentially pleasant, neutral, and aversive audio clips (e.g., windy trees, white noise, chewing). Staff were presented with one audio clip at a time and were first asked whether the audio clip was aversive or not. If they answered yes, they were then asked to rate the clip from 0-100 with 0 being "an aversive sound that you can easily ignore" and 100 being "an aversive sound that is impossible for you to ignore." The audio files of the child crying were ranked (i.e., indicated as aversive on the yes/no question) by 12 of 18 staff, and were then scored 3rd, 5th, and 6th aversive out of the seven clips (chewing was scored 1st). The mean scores for each crying clip ranged from 54-58 on the 100-point scale. The highest-rated crying clip was selected for use in the study because crying may be encountered in the treatment of destructive behavior, and because we wanted to use a clip that was relatively difficult to ignore.

Prior to the start of the study, participants completed visual and audio checks. Visual checks were used to confirm that participants could distinguish between three colors (blue, yellow, and green) that were randomly assigned to a condition of the experiment. Participants were asked to match these three colors to their name by dragging the picture to the written description (i.e., match to sample). The audio check was used to confirm that participants could hear the auditory stimuli (letters, crying clip) used throughout the study. Additionally, participants were asked to rate the aversiveness of the crying clip using the same 100-pt scale described previously. Only those individuals who scored 100% on the visual check and who rated

the aversiveness of the audio clip above 36 points (falling above the 1st quartile of responses from pilot data) were included in our data analysis; one participant was excluded based on this criterion. For participants included in the study, the average rating of the audio clip was a 66.2 (range 50-82) on the 100-pt scale.

Program Validation

The authors validated the experimental software using procedures outlined by Smith and Greer (2020). See Appendix E and F for the Basic Functionality Checklist and variables evaluated when testing the basic functionality. The current study was not initiated until the software implemented each parameter included in the Basic Functionality Checklist with 100% accuracy.

Experimental Procedures

Participants independently accessed the software described and validated by the authors through Amazon's MTurk platform (MTurk participants) or through a link sent via email by a departmental office associate (staff participants). Following the informed consent process, participants received a general description of the study:

> "We are conducting this study to learn how adults respond to the presentation of the sound of a baby crying while engaging in other tasks. We will provide specific instructions that will give you options to stop or reduce the number of times you hear the sound while playing a game. Please read the instructions carefully prior to the start of each phase."

Baseline Phase

The baseline phase lasted six minutes and was designed to mimic what a caregiver might experience prior to their child accessing behavioral services. That is, approximately every 8s throughout the phase, participants heard the aversive sound (crying clip, described earlier), lasting

5s or until the participant pressed the target “k” key. This feature of the program was designed to model the temporary suppression of that destructive behavior that may maintain the caregiver’s reaction via negative reinforcement. Prior to initiating the phase, participants were provided with the following instructions:

> “The goal of this game is to earn points and to reduce the number of times you hear the baby crying. You will earn points by correctly pressing the letter on the keyboard corresponding to the letter you see or hear. When you hear the baby crying, press the letter ‘k’ to stop the sound.”

The ambient task (see above) was running throughout the phase and continued during the sound (i.e., letters continued to appear or sound while the aversive sound was present).

Intervention Comparison

No Alternative Phase. This phase lasted 10 min and was designed to model what a caregiver might experience after being instructed to ignore or minimize attention for destructive behavior with no prompted alternative activity. Prior to initiating this phase, participants were provided with the following instructions:

> “The goal of this game is to earn points and reduce the number of times you hear the baby crying. You will continue to earn points by correctly pressing the letter on the keyboard corresponding to the letter you see or hear. When you hear the baby crying, you can press the letter ‘k’ to stop the sound, but the baby will continue crying throughout the phase. By inhibiting, or not pressing the ‘k’ key, the baby crying will occur less often over time, and you can receive bonus points.”

If participants pressed the “k” key in this condition, the sound terminated immediately but continued to occur approximately every 8s. If participants inhibited a response to the “k” key in the presence of the sound, the sound continued for 5 s, but the inter-sound interval increased

cumulatively (i.e., 8s, 16s, 32s) each time participants successfully inhibited the target "k" key. Thus, the 10-min session duration allowed for the full suppression of the aversive sound if the participant inhibited all "k" key presses; this feature of the program was designed to model extinction of destructive behavior with correct adherence to the extinction component of treatment. If participants pressed the "k" key following initial inhibition, the inter-sound interval stayed at the interval the participant had progressed to and did not increase farther unless the participant inhibited responding again. The computer program awarded 5 bonus points for inhibiting a response to the "k" key in the presence of the aversive sound throughout the phase. Bonus points were designed to model a researchers' use of additional incentives (e.g., positive feedback, tangible incentives) for correct treatment implementation during caregiver training.

Alternative Task Phase. This phase lasted 10 min and was designed to model what a caregiver might experience after being instructed to ignore or minimize attention for destructive behavior while also being provided a specific alternative task (pressing the "o" key) in which to engage. During this phase, participants were provided with the following instructions:

> "The goal of this game is to earn points and reduce the number of times you hear the baby crying. You will continue to earn points by correctly pressing the letter on the keyboard corresponding to the letter you see or hear. When you hear the baby crying, you can still press the letter 'k' to stop the sound, but the baby will continue crying throughout the session. If you press the "o" key when you hear the baby cry, crying will occur less often over time, and you can receive bonus points."

The computer program responded to "k" key pressing in an identical manner to that described during the No Alternative Task phase above. Pressing the "o" key was scored but was not intended to change the schedule of aversive sound presentations during this phase (i.e., inhibition of "k" increased the inter-sound-interval regardless of whether "o" was pressed or not); this feature of the program was designed in acknowledgement that caregiver alternative responding in

the typical environment would likely have no effect on the occurrence of destructive behavior independent of implementation of extinction. However, upon retrospective review the data outputs, this feature of the program performed differently than anticipated. The inter-sound-interval time did not increase for inhibiting "k" until the participant sampled engagement on the alternative "o" key. In other words, if a participant did not ever engage in the alternative "o" key but was successful at inhibiting "k," the aversive sound continued at the baseline rate of once every 8 s. If the participant engaged in the alternative "o" key while inhibiting "k" at some point during this phase, subsequent inhibition of "k" did increase the inter-sound-interval regardless of engagement with the alternative "o" (i.e., program ran as planned).

The computer program awarded 5 bonus points for inhibiting a response to the "k" key *and* engaging in the alternative response (i.e., pressing "o") in the presence of the aversive sound during this phase. Again, these bonus points were designed to model a researchers' use of additional incentives (e.g., positive feedback, tangible incentives) for correct treatment implementation during caregiver training.

Dependent Measures and Experimental Design

The dependent variables were the rate of "k" key pressing (target response), the rate of "o" pressing (alternative response), and response accuracy (% correct) on the ambient task (a measure of general responsiveness to the program). Errors of commission were defined by maintenance of "k" pressing during the Alternative Task or No Alternative phases of the study.

An ABAC or ACAB reversal design was used to compare the primary dependent measure (rate of "k") across baseline and intervention phases. The program randomly selected whether each participant would experience the No Alternative Phase or Alternative Task phase first in the reversal. The phases were counterbalanced across participants to better reveal any potential carryover effects across conditions.

CHAPTER 2: RESULTS

Figure 2 depicts the average target responses (i.e., "k" presses) across phases for all 14 participants. Target responding occurred at an average rate of 3.52 during the first baseline phase (SD = 3.52), 1.01 during the No Alternative phase (SD = 2.55), 4.52 during the second baseline phase (SD = 3.38), and 2.19 during the Alternative Task phase (SD = 3.16). Thus, errors of commission were generally and approximately equally reduced during the intervention phases for the group as a whole.

We supplemented this group-level analysis with visual inspection of individual data to identify intra-subject variability in our participant sample. Doing so revealed four unique response patterns (i.e., categories). We labeled the first category as *sensitive*, for participants who showed consistent "k" pressing in the baseline phase, consistent suppression of "k" pressing during the intervention phases, and high accuracy on the ambient task. These participants' response patterns appeared to be sensitive to both programed positive (i.e., point accrual in the ambient task; bonus points in the intervention phases) and negative (i.e., removal of crying) reinforcement contingencies, and/or sensitive to the rules. Participants whose responding did not appear sensitive to both contingencies fit into three distinct categories. The *positive reinforcer only* category showed sustained high accuracy on the ambient task but no "k" pressing (or a sharp decreasing trend) during baseline phases. Response patterns in this category suggested motivation by the accrual of points on the ambient task with little to no responding related to the negative reinforcement contingencies programmed for the cessation of the audio clip. The *commission errors* category included participants who continued to press the "k" key during the intervention phases. This response pattern suggested either strong control by the immediate negative reinforcer (i.e., sound removal) programmed for "k" pressing, insensitivity to the delayed reinforcer (i.e., inter-sound-interval increases) for inhibiting "k" pressing, lack of rule governance, and/or insensitivity to the bonus points for response inhibition. Finally, the *unclear*

category included participants for whom responding was inconsistent (e.g., lack of replication in the reversal to baseline) and difficult to interpret. We also analyzed rate of ambient task completion within and across groups (see Fig 4-7) and did not see any noteworthy differences (ranged from 10.5 – 32.5 per min). Across groups mean and standard deviation were as follows, *sensitive* (M = 19.31, SD = 4.05), *positive reinforcer only* (M = 15.05, SD = 1.43), *commission errors* (M = 24.98, SD = 3.92), and *unclear* (M = 20.53, SD = 4.68).

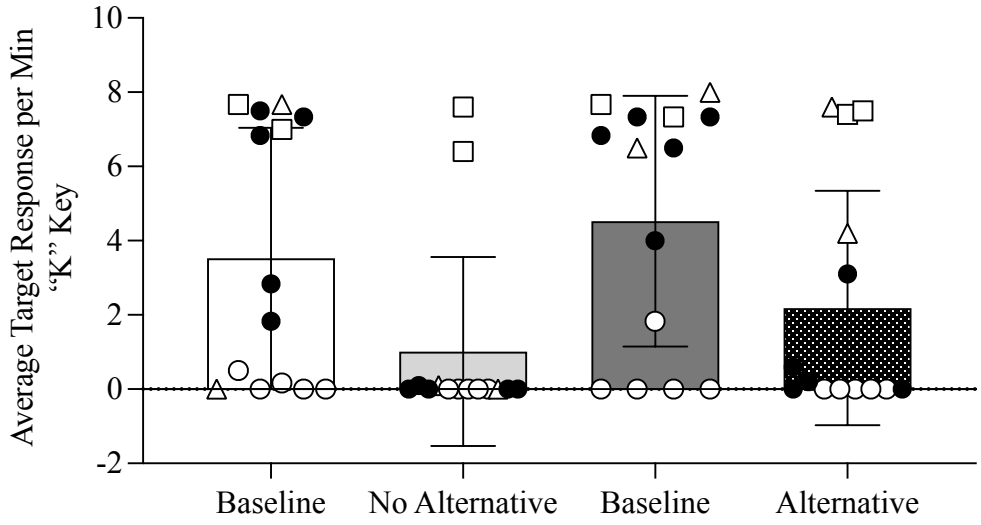

Figure 2: Average target responses across phases for all participants. N=14 participants. Bar represents mean, whiskers represent standard deviation. Closed circles = sensitive group; open circles = positive reinforcer only group; open squares = commission errors group; open triangles = unclear group.

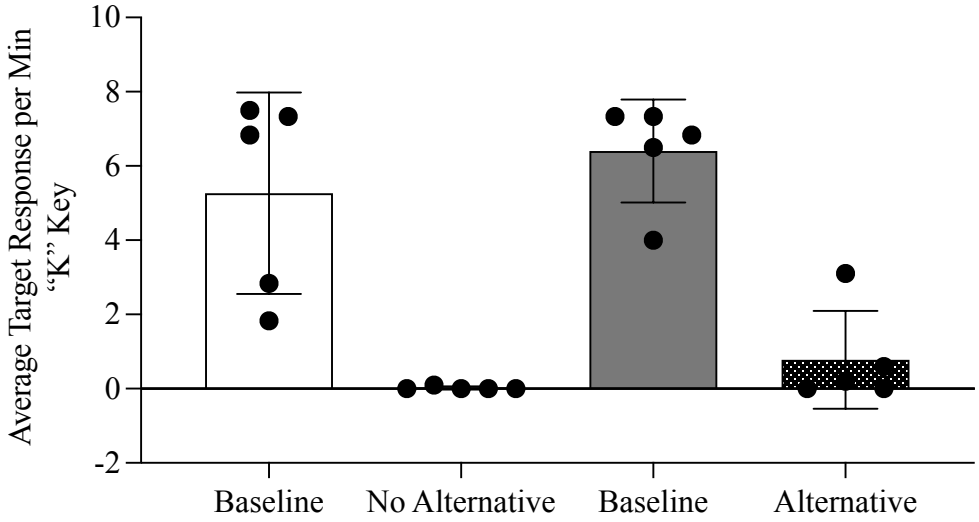

Figure 3: Average target responses across phases for Sensitive group. *n* = 5 participants. Bar represents mean, whiskers represent standard deviation.

Response patterns for five participants (staff *n* = 4; Mturk *n* = 1), matched the definition of responding *sensitive* to both positive and negative reinforcement contingencies programmed during the baseline and intervention phases. Figure 3 depicts the average rate of target responses (i.e., "k" presses) across phases for this group only. Target responding occurred at an average rate of 5.27 during the first baseline phase (SD = 2.71), 0.02 during the No Alternative phase (SD = 0.04), 6.4 during the second baseline phase (SD = 1.39), and 0.78 during the Alternative Task phase (SD 1.32). Visual inspections show clear differences in the average level of responding across baseline and intervention phases with less variability (i.e., a lower standard deviation) compared to the same analysis conducted with all participants' data (Figure 2). However, this group still showed substantial overlap in "k" responding across the experimental conditions.

Figure 4 depicts results for each participant meeting criteria to be included in the *sensitive* group across dependent measures. All participants engaged in consistent "k" presses in both baseline phases, inhibited "k" presses during both intervention phases, engaged in "o" presses during the Alternative Task phase, and had steady accuracy on the ambient task throughout the experiment. One participant ($P15_{s;}$ staff) engaged in commission errors ("k" presses) during the Alternative Task phase at lower levels than in baseline and on a decreasing trend; thus, their responses were still considered as sensitive to the programmed contingencies. Additionally, one participant ($P10_{m;}$ MTurk worker) responded on the "o" key in both the Alternative Task and No Alternative phases, showing potential carryover effects.

Of the nine participants not included in the *sensitive* group, five participants' (all Mturk workers) responses appeared to be sensitive to the *positive reinforcers only.* Figure 5 depicts results for participants in this category. Across these participants, responding to the target "k" key (i.e., negative reinforcement contingency) was low to zero, engagement with the alternative "o" key was also low to zero, and ambient task performance was high. $P5_{m}$ showed an initially high, but sharply decreasing level in target "k" key responding during baseline phases, which was

accompanied by an increasing trend in ambient task scores during this phase. This suggests a potential shift in motivation toward positive reinforcement (point earning) across the baseline phases for this participant. The programming error that occurred specifically impacted this group in that these participants inhibited pressing “k” in both intervention phases but only experienced the increased inter-sound-interval contingency in the No Alternative phase. To maintain full transparency in this feasibility study, we included all data sets.

Figure 6 depicts results for two participants’ (one staff, one Mturk worker) response patterns that were categorized as *errors of commission*. Both participants continued to press the target “k” key during the Alternative Task phase while also engaging in the alternative “o” key. Additionally, both participants continued to engage with the target “k” key during the No Alternative phase.

Finally, Figure 7 depicts results for participants whose responding was *unclear.* For these two participants (one staff, one Mturk worker) patterns of responding were inconsistent and difficult to interpret. $P7_s$ engaged in steady rates of “k” pressing during both baseline phases and the Alternative Task phase and did not press “o” consistently during the latter. This participant showed zero levels of “k” pressing during the No Alternative Task phase, during which levels of accuracy during the ambient task were highest. $P11_m$ did not press “k” during the first baseline but pressed “k” at increasing levels during the second baseline; thus, replication was not achieved.

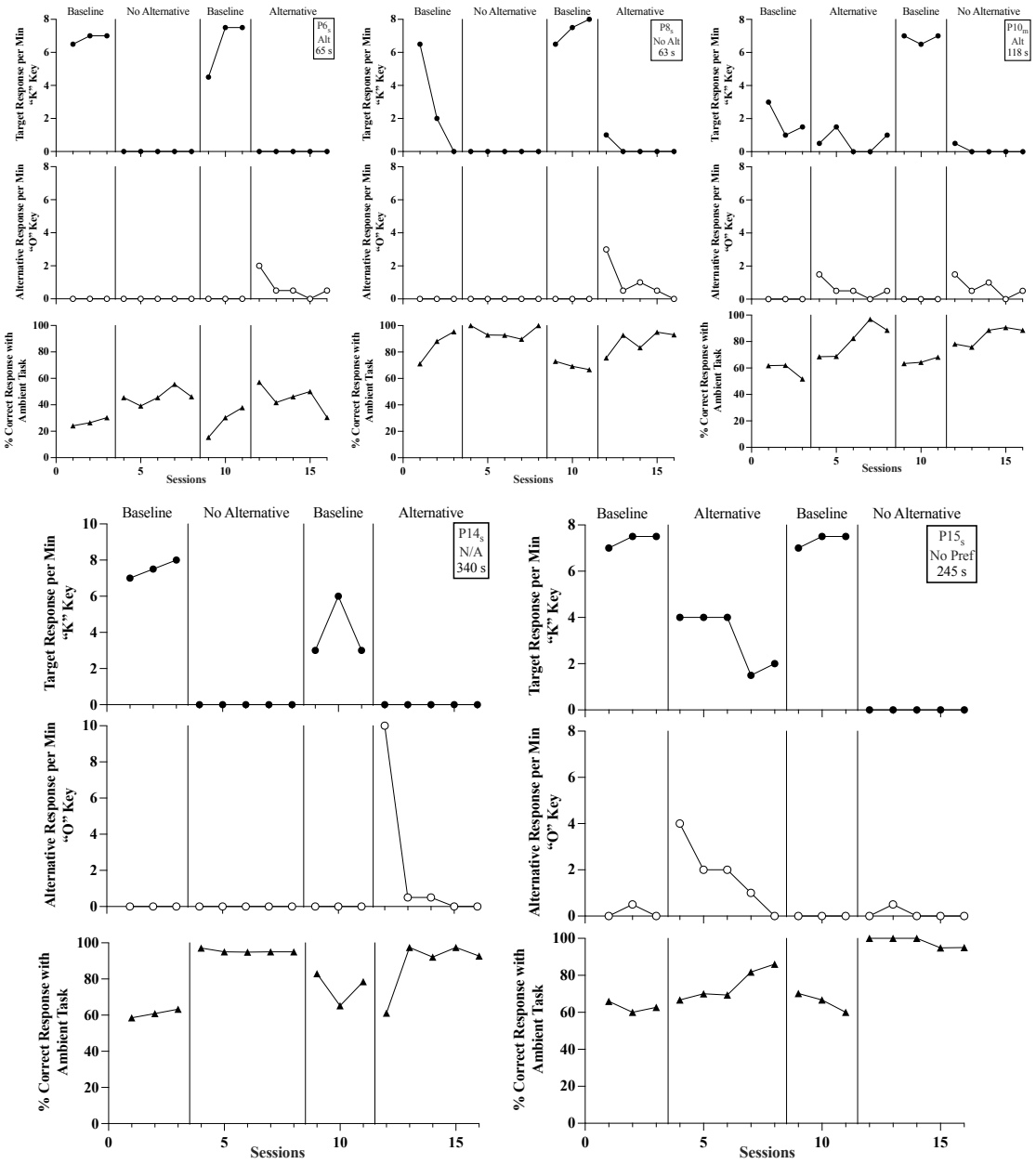

Figure 4: Dependent measures for Sensitive group. Box includes participant number, subscript $_m$ = Mturk worker, subscript $_s$ = staff, next row indicated participants selected preference (Alt = Alternative, No Alt = No Alternative, No Pref = No preference, N/A = information not available), last row indicates approximate time in seconds spent on rules pages.

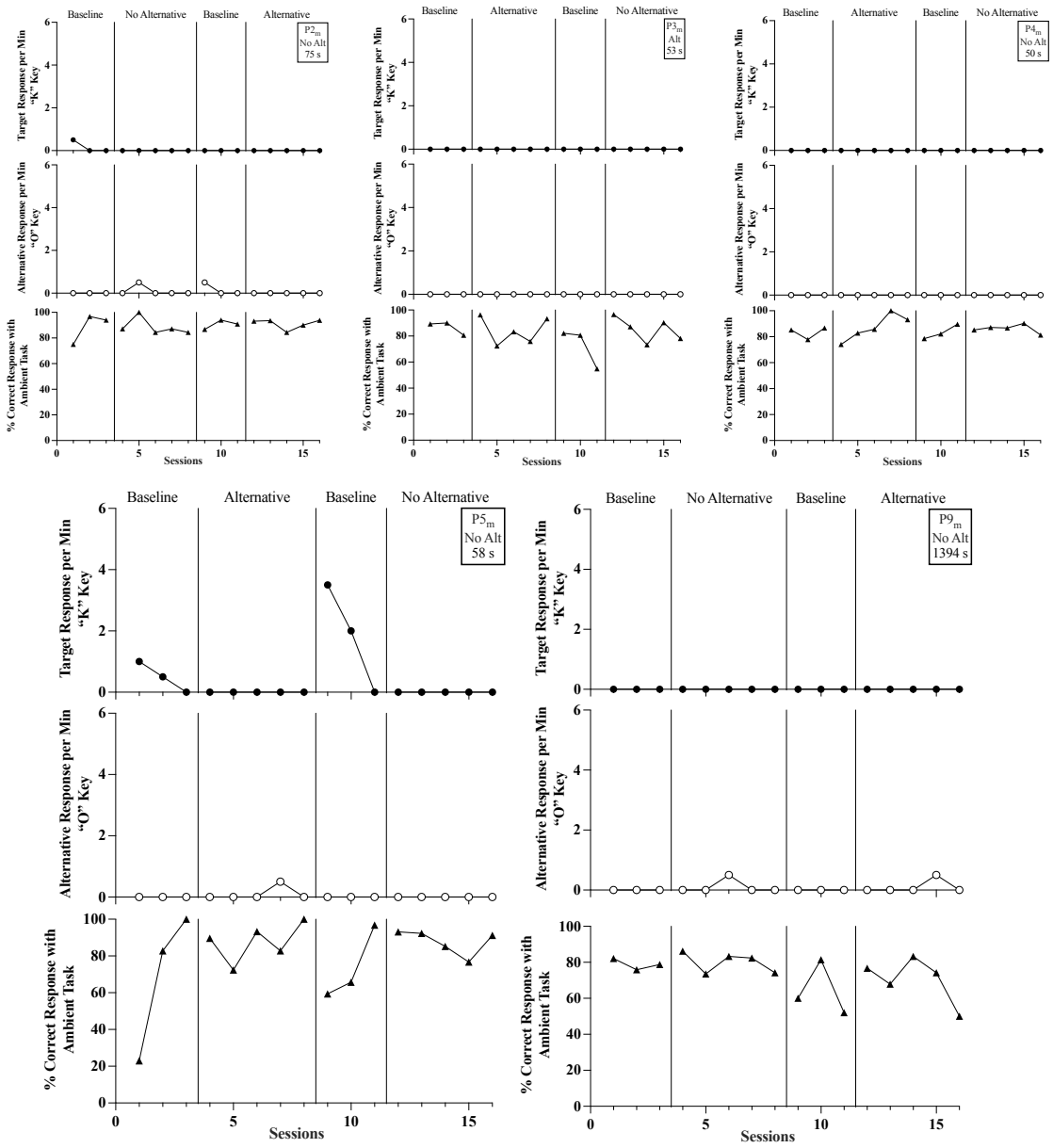

Figure 5: Dependent measures for Positive Reinforcer Only group. Box includes participant number, subscript $_m$ = Mturk worker, subscript $_s$ = staff, next row indicated participants selected preference (Alt = Alternative, No Alt = No Alternative, No Pref = No preference, N/A = information not available), last row indicates approximate time in seconds spent on rules pages.

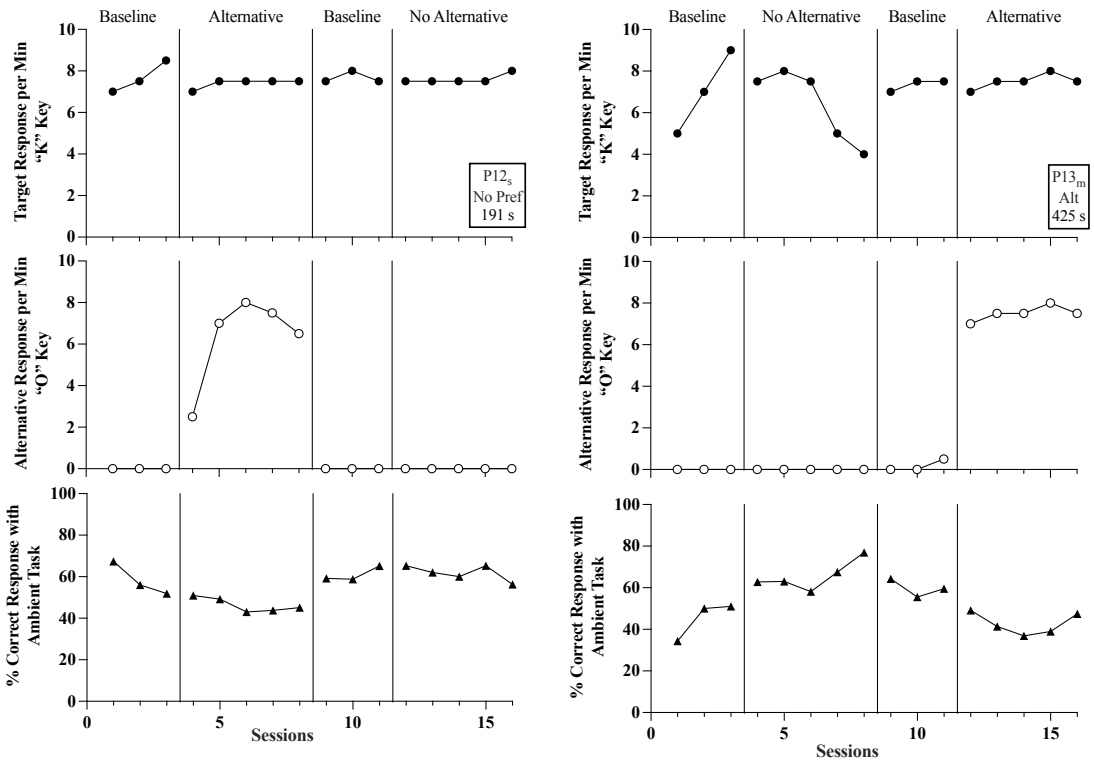

Figure 6: Dependent measures for Commission Errors group. Box includes participant number, subscript $_m$ = Mturk worker, subscript $_s$ = staff, next row indicated participants selected preference (Alt = Alternative, No Alt = No Alternative, No Pref = No preference, N/A = information not available), last row indicates approximate time in seconds spent on rules pages.

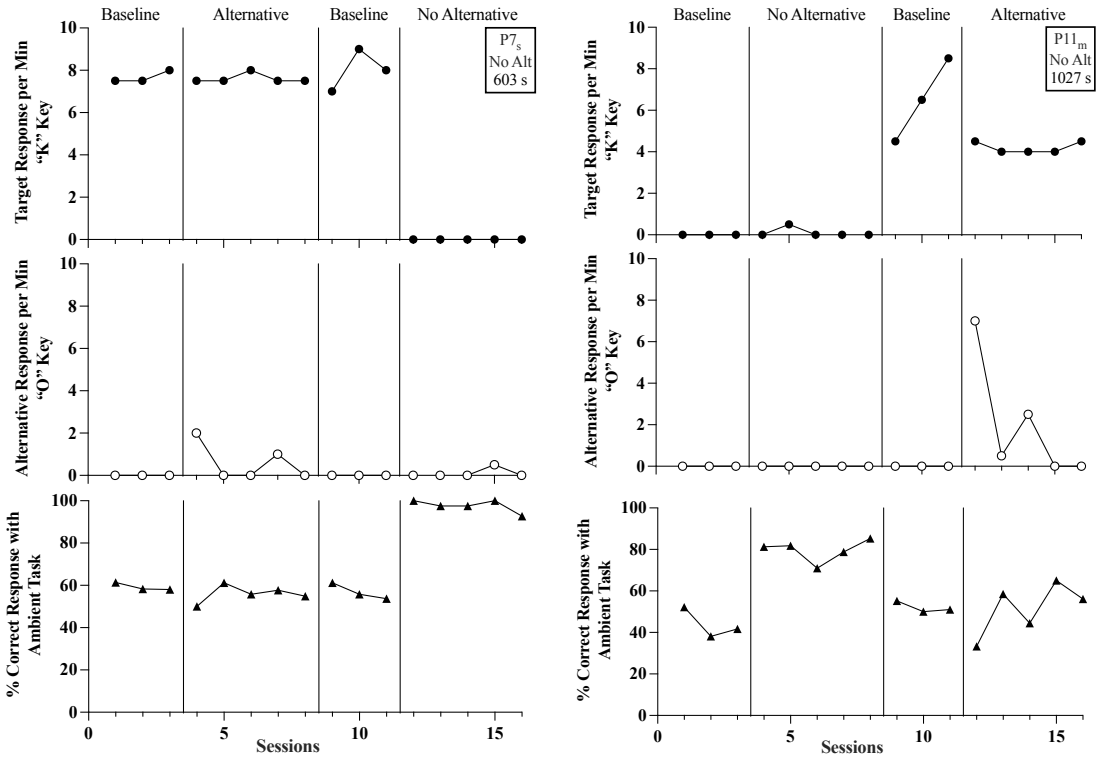

Figure 7: Dependent measures for Unclear group. Box includes participant number, subscript $_m$ = Mturk worker, subscript $_s$ = staff, next row indicated participants selected preference (Alt = Alternative, No Alt = No Alternative, No Pref = No preference, N/A = information not available), last row indicates approximate time in seconds spent on rules pages.

Figure 8 depicts reported preference across all participants (left panel) and across only those participants in the *sensitive* category (right panel). Across all participants (*n*=14), the No Alternative phase was reported as most preferred for the majority (*n* = 7). Across only those participants whose response patterns matched the programmed positive and negative contingencies (i.e., *sensitive; n*=5), the Alternative Task phase was reported as the most preferred for the 2 of the 5 participants. One participant selected no preference, one participant selected the No Alternative phase, and one participant did not complete the preference questionnaire in the sensitive group.

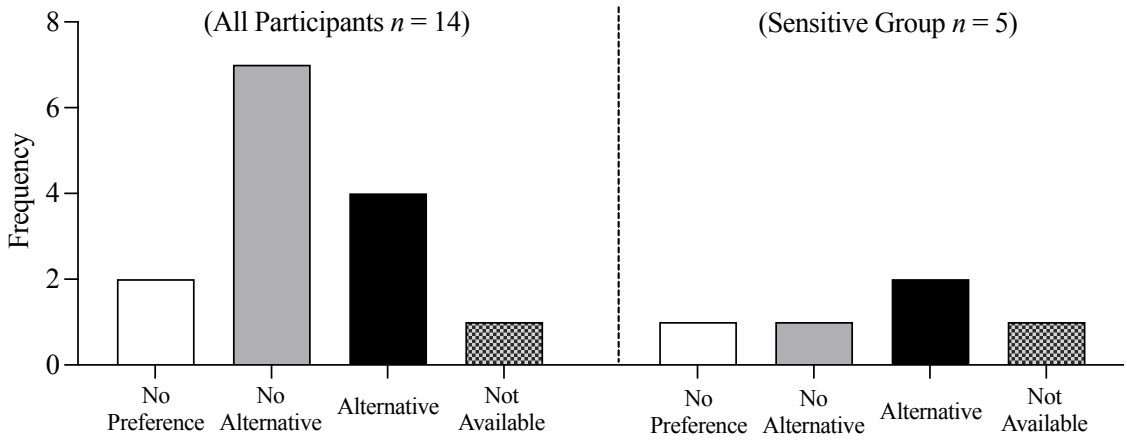

Figure 8: Participant preference for all participants and for Sensitive group. Left panel includes all participants who met criteria for participation. Right panel indicates preference for subgroup, *sensitive* group.

CHAPTER 3: DISCUSSION

The purpose of this study was twofold. First, we sought to assess the feasibility of a laboratory model for studying strategies for reducing treatment integrity errors. On average, target responding ("k" pressing in the laboratory model) was higher in baseline phases than experimental phases across all participants (see Figure 2), roughly approximating conditions in which caregivers may respond to destructive behavior and experience its temporary suppression prior to training on planned ignoring. Thus, the baseline performance in our laboratory model matched what would be considered natural caregiver responses to destructive behavior, lending confidence to our laboratory model. During experimental phases, participants showed four distinct patterns of responding that suggested sensitivity to various programmed contingencies and/or their accompanying rules. Contingencies included positive reinforcement (i.e., point accrual and backup monetary rewards) and negative reinforcement (i.e., removal or delay of aversive sound) programmed in the task. These various sensitivities provided further evidence that the laboratory model was sensitive to individual differences, and that future research on this topic should continue to include single-case design to detect these differences.

To our knowledge this is the first study to evaluate the use of a laboratory model to explore caregiver strategies to mitigate commission errors. Importantly, the use of a laboratory model to explore this context provided a safe and efficient way to produce preliminary data. The use of analog responses (i.e., keyboard presses) and consequences (i.e., points, aversive sound) allowed the researchers to model the reinforcement and extinction of destructive behavior, and errors of commission, without creating an environment that would be dangerous to children and their caregivers. In our development of the virtual environment, we attempted to include several contexts encountered by caregivers who are seeking training on treatment of their child's destructive behavior. This included negative reinforcement for caregiver reactions to destructive behavior, gradual reduction in child destructive behavior upon successful adherence to periods of

planned ignoring, positive reinforcement for child destructive behavior when errors are made, and rules. We also attempted to model daily life and promote sustained engagement with the study using an ambient task. Errors of commission occurred at some point during periods of planned ignoring for 7 out of the 14 participants indicating the need for further exploration to reduce errors of commission in this model. However, all features of the laboratory model seemingly controlled responding for only 5 of 14 participants; thus, our data should be considered preliminary and further refinements to the computer program may be necessary to understand the full range of variables responsible for caregiver treatment adherence (see below).

The second purpose of this study was to evaluate the effects of an alternative activity on adherence to planned ignoring in the laboratory model and to ascertain preference for alternative activities. However, preliminary data showed that access to the alternative task ("o" pressing) did not improve adherence to planned ignoring (i.e., "k" inhibition). That is, for the five participants who showed sensitivity to all contingencies, treatment adherence was roughly equal (n = 4) or worse (n = 1; $P15_s$) in the Alternative Task phase compared to the No Alternative phase. Although two participants reported preference for access to the alternative activity, several other participants within and outside of this group showed preference for no alternative activity or indifference. Although these outcomes seem to suggest abandoning the approach of providing caregivers with an alternative activity, we feel the outcomes require a more nuanced discussion and consideration for future study.

Ultimately, only 5 out of 14 participants responded in a manner that indicated sensitivity to both positive and negative reinforcement contingencies (Figure 4). By reviewing response patterns from the remaining 9, we can attempt to improve programming for future studies and fix known programming errors. A detailed list of specific considerations for future research on this laboratory model can be found in Table 2. This table outlines the features of the laboratory model

programmed in the current study, their rationale, and potential points of revision based on our preliminary data.

For example, one key consideration is that participants were recruited from two sources, with only the Mturk workers receiving monetary compensation. Due to concerns for the length of time of the study, we promoted continued engagement by telling Mturk participants their compensation would be based on participation in the ambient task (i.e., programmed positive reinforcement contingency). This contingency seemed to have a particularly strong impact on the Mturk workers, who fully comprised the group of individuals whose response patterns were sensitive to positive reinforcers only (see Figure 5). That is, these individuals did not engage in the target response to escape the aversive sound in baseline, precluding our ability to ask an experimental question about commission errors. Additionally, 4 out of the 5 participants whose responding appeared sensitive to both positive and negative reinforcement contingencies (Figure 4) were staff who were not provided additional compensation; the devaluing of the points for these staff may have increased their motivation to escape and avoid the aversive sound. Future iterations of this study should more systematically test the effects of removal (or reduction) of monetary reinforcement for engagement in the ambient task to evaluate how treatment integrity fluctuates based on the value of ambient tasks (e.g., if their value is high, caregivers may be more likely to ignore destructive behavior).

Table 2: Paradigm Features and Change Considerations

Paradigm Feature	Rationale	Possible Change(s)	Rationale
Lab model	Safely model scenarios in which caregivers respond to destructive behavior.	Virtual reality	Provide more realistic context to assess caregiver behavior while also maintaining safety
Visual and audio task component	Analogous to the potential variety of tasks caregivers are engaging in on a daily basis, requiring different levels of attention.	Pause the task while the crying clip plays	Decrease incentive to engage in competing tasks when the crying clip is present, similar to what a caregiver may experience when a child is exhibiting destructive behavior
Points provided for correct responses on the ambient task	Analogous to reinforcement for daily life tasks (engaging with others, chores, etc.).	Change point accrual to a variable schedule, remove points contingent on mistakes, remove points contingent on commission errors.	Continue to promote engagement in the ambient task while making it more realistic to the natural environment and better incentivizing adherence to the rules.
Monetary incentive	A primary incentive for participation in the study (Mturk workers) and a backup incentive for ambient task engagement and rule adherence.	Remove or reduce the monetary incentive or provide it contingent on other features of the task.	Reduce control by the positive reinforcer for ambient task engagement, potentially increasing sensitivity to other programmed contingencies.
Bonus point contingency	Analogous to receiving praise for engaging in the correct response during training	Increase the bonus points for correct responses; remove points (response cost) for incorrect responses.	To increase saliency of correct responses and to improve control by the positive reinforcer programmed for treatment adherence.

Crying clip of baby crying	This feature provided a safe and efficient way to simulate a singular aspect of destructive behavior that caregivers may experience.	Vary the clip to include other sounds (e.g., older child yelling, other clips of crying); conduct an aversiveness assessment across multiple clips and pick the clip rated most aversive by the participant.	Reduce potential habituation to the singular audio clip; customize the programming.
Temporary cessation of the crying clip (immediate negative reinforcement)	Analogous to the brief period of negative reinforcement caregivers may experience if they attend to attention-maintained problem behavior.	No changes suggested at this time.	N/A
Progressive inter-trial-time increases in the crying clip contingent on omission of participant responding	Analogous to the gradual decrease in destructive behavior when placed on extinction.	Return to the baseline rate of audio clip presentation following commission errors.	This change would better approximate a scenario when relapse in destructive behavior occurs due to errors of commission
Rules	Trainers typically provide caregivers with rules prior to practice implementing a treatment	Provide models and practice scenarios with feedback on correct and incorrect performance.	Due to response patterns observed, it is unclear if participants attended to and/or understood the rules. By providing additional training, we would anticipate fewer inconsistencies with responses and would be able to state with more confidence that errors occurred due to distinct experiences with the conditions.

Note. N/A=Not applicable.

We would also like to note that for the participants whose responding was sensitive to positive reinforcers only, they did not experience the increase in inter-sound-interval during the Alternative phase due to a programming error. That is, during the Alternative phase, participants in this group heard the aversive sound at baseline levels. Not experiencing the inter-sound-interval increase (due to not pressing "o" at least once) during the Alternative phase may have impacted preference for participants in this group. For the remaining four participants who consistently pressed "k" (*commission errors*) or whose responding was unclear, it is possible that the there was strong control by the immediate negative reinforcer (i.e., sound removal; see Figure 6) and/or lack of rule governance. Rule statements were provided prior to each phase but it is possible the rules were not clear, the participant did not attend to the rules, or the participant's responding was simply not affected by the rules. To address this concern, future research should consider the use of behavior skills training to teach the skills and/or provide practice opportunities for participants to engage in the desired behavior and sample the programmed contingencies before entering the experimental phases (Brookman-Frazee et al., 2009; Moore & Amado, 2021).

In addition to improving the laboratory model, future research should consider asking these same questions using other experimental designs. The current reversal design did not allow for replication of the intervention phases, or use of stability criteria, as is recommended for single subject design (Kratochwill et al., 2010). Additionally, the use of rate as a dependent measure may not be the best measure to indicate performance. Future research should consider percentage of opportunities to engage in or inhibit the target and alternative response. Due to time constraints, we were not able to graph the current data in this manner. Programmed design features were a result of time constraints of the laboratory model, our desire for each participant to contact both intervention phases (inter-subject comparison) to assess relative preference, and our interest in keeping study compensation equitable across participants. However, this lack of replication of each experimental phase further compromised our conclusions. Future research

should consider extending the task to allow for replication of experimental conditions. Other potential design options (multielement, multiple baseline across participants, group design) can also be explored, but these present their own limitations, such as increasing the potential for carryover effects (multielement), exposing participants with different histories of negative reinforcement baselines (multiple baseline), or limiting our ability to assess preference across conditions and explore intra-subject variability (group design).

A final approach to consider is exploring this experimental question in contexts where the stakes (effort, motivation, authority) are higher. The use of virtual environments in which child destructive behavior, and therapist feedback, is simulated is one direction to consider. Across disciplines, virtual reality has been used to increase motivation in physical therapy rehabilitation (Bonnechère, et al., 2016), safely expose and treat individuals with anxiety (Krijn et al., 2004), specific phobias (Botella et al., 2017), and post-traumatic stress disorders (PTSD; Gonçalves et a., 2012). By moving to closer approximations of the applied context, researchers can explore ways to individualize the prescribed alternative response (e.g., preference assessments) and test under conditions more closely resembling the caregivers experience with their own child. For example, the use of a single crying clip and arbitrary key-press responses were not like the environments caregivers operate in on a day-to-day basis. The experience of managing destructive behavior is highly effortful, physically risky, and emotionally draining; we recognize that the laboratory model decreased the effort, and likely the motivation, to escape our analog to destructive behavior as indicated by the 5 participants who never or rarely pressed “k” to escape the crying clip. In addition, the remote delivery of rules for a computer “game” may not adequately capture the social dynamics of a caregiver being trained to implement treatment by a behavior analyst, as suggested by the two participants who engaged in commission errors across the study. Virtual environments may better approximate these ecologically relevant factors.

We close by acknowledging that our study focused on a singular component of treatment packages (i.e., planned ignoring, a form of extinction) and that treatment approaches without extinction are increasing in popularity (Fritz et al., 2017; Kunnavatana et al., 2018; MacNaul & Neely 2018; Rajaraman et al., 2022). While we support continued research to create safe and sustainable treatment approaches without extinction, we do not anticipant extinction as a procedure going away. Extinction may still be used as a preventative strategy before intense behaviors emerge or in various shaping procedures used throughout behavioral acquisition programs (Cihon, 2022). Recently published variations on extinction (e.g., "kind extinction, Tarbox et al., 2023) also have involved providing an alternative to the functional reinforcer for destructive behavior; this strategy is compatible with the one evaluated in our laboratory model. The decision to include variations of extinction in any behavioral treatment should be made in consultation with trained providers. Future research on the use of treatment with and without extinction will continue to consider both reductions in child destructive behavior and adult treatment adherence. The current laboratory model provides a feasible and safe first step in exploring these many nuanced decisions.

BIBLIOGRAPHY

Athens, E. S., & Vollmer, T. R. (2010). An investigation of differential reinforcement of alternative behavior without extinction. *Journal of Applied Behavior Analysis*, *43*(4), 569-589. https://doi.org/10.1901/jaba.2010.43-569

Bonnechère, B., Jansen, B., Omelina, L., & Van Sint Jan, S. (2016). The use of commercial video games in rehabilitation: a systematic review. *International Journal of Rehabilitation Research*, *39*(4), 277-290.

Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017). Recent progress in virtual reality exposure therapy for phobias: a systematic review. *Current Psychiatry Reports*, *19*, 1-13.

Bowman, L. G., Hardesty, S. L., & Mendres-Smith, A. E. (2013). A functional analysis of crying. *Journal of Applied Behavior Analysis*, *46*(1), 317-321. https://doi.org/10.1002/jaba.4

Brookman-Frazee, L. I., Drahota, A., & Stadnick, N. (2012). Training community mental health therapists to deliver a package of evidence-based practice strategies for school-age children with autism spectrum disorders: A pilot study. *Journal of autism and developmental disorders*, *42*(8), 1651-1661. https://doi.org/10.1007/s10803-011-1406-7

Carr, E. G., & Carlson, J. I. (1993). Reduction of severe behavior problems in the community using a multicomponent treatment approach. *Journal of Applied Behavior Analysis*, *26*(2), 157-172. https://doi.org/10.1901/jaba.1993.26-157

Cihon, J. H. (2022). Shaping: A Brief History, Research Overview, and Recommendations. *Handbook of Applied Behavior Analysis Interventions for Autism: Integrating Research into Practice*, 403-415.

Drifke, M. A., Tiger, J. H., & Lillie, M. A. (2020). DRA contingencies promote improved tolerance to delayed reinforcement during FCT compared to DRO and fixed-time

schedules. *Journal of Applied Behavior Analysis*, *53*(3), 1579-1592. https://doi.org/10.1002/jaba.704

Emerson, E., Kiernan, C., Alborz, A., Reeves, D., Mason, H., Swarbrick, R., Mason, L., & Hatton, C. (2001). Predicting the persistence of severe self-injurious behavior. *Research in Developmental Disabilities*, *22*(1), 67-75. Doi:10.1016/S0891-4222(00)00062-7

Fisher, W. W., Thompson, R. H., Hagopian, L. P., Bowman, L. G., & Krug, A. (2000). Facilitating tolerance of delayed reinforcement during functional communication training. *Behavior Modification*, *24*(1), 3-29. https://doi.org/10.1177/0145445500241001

Fritz, J. N., Jackson, L. M., Stiefler, N. A., Wimberly, B. S., & Richardson, A. R. (2017). Noncontingent reinforcement without extinction plus differential reinforcement of alternative behavior during treatment of problem behavior. *Journal of Applied Behavior Analysis*, *50*(3), 590-599. https://doi.org/10.1002/jaba.395

Gabor, A. M., Fritz, J. N., Roath, C. T., Rothe, B. R., & Gourley, D. A. (2016). Caregiver preference for reinforcement-based interventions for problem behavior maintained by positive reinforcement. *Journal of Applied Behavior Analysis*, *49*(2), 215-227. https://doi.org/10.1002/jaba.286

Ghaemmaghami, M., Hanley, G. P., & Jessel, J. (2021). Functional communication training: From efficacy to effectiveness. *Journal of Applied Behavior Analysis*, *54*(1), 122-143. https://doi.org/10.1002/jaba.762

Gonçalves, R., Pedrozo, A. L., Coutinho, E. S. F., Figueira, I., & Ventura, P. (2012). Efficacy of virtual reality exposure therapy in the treatment of PTSD: a systematic review. *PloS One*, *7*(12), 1-7. Doi:10.1371/journal.pone.0048469

Gregori, E., Rispoli, M., Neely, L., Lory, C., Kim, S. Y., & David, M. (2021). Training direct service personnel in functional communication training with adults with disabilities. *Journal of Developmental and Physical Disabilities*, *33*, 669-692.

Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquisto, J., & LeBlanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis*, *31*(2), 211-235. https://doi.org/10.1901/jaba.1998.31-211

Holden, B., & Gitlesen, J. P. (2006). A total population study of challenging behaviour in the county of Hedmark, Norway: Prevalence, and risk markers. *Research in Developmental Disabilities, 27*(4), 456-465. Doi: 10.1016/j.ridd.2005.06.001

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. *What works clearinghouse*. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kunnavatana, S. S., Bloom, S. E., Samaha, A. L., Slocum, T. A., & Clay, C. J. (2018). Manipulating parameters of reinforcement to reduce problem behavior without extinction. *Journal of applied behavior analysis*, *51*(2), 283-302. https://doi.org/10.1002/jaba.395

Lloveras, L. A., McKeown, C. A., Lichtenberger, S. N., Sellers, T. P., & Vollmer, T. R. (2023). Recommendations Regarding Use of the Term "Ignore" in Applied Behavior Analysis. *Perspectives on Behavior Science*, 1-10. https://doi.org/10.1007/s40614-023-00373-2

Lowe, K., Allen, D., Jones, E., Brophy, S., Moore, K., & James, W. (2007). Challenging behaviours: Prevalence and topographies. *Journal of Intellectual Disability Research, 51(*8), 625-636. Doi: 10.1111/j.1365-2788.2006.00948.x

Luczynski, K. C., & Hanley, G. P. (2009). Do children prefer contingencies? An evaluation of the efficacy of and preference for contingent versus noncontingent social reinforcement during play. *Journal of Applied Behavior Analysis*, *42*(3), 511-525. https://doi.org/10.1901/jaba.2009.42-511

MacNaul, H. L., & Neely, L. C. (2018). Systematic review of differential reinforcement of alternative behavior without extinction for individuals with autism. *Behavior Modification, 42*(3), 398-421. https://doi.org/10.1177/0145445517740321

Melanson, I. J., & Fahmie, T. A. (2023). Functional analysis of problem behavior: A 40-year review. *Journal of Applied Behavior Analysis*, *56*(2), 262-281. https://doi.org/10.1002/jaba.983

Miller, J. R., Lerman, D. C., & Fritz, J. N. (2010). An experimental analysis of negative reinforcement contingencies for adult-delivered reprimands. *Journal of Applied Behavior Analysis*, *43*(4), 769-773. https://doi.org/10.1901/jaba.2010.43-769

Mitteer, D. R., Greer, B. D., Fisher, W. W., Briggs, A. M., & Wacker, D. P. (2018). A laboratory model for evaluating relapse of undesirable caregiver behavior. *Journal of the Experimental Analysis of Behavior*, *110*(2), 252-266. https://doi.org/10.1002/jeab.462

Moore, T. R., & Amado, R. S. (2021). A conceptual model of treatment adherence in a behavior analytic framework. *Education & Treatment of Children, 44*(1), 1–17. https://doi.org/10.1007/s43494-020-00032-0

Nuemann & waters, 2006 Neumann, D. L., & Waters, A. M. (2006). The use of an unpleasant sound as an unconditional stimulus in a human aversive Pavlovian conditioning procedure. *Biological Psychology*, *73*(2), 175-185. https://doi.org/10.1016/j.biopsycho.2006.03.004

Pipkin, C. S. P., Vollmer, T. R., & Sloman, K. N. (2010). Effects of treatment integrity failures during differential reinforcement of alternative behavior: A translational model. *Journal of Applied Behavior Analysis*, *43*(1), 47-70. https://doi.org/10.1901/jaba.2010.43-47

Rajaraman, A., Hanley, G. P., Gover, H. C., Staubitz, J. L., Staubitz, J. E., Simcoe, K. M., & Metras, R. (2022). Minimizing escalation by treating dangerous problem behavior within an enhanced choice model. *Behavior analysis in practice*, *15*(1), 219-242. https://doi.org/10.1007/s40617-020-00548-2

Romano, L. M., & St Peter, C. C. (2017). Omission training results in more resurgence than alternative reinforcement. *The Psychological Record*, *67*(3), 315-324. https://doi.org/10.1007/s40732-016-0214-z

Saini, V., Miller, S. A., & Fisher, W. W. (2016). Multiple schedules in practical application: Research trends and implications for future investigation. *Journal of Applied Behavior Analysis*, *49*(2), 421-444. https://doi.org/10.1002/jaba.300

Sanetti, L. M. H., & Collier-Meek, M. A. (2019). *Supporting successful interventions in schools: Tools to plan, evaluate, and sustain effective implementation*. Guilford Publications.

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, *38*(4), 445.

Smith, S. W., & Greer, B. D. (2022). Validating human-operant software: A case example. *Behavior Analysis: Research and Practice 22(4), 389–403*. https://doi.org/10.1037/bar0000244

Stocco, C. S., & Thompson, R. H. (2015). Contingency analysis of caregiver behavior: Implications for parent training and future directions. *Journal of Applied Behavior Analysis*, *48*(2), 417-435. https://doi.org/10.1002/jaba.206

Tarbox, C., Tarbox, J., Bermudez, T.L. *et al.* Kind Extinction: A Procedural Variation on Traditional Extinction. *Behavior Analysis Practice*, 1-11. https://doi.org/10.1007/s40617-023-00833-w

Thompson, R. H., Bruzek, J. L., & Cotnoir-Bichelman, N. M. (2011). The role of negative reinforcement in infant caregiving: An experimental simulation. *Journal of Applied Behavior Analysis*, *44*(2), 295-304. https://doi.org/10.1901/jaba.2011.44-295

Tiger, J. H., Hanley, G. P., & Bruzek, J. (2008). Functional communication training: A review and practical guide. *Behavior Analysis in Practice*, *1*(1), 16-23. https://doi.org/10.1007/BF03391716

Trump, C. E., Ayres, K. M., Quinland, K. K., & Zabala, K. A. (2020). Differential reinforcement without extinction: A review of the literature. *Behavior Analysis: Research and Practice*, *20*(2), 94. https://doi.org/10.1037/bar0000169

Vollmer, T. R., Peters, K. P., Kronfli, F. R., Lloveras, L. A., & Ibañez, V. F. (2020). On the definition of differential reinforcement of alternative behavior. *Journal of Applied Behavior Analysis*, *53*(3), 1299-1303. https://doi.org/10.1002/jaba.701

Wathen, S. N., & Podlesnik, C. A. (2018). Laboratory models of treatment relapse and mitigation techniques. *Behavior Analysis: Research and Practice*, *18*(4), 362. http://dx.doi.org/10.1037/bar0000119

APPENDIX A: CONSENT FORM

Title of this Research Study
Human Response Persistence Evaluation Using Internet-Based Software

You are invited to participate in this research study that is being conducted by Jessie Weber, a graduate student at the University of Nebraska Medical Center. The information in this consent form is provided to help you make an informed decision whether or not to participate. If you have any questions, please do not hesitate to ask.

The purpose of the research is to test how people respond to the sound of an infant crying and task persistence. You will be asked to play a matching game made up of auditory and visual letters appear on the screen or sound through your headphones and you will select to the letter on your keyboard. You will receive points for each letter you match. During the experiment, you will hear an infant crying and be given instructions on how to stop the sound or make it happen less often. By following the instructions for each section, you will hear the infant crying less.

The information will be anonymously collected. No one will know which responses are yours. Participation in the study will take about 60 minutes with the opportunity for self-paced breaks. We anticipate that up to 200 subjects will take part in the study.

The only risks and discomforts you might experience by taking part in this research include boredom and minor frustration from hearing the infant crying. There are not direct benefits for you from taking part in this study.

You will be paid up to $5. The base rate for completing the experiment will be $0.10. You will be paid a bonus based on your performance during the experimental task. If you score more points by matching the letters correctly, you will be paid more.

The research is anonymous. Your Amazon MTurk Worker ID will be kept for the duration of the study to ensure that nobody completes the study more than once, but these data will be destroyed upon completion of the study. No other information will be collected that can identify who you are. As an additional safeguard, we will store all data on encrypted, password-protected computers and servers at the University of Nebraska Medical Center. Only researchers working on this experiment will have access to the data. Data on your performance in the study that is not personally identifiable will be kept for seven years following completion of the study. After the study is over the information collected for this research will not be used or distributed to investigators for other research.

The research team and the Institutional Review Board at the University of Nebraska Medical Center are the only parties that may see the data, except as may be required by law. If the findings of this research are professionally presented or published, only group results will be stated.

It is your choice whether you take part in the research. You may choose to take part, not to take part or you may change your mind and withdraw from the study at any time. If you do not want to enter the study or decide to stop taking part, your relationship with the study staff will not change, and you may do so without penalty and without loss of benefits to which you are otherwise entitled. Please note, however, that once you have submitted your responses, you may no longer withdraw them as we will not know which ones yours are.

If you have questions about taking part in this study, you can contact the Principal Investigator: Jessie Weber, Munroe Meyer Institute 6902 Pine St. Omaha NE, 68106. You can also contact Jessie Weber by email at jessie.weber@unmc.edu.

If you have questions about your rights as a research subject, you can contact the IRB at: University of Nebraska Medical Center IRB Phone Number (402) 559-6463 or email at irbora@unmc.edu

Please print a copy of this consent form for your records. If you are 19 years of age or older, understand the statements above, and consent to take part in the study, click on the "I Agree" button to begin the research. If not, please click on the "I Do Not Agree" button which will exit you from this screen/program.

Principal Investigator: Jessie Weber, jessie.weber@unmc.edu

APPENDIX B: CONSENT QUIZ

IRB # 0806-22-EX

Prospective Participants will be required to correctly answer 100% of the following questions to participate in the research:

Do you have to participate in this research?
- A. Yes
- B. No

If you start the experiment, do you have to finish the experiment?
- A. Yes
- B. No

When can you withdraw consent?
- A. I cannot withdraw consent
- B. Only before I start the experiment
- C. Only after I finish the experiment
- D. At any time

How can you withdraw my consent?
- A. Contacting the primary investigator by email at jessie.weber@unmc.edu
- B. Contacting an administrator at Amazon Mechanical Turk
- C. Closing the window to the experiment on my computer
- D. Restarting my computer

What will your compensation be?
- A. $1
- B. $2
- C. Up to $5 based on performance during the experiment
- D. $20

What are the other direct benefits of you participation?
- A. There are no direct benefits
- B. I will receive treatment for an illness
- C. I will learn something new about myself
- D. I will be smarter

What are the risks involved with your participation?
- A. I might get sick
- B. I might lose money
- C. I might see something that will scare me
- D. There are no significant risks

How long will your participation last?
- A. About 15 min
- B. About 120 min
- C. About 30 min

D. About 60 min

What will you be doing in the experiment?

A. Reading stories and answering questions
B. Spelling words
C. Clicking buttons
D. Matching letters and hearing an infant crying

Will any personally identifiable information be collected about you?

A. Yes
B. No

What should you do if you have questions about the study?

A. Contact the principal investigator, Jessie Weber, at jessie.weber@unmc.edu
B. Contact my doctor
C. Contact a pharmacist
D. Contact a family member

What should you do if you have questions about your rights as a research participant?

A. Contact the principal investigator, Jessie Weber, at jessie.weber@unmc.edu
B. Contact the University of Nebraska Medical Center's IRB at (402)-559-6463 or write to them at 987830 Nebraska Medical Center Omaha, NE 68198-7830
C. Any of the above

APPENDIX C: DEMOGRAPHIC QUESTIONNAIRE

Participants will be asked to answer the following questions at the end of the research study:

What gender do you identify as?

A. Female
B. Male
C. (Fill in)
D. Prefer not to say

What is your age?

A. 19-30
B. 31-40
C. 41-50
D. 50+
E. Prefer not to say

Please specify your ethnicity. (Pick multiple)

A. Caucasian
B. African American
C. Latino or Hispanic
D. Asian
E. Native American
F. Native Hawaiian or Pacific Islander
G. Two or more
H. Other/Unknown
I. Prefer not to say

Are you married?

A. Yes
B. No
C. Prefer not to say

Do you have children?

A. Yes
B. No
C. Prefer not to say

What age are your children?

A. 0-2 years old
B. 3-5 years old
C. 6-10 years old
D. 10+ years old
E. Prefer not to say

Which condition did you prefer?

A. No alternative task available (yellow)
B. Alternative task available (green)
C. No preference

Please report if you multitasked or engaged in other activities while completing this study? This will not impact your compensation.

A. Yes
B. No
C. Prefer not to answer

Where there any rules or strategies you used other than those provided to gain points?

[Fill in the blank]

Where there any rules or strategies you used to ignore or stop the noise other than those provided?

[Fill in the blank]

APPENDIX D: DEBRIEFING FORM

IRB # 0806-22-EX

Title of this Research Study

Human Response Persistence Evaluation Using Internet-Based Software

Sometimes in behavioral research it is necessary for the experimenter to tell participants incorrect information because telling the participants all of the true information would change how they would behave during the experiment. In the experiment you just completed, you were told incorrect information before the experiment started.

You were told that you would get more money at the end of the experiment if you earned more points during the experiment. Instead, everyone who participated earned the same amount of money. We paid everyone the maximum amount of money that we said they could earn regardless of how many points they earned during the experiment.

We told you that you would get more money if you earned more points to make sure that you would be motivated to get points throughout the entire experiment. We did this because in our experiment we wanted to see how you persisted when presented with an aversive sound and how you chose to remove or reduce the sound. To do this, we needed to make it seem like the points that you were getting were rewarding by telling you that they would be worth money.

However, to ensure that every participant is treated equally, everyone who completes the same experiment should get paid the same amount of money, even if they behave differently during the experiment. Thus, we paid everyone the same amount, the maximum amount that we said they could earn.

When people know exactly what the researcher is studying, they often change their behavior, thus making their responses unusable for drawing conclusions about human nature and experiences. For this reason, we ask that you please not discuss this study with others who might participate any time after you.

If your participation in this study has in any way upset you, you may contact University of Nebraska Medical Centers Counseling Services at (402)-559-7276. Please be aware that any cost in seeking medical assistance is at your own expense.

If you have any questions about this study, feel free to ask the researcher, Jessie Weber, by email at jessie.weber@unmc.edu

If you have questions about your rights as a research subject, you can contact the IRB at: University of Nebraska Medical Center IRB Phone Number (402) 559-6463 or email at irbora@unmc.edu

Now that you understand the true nature of our study, we would like to give you the chance to refuse the use of your data for our research purposes. You are free to ask us not to use your data in our study analysis. If you have any concerns about your participation or the data you provided in light of this disclosure, please discuss this with us. We will be happy to provide any information we can to help answer questions you have about this study. Please again accept our appreciation for your participation in this study.

By clicking the “AGREE” button below, you acknowledge that you have read this debriefing form and you agree to allow the use of your data for research purposes.

By clicking the “DISAGREE” button below, you acknowledge that you have read this debriefing form and you would like your data to be immediately withdrawn and destroyed.

[Two buttons will appear below the text, one saying “AGREE” and one saying “DISAGREE.” This will indicate to the researcher whether to keep or destroy the data]

APPENDIX E: BASIC FUNCTIONALITY CHECKLIST

Parameter	Correct Functionality	
Point total visible	Yes	No
Ambient task target visible	Yes	No
Ambient task target audible	Yes	No
Aversive sound audible	Yes	No
Target key press results in aversive sound removal	Yes	No
Points add correctly	Yes	No
Points withheld correctly	Yes	No
Background colors	Yes	No

Note. Only select "Yes" when the software implements each parameter with 100% accuracy throughout the entire experiment (adapted from Smith and Greer, 2022). Use Basic Functionality to evaluate variables listed in Appendix F.

APPENDIX F: BASIC FUNCTIONALITY VARIABLES EVALUATED

Hardware		Different Laptops	Correct Functionality	
Operating Systems		Windows		
		macOS	Yes	No
Web browsers		Internet Explorer	Yes	No
		Microsoft Edge	Yes	No
		Google Chrome	Yes	No
		Firefox	Yes	No
		Safari	Yes	No
Stress test	Internet speed	High-speed	Yes	No
		Low-speed	Yes	No
	Internet traffic	High-traffic	Yes	No
		Low-traffic	Yes	No

Note. Check box next to a variable when Basic Functionality Checklist is scored as 100% "Yes" for that variable.