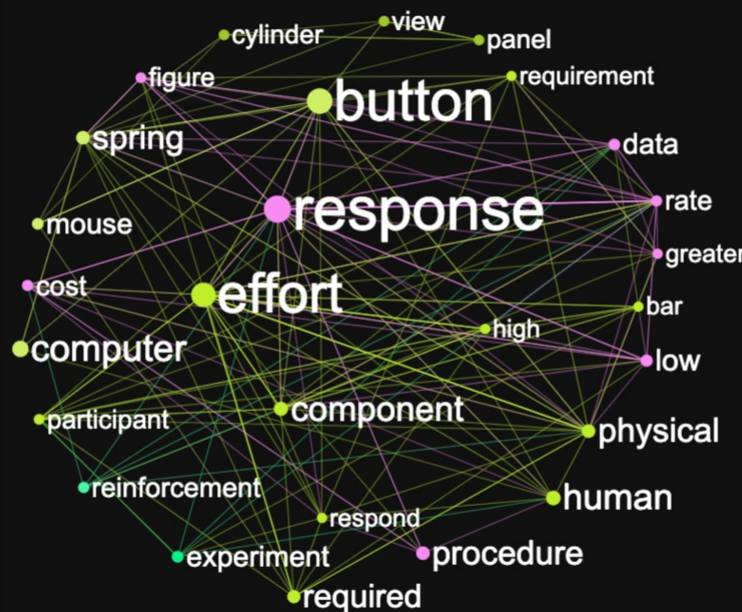
The Experimental Analysis of Human Behavior Special Interest Group Presents

THE BULLETIN

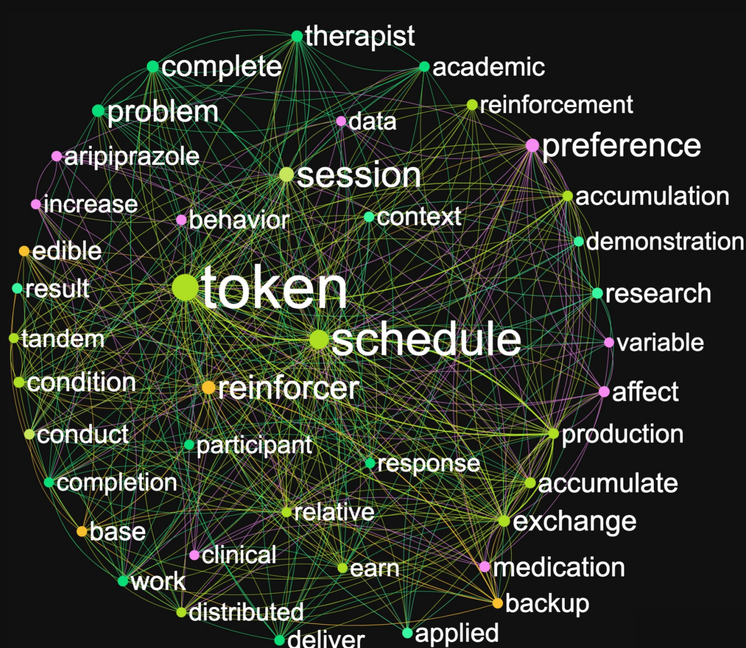
Network visualization of word sequences for all articles contained in this issue with degree greater than 20.

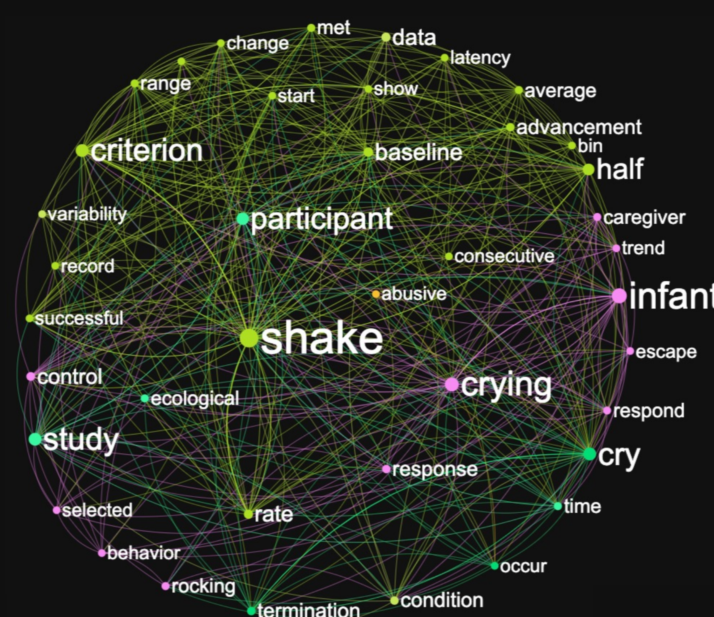
Volume 33 Spring 2023 Number 1

TECHNICAL REPORT



RESEARCH REPORTS





THE EXPERIMENTAL ANALYSIS OF HUMAN BEHAVIOR BULLETIN

About the Bulletin:

- The EAHB Bulletin (ISSN: 1938-7237) is published electronically on an ongoing basis by the Experimental Analysis of Human Behavior Special Interest Group (EAHG SIG), a group organized under the auspices of the Association for Behavior Analysis, International (ABAI). The EAHB SIG is dedicated to promoting and supporting basic behavior analytic research with human participants. Articles in the Bulletin represent the views of the authors. They are not intended to represent the approved policies of the EAHB SIG or ABAI, or the opinions of the membership of the EAHB SIG or ABAI. For more information, write to the Editor.

- The Bulletin intends to focus on the demonstration of experimental control through thoughtful experimental designs, not noteworthy or systematic findings. Therefore, we welcome manuscripts with null or unsystematic findings that may serve as fodder for future research and examples of methodological "lessons learned". In addition to experimental research on human behavior, we welcome conceptual papers, reviews, and technological papers, especially those that make inspire and improve the quality of future research in the experimental analysis of human behavior.

Types of Manuscripts Accepted

- Brief Reports (≤ 2,000 words): This section is for manuscripts that make an original, and valid contribution to the experimental analysis of human behavior. Examples include pilot studies, empirical reports, literature reviews, or theoretical treatises that address issues of interest the experimental analysis of human behavior. These submissions undergo a rigorous peer-review process and, once published, are considered final publications. As such, these reports have the status as work published in print journals.

- Research Reports: This section is for manuscripts related to the techniques and findings of studying human behavior. An experimental analysis refers to studies that systematically manipulate the environment surrounding and consequences of human behavior and measure changes in behavior using arrangements that demonstrate experimental control. These analyses can involve a large range of human populations, research designs, experimental apparatuses, and behavioral subject matter. These topics include, but are by no means limited to, discounting, stimulus control, stimulus equivalence, human-operant research, etc.

- Laboratory Descriptions (≤ 2,000 words): This section is intended for brief descriptions of laboratory procedures. Examples include novel methods or preparations, innovative data analytic methods, and descriptions of research strategies. These submissions are subject to editorial review only. These reports are not considered final publications.

- Technical Information: This section is intended for brief descriptions of technical information unique to the study of the behavior of human participants. Examples include, but are not limited to, descriptions of particularly effective programming code, innovative methods for stimulus presentation, response detection, or contingency management with human participants.

TECHNICAL REPORT

A RELATIVELY LOW-COST EQUIPMENT TO INVESTIGATE PHYSICAL EFFORT IN HUMANS

Raquel Fernanda Ferreira Lacerda[1]; André Luiz[1,2,3]; Carlos Eduardo Costa[1]

[1]UNIVERSIDADE ESTADUAL DE LONDRINA
[2]UNIVERSIDADE POSITIVO – FACULDADE LONDRINA
[3]INSTITUTO CONTINNUM

Response cost is a concept used to represent multiple relations among events (see Soares, Costa, Aló, Luiz, & Cunha, 2017 for a review of this literature). For example, Luce, Christian, Lipsker, and Hall (1981) have pointed out three procedures mainly called response-cost: (a) the increase in the physical effort required to respond (e.g., Alling & Poling, 1995, Experiment 1; Skinner & Morse, 1958; Solomon, 1948); (b) changes in the programmed contingency such as an increase in the ratio required to produce a consequence (e.g., Powell, 1968; Weiner, 1966; Winograd, 1965); (c) contingent reinforcement loss, such as point loss (e.g., Bolívar & Dallery, 2020; Cunha, Cordeiro, & Costa, 2018; Okouchi, 2015; Pietras & Hackenberg, 2005; Weiner, 1962, 1969). The current manuscript aims to describe relatively low-cost equipment that allows the investigation of the first response-cost procedure using humans as subjects.

The experimental analysis of human operant behavior can be seen as an intermediate field between non-human animal research and the development of solutions to social problems (Lattal & Perone, 1998). In this field, humans are frequently exposed to computer tasks controlled by software (Becker, 2011; Cabello et al., 2002, 2003; Costa & Banaco, 2002; Peirce et al., 2019; Roche & Dymond, 2003; Ruiz & Bermúdez, 2018) and touch on the computer's screen (e.g., Dube & McIlvane, 2001; Okouchi, 2007, 2015) or presses on the mouse button (e.g., Kestner, Romano, St. Peter, & Mesches, 2018; Lacerda, Suarez, & Costa, 2017) are often recorded as responses. Commonly, research software allows the investigation of the increase in the ratio required to produce a consequence and contingent reinforcement loss, two out of three response-cost procedures previously described. However, such software cannot directly require levels of physical effort (the first response-cost procedure) on responding. Therefore, aiming to study the effects of physical effort on human behavior, we build a spring button that can be used as a response button and allows the experimenter to manipulate levels of physical effort required to respond.

We want to thank Dr. Jônatas Lacerda, and Ailton Vaccare for helping us to build the Spring Button described here. Corresponding author: André Luiz: caecosta@uel.br

DOI: 10.17605/OSF.IO/DQK2V

THE SPRING BUTTON

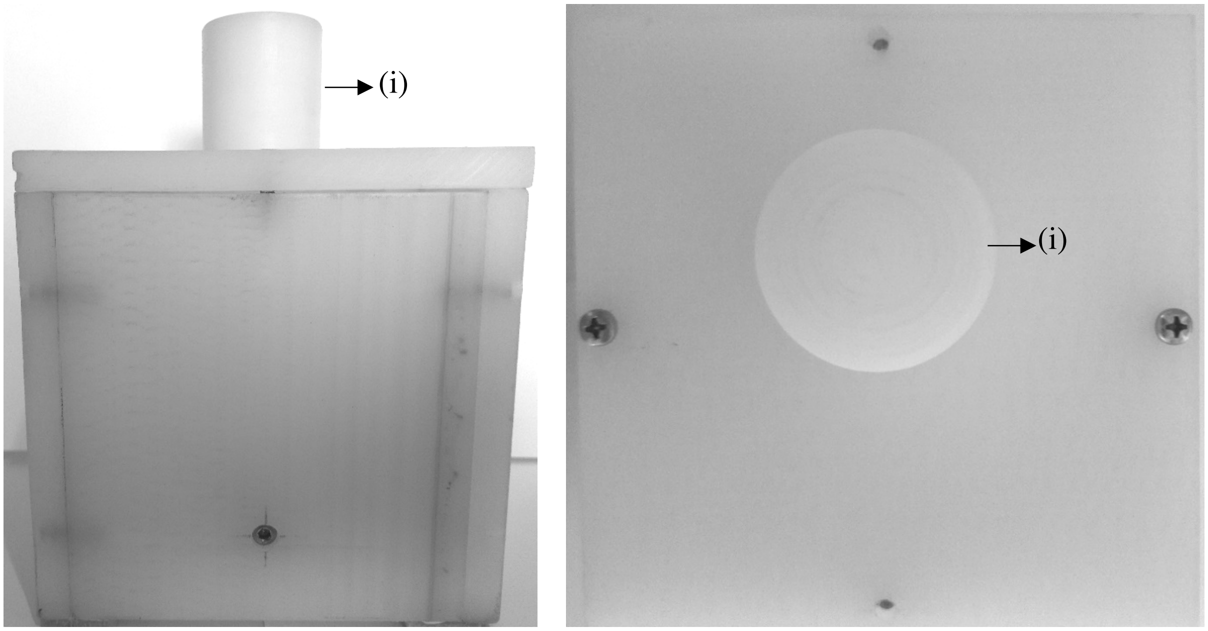
The Spring Button (Figure 1) consists of an 13 cm (height) X 13 cm (length) X 13 cm (wide) nylon box. Nevertheless, the button's material does not need to be nylon; wood or acrylic, for example, can replace it. At the inner bottom of the equipment, a mouse's circuit board with an optical system and a USB connection was fixed using Velcro tape (Figure 2).

At the top of the equipment, a cylinder with a diameter of 3.53 cm, that when pressed 3.5 cm down activated the microswitch, which, in turn, started the mouse's circuit board. We used a microswitch of 15A with three terminals (but only two terminals were used[1]).

[1] We used the Normally Open (NO) and the Common (COM) terminals. These terminals were connected to the left mouse button connection in the mouse's circuit board using two 0.5 mm flexible copper wires. The yellow wire connected the COM terminal to the first left mouse button connection, and the black wire connected the NO terminal to the second left mouse button connection.

Figure 1. *Spring button side view on Panel A and top view on Panel B. (i) cylinder.*

A B



C D

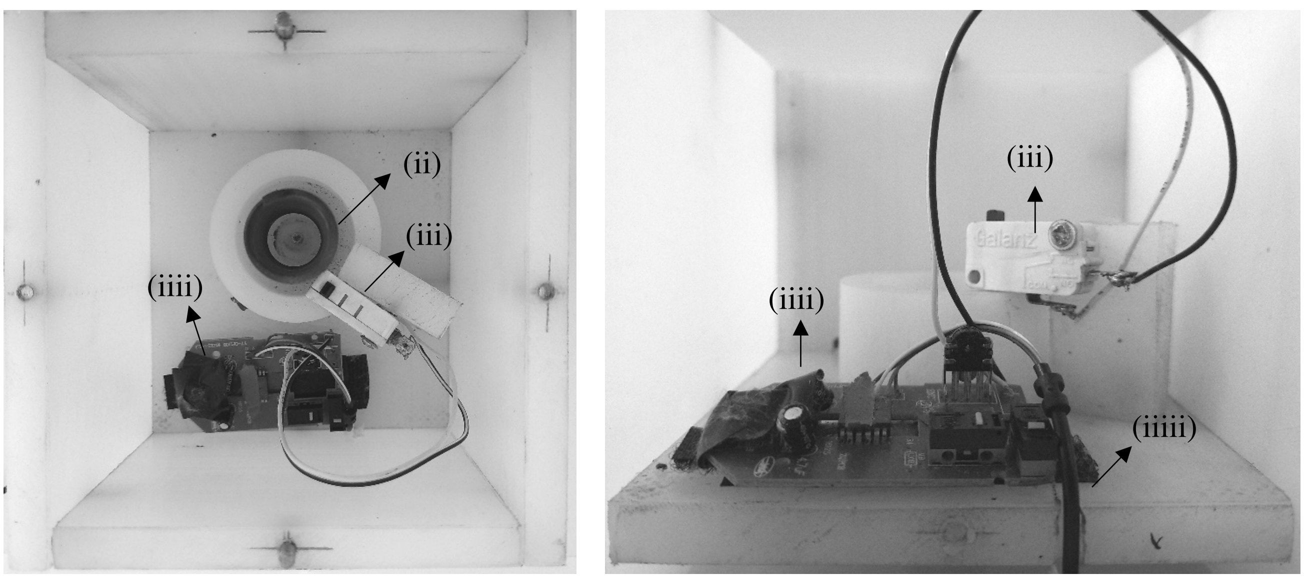


Figure 2. Spring Button interior top view on the Panel C and interior side view on the Panel D. (ii) steel spring (the spring is removable); (iii) microswitch; (iiii) mouse's circuit board; (iiiii) USB cable.

A steel spring placed between the cylinder and the mouse's circuit board required different levels of physical effort to press the button. The physical effort requirements imposed by the spring were measured according to Hooke's Law (Aranha et al., 2016). Different levels of physical effort can be required using other springs with different stiffness (e.g., 30, 50, 90, and 110 N, see Figure 3).

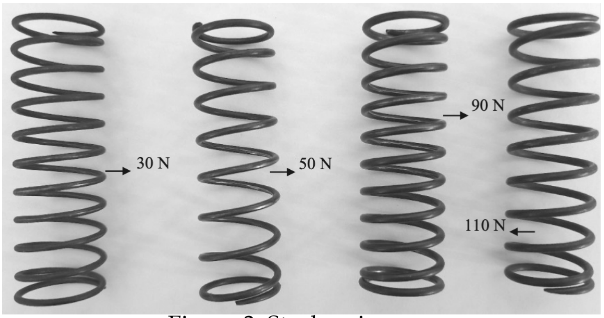


Figure 3. Steel springs.

A USB connection cable connected the button to the computer; thus, the computer records pressures on the Spring Button as pressures on the left mouse button, and the spring directly imposes a physical force requirement on responding (Figure 4). More than one Spring Button can be used simultaneously in computers with more than one USB connection (see Luiz et al., 2020 for an example of it).

USING THE SPRING BUTTON

In our laboratory, Luiz (2017) used two Spring Buttons to examine the effects of two levels of physical effort on resistance to change in humans. Each button served as the response button for one component of a two-component multiple schedule of reinforcement. Below we present the data from the last four Baseline (BL) 30-min sessions of two participants of Luiz's experiment. We chose the sessions used by Luiz to calculate response-rate stability. His BL consisted of a two-component multiple variable interval (VI) VI schedule with equal reinforcement rates. In one component, the physical effort required was 10 N for both participants (Low-Effort Component), and in the other, the physical effort required was 50 N for Participant 1 and 30 N for Participant 2 (High-Effort Component).

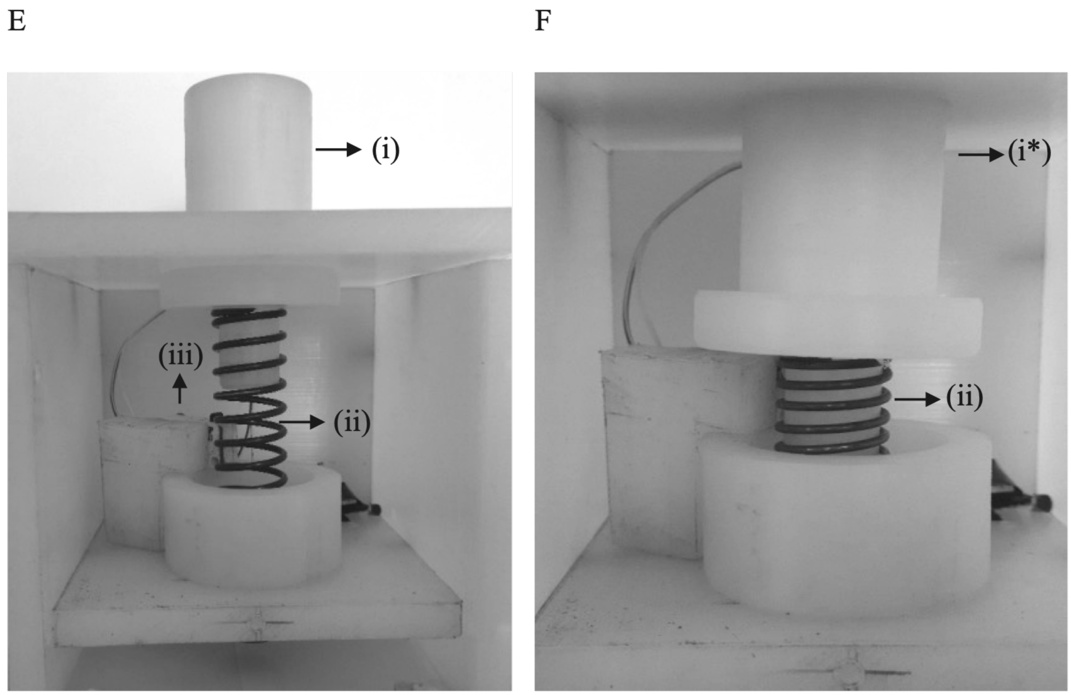


Figure 4. Spring Button's interior side view on Panel E and interior side view with the cylinder pressed down on Panel F. (i) cylinder; (i*) cylinder pressed down; (ii) steel spring; (iii) microswitch.

Figure 5 shows the response rates (responses per min) in the Low-Effort and the High-Effort Components for P1 and P2 during the last four BL sessions of Luiz's (2017) experiment. For both participants, response rates were always higher in the Low-Effort Component.

Figure 6 shows the proportional differences between the response rate during Low-Effort and the High-Effort Components for Participants 1 and 2. We obtained these proportional differences by dividing the response rate during Low-Effort Component by the response rate during the High-Effort Component in each session for both participants. Thus, data above or below 1.0 indicates higher response rates in the Low-Effort and High-Effort Component, respectively. In addition, response rates were always higher during the Low-Effort Component, and the 50 N vs. 10 N produced a greater difference between the response rates than the 30 N vs. 10 N.

Corroborating experiments with non-humans (e.g., Alling & Poling, 1995; Chung, 1965) data from Figure 5 show that the greater the physical-force requirement, the lower is the response rate. Additionally, Figure 6 suggests that the greater the difference between two physical-force requirements, the greater the proportional differences between response rates. These data indicate that the spring button is a viable alternative for researchers aiming to study the effects of physical effort on human responding.

[2] Figures 5 and 6 were made for the present manuscript based on data from Luiz (2017).

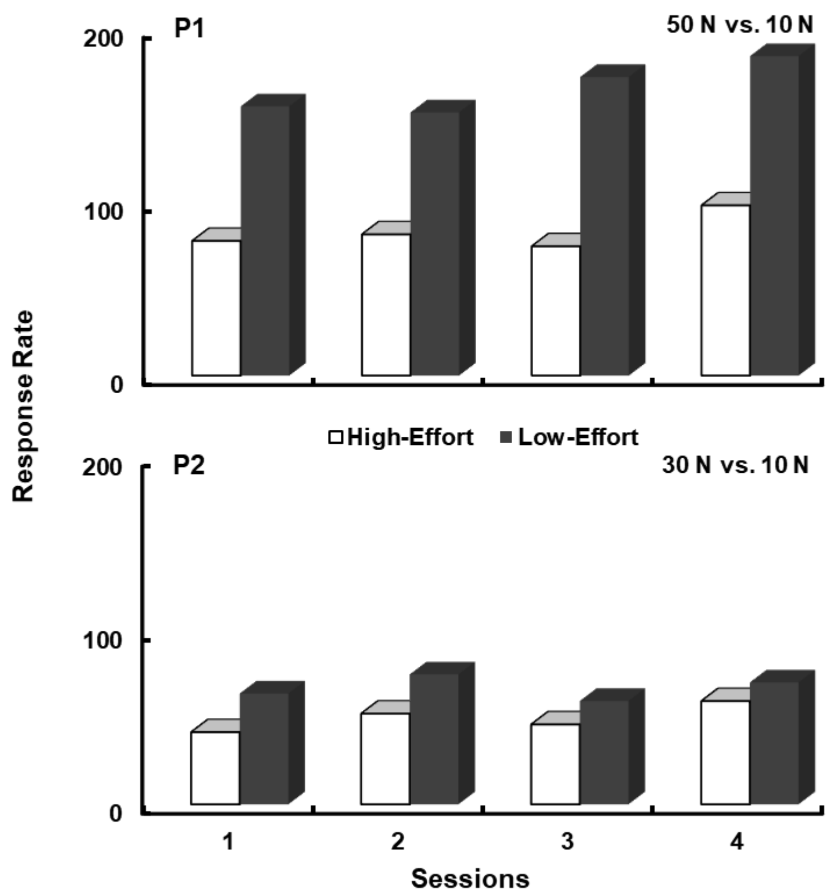


Figure 5. Response rates (responses per min) in the Low-Effort (white bars) and the High-Effort (gray bars) Components for P1 and P2 during the last four BL sessions of Luiz's (2017) experiment.

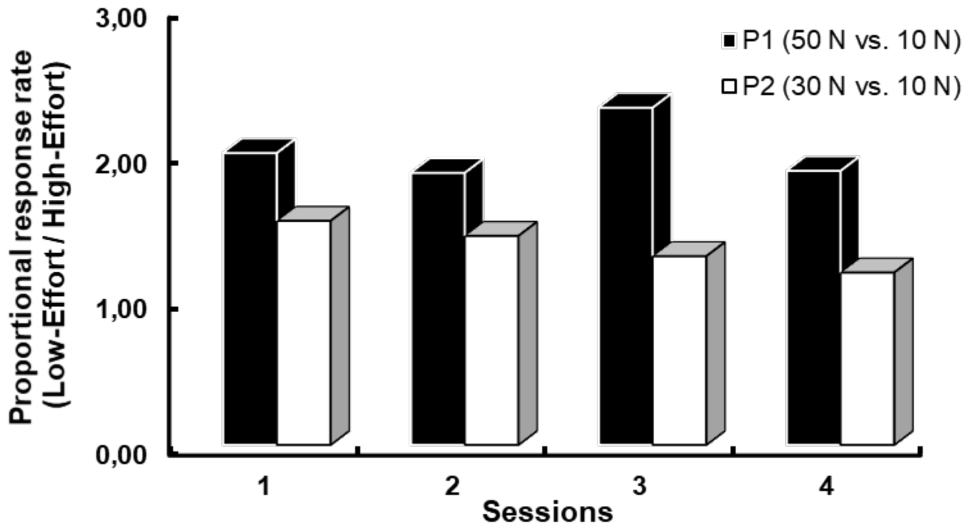


Figure 6. Proportional differences between the response rate during Low-Effort and the High-Effort Components for Participants 1 (black bars) and 2 (white bars) from Luiz's (2017) experiment. Data above or below 1.0 indicates higher response rates in the Low-Effort and High-Effort Component, respectively.

REFERENCES

Alling, K., & Poling, A. (1995). The effects of differing response-force requirements on fixed-ratio responding of rats. *Journal of the Experimental Analysis of Behavior*, *63*(3), 331–346. https://doi.org/10.1901/jeab.1995.63-331

Aranha, N., De Oliveira, J. M., Bellio, L. O., & Bonventi, W. (2016). A lei de Hooke e as molas não-lineares, um estudo de caso. *Revista Brasileira de Ensino de Fisica*, *38*(4). https://doi.org/10.1590/1806-9126-RBEF-2016-0102

Becker, R. M. (2011). *ProgRef V4: Um software para coleta de dados em programas de reforço com humanos* [Universidade Estadual de Londrina]. http://www.bibliotecadigital.uel.br/document/?code=vtls000167971

Bolívar, H. A., & Dallery, J. (2020). Effects of response cost magnitude on resurgence of human operant behavior. *Behavioural Processes*, *178*, 104187. https://doi.org/10.1016/j.beproc.2020.104187

Cabello, F., Barnes-Holmes, D., O'Hora, D., & Stewart, I. (2002). Technical paper using visual basic in the experimental analysis of human behavior : a brief introduction. *Experimental Analysis of Human Behavior Bulletin*, *20*, 18–21.

Cabello, F., Barnes-Holmes, D., O'Hora, D., & Stewart, I. (2003). Technical paper computerized voice production and recognition using visual basic. *Experimental Analysis of Human Behavior Bulletin*, *21*, 30–34.

Capehart, J., Viney, W., & Hulicka, I. M. (1958). The effect of effort upon extinction. *Journal of Comparative and Physiological Psychology*, *51*(4), 505–507. https://doi.org/10.1037/h0045753

Chung, S.-H. (1965). Effects of effort on response rate. *Journal of the Experimental Analysis of Behavior*, *8*(1), 1–8. https://doi.org/10.1901/jeab.1965.8-1

Costa, C. E., & Banaco, R. A. (2002). ProgRef v3: sistema computadorizado para coleta de dados sobre programas de reforço com humanos – recursos básicos. *Revista Brasileira de Terapia Comportamental e Cognitiva*, *4*(2). https://doi.org/10.31505/rbtcc.v4i2.116

Cunha, T. R., Cordeiro, P. R., & Costa, C. E. (2018). Extensão da história em FR e o desempenho subsequente em FI com custo. *Suma Psicológica*, *25*(1), 41–50. https://doi.org/10.14349/sumapsi.2018.v25.n1.5

Dube, W. W., & McIlvane, W. J. (2001). Behavioral momentum in computer-presented discriminations in individuals with severe mental retardation. *Journal of the Experimental Analysis of Behavior*, *75*(1), 15–23. https://doi.org/10.1901/jeab.2001.75-15

Johnson, N., & Viney, W. (1970). Resistance to extinction as a function of effort. *Journal of Comparative and Physiological Psychology*, *71*(1), 171–174. https://doi.org/10.1037/h0028980

Kestner, K. M., Romano, L. M., St. Peter, C. C., & Mesches, G. A. (2018). Resurgence Following Response Cost in a Human-Operant Procedure. *The Psychological Record*, *68*(1), 81–87. https://doi.org/10.1007/s40732-018-0270-7

Kissi, A., Hughes, S., De Schryver, M., De Houwer, J., & Crombez, G. (2018). Examining the Moderating Impact of Plys and Tracks on the Insensitivity Effect: a Preliminary Investigation. *The Psychological Record*, *68*(4), 431–440. https://doi.org/10.1007/s40732-018-0286-z

Larcerda, R. F. F., Suarez, C. J. & Costa, C. E. (2017) Distribuição dos intervalos e taxa de reforço em múltiplo VI VI e resistência à extinção. *Acta Comportamentalia, 25*, 139-155.

Lattal, K. A., & Perone, M. (1998). The Experimental Analysis of Human Operant Behavior. In *Handbook of Research Methods in Human Operant Behavior* (pp. 3–14). Springer US. https://doi.org/10.1007/978-1-4899-1947-2_1

Luce, S. C., Christian, W. P., Lipsker, L. E., & Hall, R. V. (1981). Response cost: A case for specificity. *The Behavior Analyst*, *4*(1), 75–80. https://doi.org/10.1007/BF03391855

Luiz, A. (2017). *Resistência do comportamento à mudança em função de diferentes forças exigidas para a ocorrência da resposta* [Universidade Estadual de Londrina]. http://www.bibliotecadigital.uel.br/document/?code=vtls000220705

Luiz, A., Costa, C. E., dos Santos, J. R., & Tsutsumi, M. M. A. (2020). Resistance to change as function of different physical-effort requirements in humans. *Behavioural Processes*, *176*(February), 104123. https://doi.org/10.1016/j.beproc.2020.104123

Miller, L. K. (1968). Escape from an effortful situation. *Journal of the Experimental Analysis of Behavior*, *11*(5), 619–627. https://doi.org/http://dx.doi.org/10.1901/jeab.1968.11-619

Okouchi, H. (2007). An exploration of remote history effects in humans. *Psychological Record*, *57*(2), 241–263. https://doi.org/10.1007/BF03395575

Okouchi, H. (2015). Resurgence of Two-Response Sequences Punished By Point-Loss Response Cost in Humans. *Mexican Journal of Behavior Analysis*, *41*(2), 137–154. https://doi.org/10.5514/rmac.v41.i2.63744

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pietras, C. J., & Hackenberg, T. D. (2005). Response-cost punishment via token loss with pigeons. *Behavioural Processes*, *69*(3), 343–356. https://doi.org/10.1016/j.beproc.2005.02.026

Powell, R. W. (1968). The effect of small sequential changes in fixed-ratio size upon the post-reinforcement pause. *Journal of the Experimental Analysis of Behavior*, *11*(5), 589–593. https://doi.org/10.1901/jeab.1968.11-589

Premack, D., Schaeffer, R. W., & Hundt, A. (1964). Reinforcement of drinking by running: effect of fixed ratio and reinforcement time. *Journal of the Experimental Analysis of Behavior*, *7*(1), 91–96. https://doi.org/10.1901/jeab.1964.7-91

Quartermain, D. (1965). Effect of Effort on Resistance to Extinction of the Bar-pressing Response. *Quarterly Journal of Experimental Psychology*, *17*(1), 63–64. https://doi.org/10.1080/17470216508416409

Roche, B., & Dymond, S. (2003). Computer-interfacing for dummies: Interfacing peripheral devices to your Macintosh PC. *Experimental Analysis of Human Behavior Bulletin*, *21*, 8–11.

Ruiz, J. A., & Bermúdez, K. (2018). Nota técnica: diseño y conducción de experimentos conductuales mediante el uso de scratch. *Revista Mexicana de Análisis de La Conducta*, *44*(2), 152–172. https://doi.org/10.5514/rmac.v44.i2.68539

Skinner, B. F., & Morse, W. H. (1958). Fixed-interval reinforcement of running in a wheel. *Journal of the Experimental Analysis of Behavior*, *1*(4), 371–379. https://doi.org/10.1901/jeab.1958.1-371

Soares, P. G., Costa, C. E., Aló, R. M., Luiz, A., & Cunha, T. R. de L. (2017). Custo da resposta: Como tem sido definido e estudado? *Perspectivas Em Análise Do Comportamento*, *8*(2), 258–268. https://doi.org/10.18761/pac.2017.020

Solomon, R. L. (1948). Effort and extinction rate: a confirmation. *Journal of Comparative and Physiological Psychology*, *41*(2), 93–101. https://doi.org/10.1037/h0057127

Weiner, H. (1962). Some effects of response cost upon human operant behavior. *Journal of the Experimental Analysis of Behavior*, *5*(2), 201–208.

https://doi.org/10.1901/jeab.1962.5-201

Weiner, H. (1966). Preference and Switching under Ratio Contingencies with Humans. *Psychological Reports*, *18*(1), 239–246. https://doi.org/10.2466/pr0.1966.18.1.239

Weiner, H. (1969). Controlling human fixed-interval performance. *Journal of the Experimental Analysis of Behavior*, *12*(3), 349–373. https://doi.org/10.1901/jeab.1969.12-349

Winograd, E. (1965). Escape behavior under different fixed ratios and shock intensities. *Journal of the Experimental Analysis of Behavior*, *8*(2), 117–124. https://doi.org/10.1901/jeab.1965.8-117

ANALYSIS OF PREFERENCE FOR TOKEN ACCUMULATION IN HUMANS: TWO NOVEL DEMONSTRATIONS

John Michael Falligant, Joelle Krantz, Jonathan D. Schmidt

KENNEDY KRIEGER INSTITUTE & JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE

Although a variety of basic research has examined variables that affect token accumulation in token-reinforcement contexts, there is relatively little translational research in this area. Through two brief demonstrations, the purpose of the current study was to a) replicate basic findings which suggest token accumulation decreases as a function of increasing token-production schedules and b) examine how preferences for accumulated token exchange-production schedules are influenced by interactive effects of psychotropic medications and classes of stimuli used as backup reinforcers. Apart from extending basic token research findings to applied contexts, these two translational demonstrations may serve as a proof of concept for future applied token accumulation research.

Keywords: accumulation; atypical antipsychotic; exchange-production schedules; preference; tokens

There are three components of token-reinforcement procedures that may influence organisms' preferences for (and the efficacy of) various token arrangements. These components involve the token-production schedule, the token-exchange schedule, and the token exchange-production schedule. The token-production schedule specifies the number of responses required to earn a token. For example, under a fixed-ratio (FR) 5 token-production schedule, one token would be delivered following every five responses. The token-exchange schedule specifies the schedule by which tokens are exchanged for backup reinforcers (Hackenberg, 2009); in other words, these schedules specify how much the token is worth. An FR-1 token-exchange schedule, for example, would specify that each token is exchangeable for one unit of the backup reinforcer; a FR-5 token-exchange schedule would specify that each token is worth five units of the backup reinforcer (e.g., Falligant & Kornman, 2019). Finally, the token exchange-production schedule specifies the number of tokens that must be earned before they can be exchanged for backup reinforcers (e.g., DeLeon et al., 2014). For example, under an FR-10 exchange-production schedule (i.e., an accumulated schedule), tokens cannot be exchanged for backup reinforcers until the individual has accumulated 10 tokens. In contrast under a FR-1 exchange-production schedule (i.e., a distributed schedule), each token can be exchanged as soon as it is earned.

Given that exchange-production schedules may affect the magnitude, duration, or continuity of reinforcer access, as well as relative work requirements and commensurate delays to reinforcement in token-reinforcement contexts (Hackenberg, 2009), these schedules are the focus of much interest in basic research contexts (e.g., Bullock & Hackenberg, 2006). Recently, applied researchers have also studied different parameters of exchange-production schedules in clinical contexts. For example, DeLeon et al. (2014) demonstrated that, among a sample of four individuals with intellectual and developmental disabilities (IDDs), accumulated exchange-production schedules were preferred relative to distributed exchange-production schedules when tokens were exchanged

Correspondence author: John Michael Falligant, Kennedy Krieger Institute, 707 N. Broadway, Baltimore, MD 21205. Email: Falligant@kennedykrieger.org.

DOI: 10.17605/OSF.IO/CMT4P

for activity-based reinforcers. Additionally, all four participants' rates of work completion were considerably faster in accumulated schedules relative to distributed schedules when they earned tokens that were exchanged for activity-based reinforcers.

A variety of research has examined factors that influence preferences for accumulated exchange-production schedules and token accumulation in basic research preparations with nonhuman organisms (e.g., Yankelevitz et al., 2008). However, comparatively little research in these areas has been conducted in translational and applied contexts with humans. Identification of factors that affect token accumulation and preferences for larger exchange-production schedules in research contexts with humans has both scientific and clinical value, allowing researchers to a) further explicate variables that affect "self-control" (i.e., preference for delayed, denser schedules of reinforcement relative to more immediate, leaner schedules of reinforcement), and b) identify conditions in which token accumulation is more or less likely to occur. To the extent that clinicians can promote token accumulation in applied situations, clients contact greater periods of reinforcement and learn important self-control skills. Though relatively unexplored in applied preparations, two contextual variables that may affect token accumulation involve a) differences in token-production schedules, and b) interactive effects between psychotropic medications and affinity for classes of stimuli used as backup reinforcers.

Recently, Glodowski et al. (2019) compared token and tandem schedules of reinforcement on response patterns with adolescents with autism. Their results were partially consistent with basic findings suggesting that tokens may suppress responding (relative to tandem schedules) under increasing token-production schedule values (e.g., Bullock & Hackenberg, 2015; Gadaire et al., 2019). Relatedly, Yankelevitz et al. (2008) found that token accumulation may decrease as a function of increasing token-production schedule values, and accumulation may be enhanced in token reinforcement (relative to tandem schedules of reinforcement) conditions. Thus, it is unknown if a) accumulation is diminished under leaner token-production values, b) differences in accumulation under token and tandem schedules occur, and c) whether such differences may be more likely to occur under relatively dense token-production schedules.

In addition, preliminary research indicates that token accumulation may vary based on the type of available backup reinforcers (i.e., edible vs. activity). That is, for some individuals, accumulated schedules may be preferred for activity-based backup reinforcers but not for edible reinforcers (DeLeon et al., 2014). However, it is unknown how other clinical variables, including use of psychotropic medication, may also affect preferences for different exchange-production schedules. The impact of medication on schedule preferences is worth exploring given a) the widespread use of psychotropic medication (in particular, antipsychotic medication) for individuals with IDD and disruptive behavior, and b) the effects of atypical antipsychotics on relevant establishing operations (i.e., increased appetite, insulin insensitivity; Parsons et al., 2009) that may affect the value of edible reinforcers. Thus, changes in the administration of these agents may affect the reinforcing value of edible stimuli and produce concomitant changes in preferences for exchange-production schedules.

Together, the purpose of the current study was to examine whether token accumulation decreases as a function of increasing token-production schedules (e.g., Yankelevitz, 2008; Demonstration 1), as well as replicate results from DeLeon et al. (2014) and parametrically evaluate the effects of dosage changes of aripiprazole on exchange-production schedule preferences (Demonstration 2). Though these are preliminary investigations, these two demonstrations may serve as a proof of concept to build upon for future token accumulation research.

METHOD

Participants and Setting

Nick was a 12-year-old male diagnosed with autism spectrum disorder (ASD) admitted to an inpatient hospital unit for the assessment and treatment of aggression and disruptive behavior. James was an eight-year-old male diagnosed with high-functioning ASD admitted to an outpatient clinic for assessment and treatment of aggression and disruptive behavior. Results from a functional analysis indicated Nick's problem behavior was maintained by social attention and escape from demands, and James' problem behavior was maintained by escape from demands. Both participants communicated vocally using full sentences and had completed token training as part of behavioral treatment for severe problem behavior (data available from corresponding author).

Sessions were conducted in clinic rooms (approximately 8m x 8m) two days per week in the afternoon for approximately 45 to 90 min per day (allowing for multiple sessions per day). Rooms contained two chairs, a desk, and relevant session materials. The therapist used an erasable marker or pencil to provide tokens (tallies) on the token board (laminated sheet of paper or piece of blank paper). Academic materials (i.e., addition and subtraction worksheets) were obtained from participants' existing educational programs.

Nick was prescribed various doses of aripiprazole as part of an ongoing clinical medication trial ranging from 7 mg to 17.5 mg per day. Note that neither the timing of aripiprazole administration (morning or evening), nor the proximity of mealtime to the administration of aripiprazole affects metabolism of the drug (e.g., Davie et al., 2004). Sessions were not conducted until a minimum of four days had passed following each medication increase (see Davies et al., 2004 for a review of aripiprazole pharmacokinetics).

Response Measurement and Interobserver Agreement

Paper-and-pencil data collection was used to record the frequency of completion within 10-s intervals across sessions. Frequency data for work completion (i.e., each academic problem) were recorded and converted to rate (responses per min) for each session. Completion was defined as any instance of the participant finishing the academic task within 30 s of initiating the demand (independently or with a vocal-model prompt) in the absence of problem behavior. For Nick, we also collected frequency data for selections for the accumulated, distributed, and control schedules within the modified concurrent-chains preference assessment.

Interobserver agreement (IOA) was calculated on an interval-by-interval basis for the token evaluation and on a trial-by-trial basis for the token accumulation and modified concurrent-chain preference assessment. An agreement was defined as both observers recording the same response during each interval or trial. Interobserver agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements then converting this fraction to a percentage by multiplying by 100. IOA was collected on 37% of sessions with Nick and averaged 99.8% (range, 97%-100%); IOA was collected on 39% of sessions with James and averaged 99.3% (range, 80%-100%).

DEMONSTRATION 1

Token Accumulation Assessment

Token condition. In this condition, James earned tokens for competing mastered academic demands (e.g., math worksheet problems, spelling problems) each session. Following a correct response (either independently or following a model prompt if the participant made an initial error), the experimenter delivered tokens according to the specified token-production schedule for each academic problem completed. Earned tokens were placed in front of James in a clear container. Each token was worth one small edible or 30-s access to an activity-based reinforcer. Backup reinforcers were identified based on results of previously conducted stimulus preference assessments and other clinical data; they were selected at each exchange opportunity. At the start of each session, the therapist detailed the contingencies to James, and said "It's time to do some work. You can do as many of these problems as you want. Let me know when you're done working." Sessions were terminated after James emitted a communicative response terminating the session (e.g., "I'm done") or 1 min elapsed in the absence of completion of an academic task. The participant would exchange the tokens by placing them in the therapist's outreached hand. All problem behavior was ignored. Sessions were conducted for the following token-production components: FR 1, FR 2, FR 5, VR 2, VR 5. These were conducted in six-session blocks, in which token condition sessions alternated with tandem condition sessions (see below) on a quasi-random basis within each block (see Table 1).

Tandem condition. These sessions were identical to the token condition sessions, except the therapist did not deliver tokens for completion of academic tasks—instead, the therapist tracked the number of tasks that were completed, and delivered the commensurate number of backup reinforcers at the end of session. At the start of each session, the therapist detailed the contingencies to James, and said "It's time to do some work. You can do as many of these problems as you want. I will keep track of the problems you complete and tell you how much you have earned at the end. Let me know when you're done working."

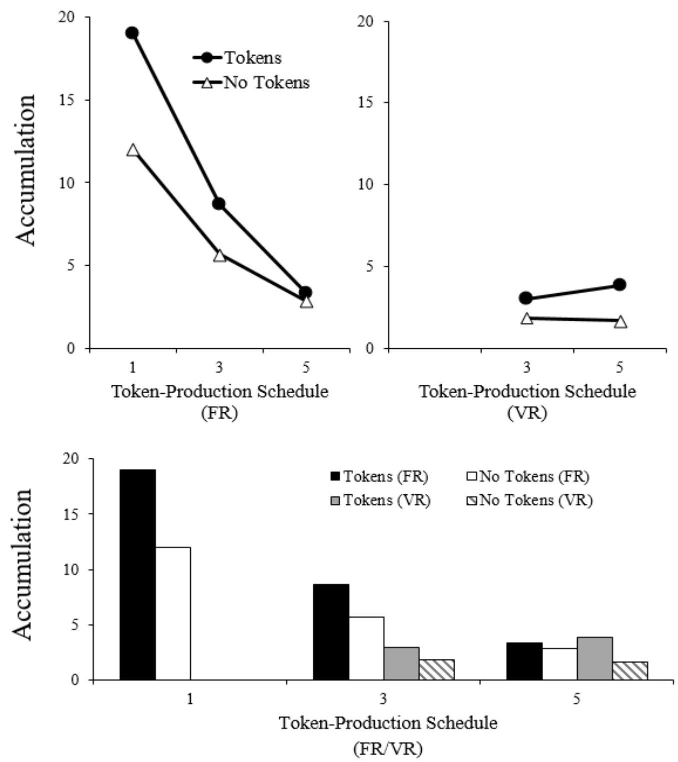
RESULTS AND DISCUSSION

Under FR token-production schedules, James consistently accumulated more reinforcers in the token condition relative to the tandem condition; across both conditions, his mean token accumulation varied inversely with the token-production schedule (Figure 1). A similar pattern emerged under VR token-production schedules, although decreases in accumulation were more pronounced in VR 3 and VR 5 components relative to FR 3 and FR 5 components. Together these

Table 1.
Programmed and obtained schedule values.

Condition	Token	Tandem	Obtained VR Value
FR1	3	3	-
FR3	3	3	-
FR5	3	3	-
FR1	3	3	-
VR3	3	3	3.1
VR5	3	3	5.08
FR1	3	3	-
FR3	3	3	-
FR5	3	3	-
FR1	3	3	-
VR3	3	3	3.3
VR5	3	3	5

Figure 1.
Token accumulation across schedule values.



results replicate Yankelevitz et al. (2008), indicating that mean token accumulation decreases as a function of increasing token-production schedules. However, similar to Yankelevitz et al., mean accumulation was greater in the token condition relative to the no-token (i.e., tandem) condition. Mean differences in responding between the token and tandem schedules primarily occurred under dense (i.e., FR 1, FR 3) token-production schedules; there were minimal differences in reinforcer accumulation between the token and tandem schedules under leaner (FR 5) schedules (cf. Glodowski et al., 2019). Though it would be premature to draw conclusions for clinical practice from this demonstration, this preparation may serve as a useful proof of concept for future research and replications in this area. Additional research might also evaluate the demand elasticity of tokens earned under different token-production and schedule arrangements (FR vs VR, token vs tandem; e.g., Argueta et al., 2019) to identify inelastic areas of demand for tokens or backup reinforcers in order to maximize work-reinforcer ratios.

DEMONSTRATION 2

Token Evaluation and Concurrent-Chains Preference Assessment

Procedures for this evaluation were modeled from those described by DeLeon et al. (2014). Briefly, we used a within-subject ABAB reversal with embedded multielement design followed by a modified concurrent-chains preference assessment in which Nick selected the exchange-production schedule for each session. At the beginning of each choice trial, the therapist stated, “It’s time to do some work. Which way would you like to work and earn tallies?" Following the selection, the therapist implemented the corresponding condition as described below. Each schedule condition was signaled with a vocal instruction and a laminated sheet of paper (21 cm x 28 cm) placed in front of Nick. Tokens were always exchanged for edible backup reinforcers. Nick would exchange the tokens by placing the token board in the therapist's outreached hand. Backup reinforcers were identified based on results of previously conducted stimulus preference assessments and other clinical data; they were selected at each exchange opportunity.

The effects of aripiprazole on Nick's preference for accumulated and distributed token exchange-production schedules were evaluated using a quasi-experimental parametric approach. That is, the concurrent-chains preference assessment was conducted at three different points during the course of multiple scheduled medication adjustments (in which his daily aripiprazole dosage was increased from 7.5 mg to 15 mg to 17.5 mg) over the course of a 21-day period. These medication changes were made by Nick's psychiatrist in the course of ongoing medical services.

Control. The control condition was signaled by a picture of an "X" on Nick's desk. These sessions served as the baseline phase in the token evaluation. Prior to the start of each session, the therapist placed the token board in front of the participant and stated, "It's time to do some work. You can do these problems if you want, but you will not earn any tokens." The therapist then placed an academic worksheet in front of Nick. The therapist delivered neutral praise (e.g., "good") for each problem Nick completed. If Nick completed a problem incorrectly, the therapist provided a vocal-model prompt (e.g., "12 plus 12 equals 24"). If the participant answered correctly following the prompt, the therapist scored the response; if Nick answered incorrectly following the prompt, the therapist did not score the response as complete and prompted Nick to complete the next problem. Sessions ended after either (a) 10 min expired, (b) 1 min elapsed without completing any work, or (c) Nick complied with 10 demands (whichever occurred first).

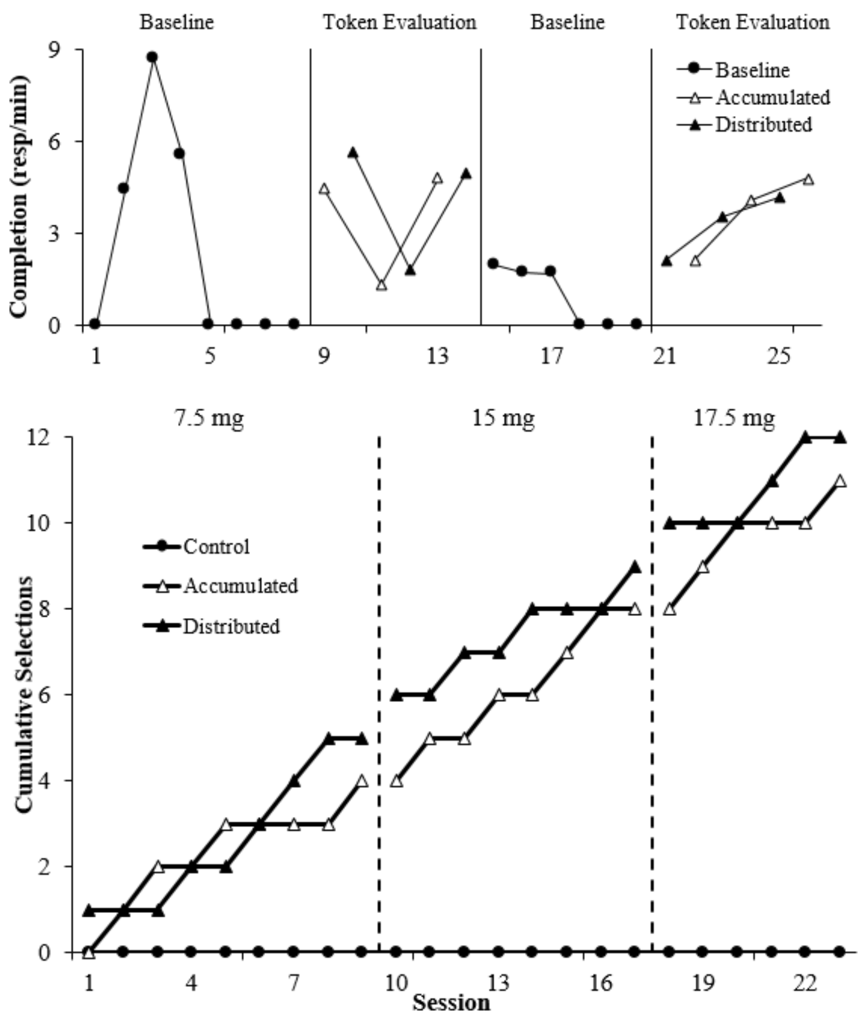
Distributed. The distributed condition was signaled by a picture of a single coin. Prior to the start of session, the therapist placed the token board in front of Nick and reviewed the token exchange-production procedure (e.g., "When you complete a problem, you will get one tally right away to trade for one small piece of snack"). For each problem that Nick completed under this schedule, the therapist delivered a tally on the token board and neutral praise by saying "you earned a token." If Nick completed a problem incorrectly, the therapist provided a vocal-model prompt. If he answered correctly following the prompt, the therapist scored the response as complete and delivered a token; if Nick answered incorrectly following the prompt, the therapist did not score the response as complete or deliver a token, and prompted Nick to complete the next problem. As soon as the response requirement was met (1 token), the therapist paused the session timer and provided one small edible. The timer resumed once Nick consumed the edible.

Accumulated. The accumulated condition was signaled by a picture of a stack of coins on Nick's desk. Prior to the start of session, the therapist placed the token board in front of Nick and reviewed the token exchange-production procedure (e.g., "When you complete a problem you will get one tally, and you can trade all of your tallies in after you have finished working"). For each problem that Nick completed, the therapist delivered a tally on the token board and neutral praise by saying "you earned a token." If Nick completed a problem incorrectly, the therapist provided a vocal-model prompt. If the participant answered correctly following the prompt, the therapist scored the response as complete and delivered a token; if Nick answered incorrectly following the prompt, the therapist did not score the response as complete or deliver a token, and prompted him to complete the next problem. As soon as the response requirement was met (10 tokens), the session timer stopped and Nick selected his 10 edibles to consume.

RESULTS AND DISCUSSION

Rates of work completion are displayed in Figure 2 across baseline and token-evaluation sessions. During initial baseline sessions, Nick's rates of work completion were variable (see Glodowski et al., 2019) but generally very low and stabilized at 0 for multiple consecutive sessions (*M* = 2.3); rates increased in the subsequent accumulated (*M* = 3.5) and distributed (*M* = 4.1) token evaluation condition sessions. Rates of work completion decreased in the return to baseline (*M* = 0.9) before increasing again in the following accumulated (*M* = 3.7) and distributed (*M* = 3.3) token

Figure 2.
Efficacy of and preference for exchange-production schedule arrangements.



evaluation condition sessions. Nick's cumulative selections for accumulated and distributed exchange-production schedules during the modified concurrent-chains preference assessment are displayed in Figure 2. Nick selected the distributed exchange-production schedule on 5 of 9 (7.5 mg), 4 of 8 (15 mg), and 3 of 6 (17.5 mg) sessions, indicating a relative indifference between accumulated and distributed exchange-production schedules across medication dosages

Similar to two participants from DeLeon et al. (2014), accumulated schedules were not associated with increased work completion relative to distributed schedules with edible-based backup reinforcers. Moreover, there was not a strong preference for one schedule over the other. Interestingly, relative preferences for accumulated and distributed schedules did not vary despite two separate increases in Nick's aripiprazole dosage. This outcome supports the hypothesis that accumulated schedules may be preferable to the extent that they enhance continuity of reinforcer access to activity-based stimuli but not necessarily other stimuli for which continuity of access is less important (i.e., food). These outcomes *may* indicate that preferences for exchange-production schedules are stable and may remain fairly consistent across changes in different organismic states (e.g., changes in satiety or hunger). To the degree these findings are replicated in future research, these results could suggest that aripiprazole does not necessarily alter the reinforcing or appetitive value of tokens earned under different exchange-production schedules. However, given very small sample size and fact that there is no comparison to non-food reinforcers (in addition to other weaknesses; e.g., lack of reversals of medication dosages), it would be premature to comment on the extent to which preferences for these exchange-production schedules is affected by the type of backup reinforcer (i.e., food-based) and drug-related changes. Regardless, the methods utilized within this demonstration suggest how preferences for token accumulation via exchange-production schedules across different medication changes may be evaluated using a similar approach in future research.

REFERENCES

Argueta, T., Leon, Y., & Brewer, A. (2019). Exchange schedules in token economies: A preliminary investigation of second-order schedule effects. *Behavioral Interventions*, *34*(2), 280-292. https://doi.org/10.1002/bin.1661

Bullock, C. E., & Hackenberg, T. D. (2006). Second-order schedules of token reinforcement with pigeons: Implications for unit price. *Journal of the Experimental Analysis of Behavior*, *85*, 95-106. https://doi.org/10.1901/jeab.2006.116-04

Bullock, C. E., & Hackenberg, T. D. (2015). The several roles of stimuli in token reinforcement. *Journal of the Experimental Analysis of Behavior*, *103*, 269-287. https://doi.org/10.1002/jeab.117

Davies, M. A., Sheffler, D. J., & Roth, B. L. (2004). Aripiprazole: A novel atypical antipsychotic drug with a uniquely robust pharmacology. *CNS Drug Deviews*, *10*, 317-336. https://doi.org/10.1111/j.1527-3458.2004.tb00030.x l

DeLeon, I. G., Chase, J. A., Frank-Crawford, M. A., Carreau-Webster, A. B., Triggs, M. M., Bullock, C. E., & Jennett, H. K. (2014). Distributed and accumulated reinforcement arrangements: Evaluations of efficacy and preference. *Journal of Applied Behavior Analysis*, *47*, 293-313. https://doi.org/10.1002/jaba.116

Falligant, J. M., & Kornman, P. T. (2019). Preferences for accumulated and distributed token exchange-production schedules: A behavior-economic analysis. *Behavior Analysis: Research and*

Practice. 19(4), 373–378. https://doi.org/10.1037/bar0000159

Gadaire, D. M., Senn, L., Albert, K. M., Robinson, T. P., Passage, M., Shaham, Y., & Topmouglu, B. Differential effects of token production and exchange on responding of children with developmental disabilities. *Learning and Motivation*. Advance online publication. https://doi.org/10.1016/j.lmot.2020.101694

Glodowski, K. R., Bourret, J., Ivy, J. W., Seaver, J. P., Jackson, J., Stine, J. M., & Nuzzolilli, A. E. (2020). A comparison of token and tandem schedules of reinforcement on response patterns for adolescents with autism. *Behavioral Interventions*, *35*(1), 25-37. https://doi.org/10.1002/bin.1704

Hackenberg, T. D. (2009). Token reinforcement: A review and analysis. *Journal of the Experimental Analysis of Behavior*, *91*, 257-286. https://doi.org/10.1901/jeab.2009.91-257

Parsons, B., Allison, D. B., Loebel, A., Williams, K., Giller, E., Romano, S., & Siu, C. (2009). Weight effects associated with antipsychotics: A comprehensive database analysis. *Schizophrenia Research*, *110*, 103-110. https://doi.org/10.1016/j.schres.2008.09.025

Yankelevitz, R. L., Bullock, C. E., & Hackenberg, T. D. (2008). Reinforcer accumulation in a token-reinforcement context with pigeons. *Journal of the Experimental Analysis of Behavior*, *90*, 283-299. https://doi.org/10.1901/jeab.2008.90-283

AN ANALOGUE ASSESSMENT OF CAREGIVER ESCAPE FROM INFANT CRYING BASED ON RATE OF SHAKING

Benjamin N. Witts
ST. CLOUD STATE UNIVERSITY

Jennifer L. Bruzek
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

For many parents, dealing with infant crying is just one of the many tasks that accompanies child rearing. In many cases, infant crying is relatively easy to alleviate. Because infant crying provokes caregiver responding, it is easy enough to conceptualize this act as a stimulus that evokes caregiver reaction. As a stimulus, crying varies in its characteristics (e.g., decibel, pitch, duration) between infants and throughout infancy for any given child.

The variation in cry characteristics seen over the course of infancy can be understood through research on the normal crying curve. In the normal crying curve—which is documented cross-culturally and develops regardless of factors like pre-term birth—infant crying tends to develop an n-shaped pattern of cry "intensity" (e.g., frequency, episode length; see Barr, 1990 for review). In other words, total daily crying duration rises steadily and can increase more sharply beginning around 6 weeks, peaking at about 8 weeks, thereafter decreasing until 4 months of age, at which point the daily total remains fairly steady. It is around the time of the 8-week peak that colicky crying, considered to be on the extreme end of the crying spectrum, is most often reported (see Barr, 2001 for review).

This phenomenon, colicky crying, deserves a bit more attention here, even though it is not the main thrust of this paper. Colicky crying is a relatively extreme stimulus presentation that, undoubtedly, will produce different caregiver responses than less extreme stimulus presentations (e.g., Fujiwara et al., 2011). Colicky crying can be difficult to define, with consensus being elusive over the decades. For our purposes, colicky crying is defined as crying that is long in duration, is seemingly immune to intervention, cannot be otherwise explained by a medical condition (see also Barr, 1993; St James-Roberts et al., 1996), and occurs most notably in children under 5 months of age (see Mayo Clinic Staff, 2014) usually during early evening hours (Barr et al., 2006).

Colicky crying is present in most cultures, with an annual prevalence between 5-19% of infants (Lucassen et al., 2001), and can have a profound impact on child and family wellbeing. For example, colicky infants have a higher prevalence of developmental delays at 6 months (Sloman et al., 1990) and 5 years of age (Rao et al., 2004). Colicky crying might be correlated with postpartum depression (Radesky et al., 2013) and distress in the family (Fujiwara et al., 2011; Maxted et al., 2005; Megel et al., 2011), among other concerns.

At its worst, colicky crying is correlated with child abuse and abusive head trauma (formerly shaken baby syndrome; Adamsbaum et al., 2010; Barr, 2014; Lee et al., 2007; Levitzky & Cooper, 2000; Reijneveld et al., 2004), which can lead to death (see also Barr et al., 2006, Lee et al., 2007). In a detailed account, Adamsbaum et al. (2010) analyzed 112 cases of abusive head

Benjamin N. Witts, Department of Community Psychology, Counseling, and Family Therapy, St. Cloud State University; Jennifer L. Bruzek, Department of Curriculum & Instruction, The University of Alabama in Huntsville.

Benjamin N. Witts ORCID: https://orcid.org/0000-0002-5554-3147. Jennifer Bruzek ORCID: https://orcid.org/0000-0001-7604-6492.

Correspondence concerning this article should be addressed to Benjamin N. Witts, EB-B210, St. Cloud State University, 720th Avenue, St. Cloud, MN 56301-4498 (email: bnwitts@stcloudstate.edu).

This work was supported by a St. Cloud State University Mid-Career Faculty Grant.

DOI 10.17605/OSF.IO/JXZAF

trauma that were referred to French courts for investigation between January 2002 and May 2009. In their report, 29 cases involved a confession of the abuse, and these cases consisted of 7 girls and 22 boys between the ages of 1 month and 7 years (M_{age} = 8 months; 27 under 12 months). Adamsbaum et al. revealed two important themes. First, all abusers reported shaking (much of which was more intense than would occur with normal rocking or bouncing). This shaking was relatively long lasting (e.g., several min) and repetitive. The repetition of the act suggests that, at least for some, shaking was more than a one-time impulsive reaction to an aversive stimulus. Second, for many, the shaking had been effective at stopping the infant's crying by inducing sleep[1], which suggests the maintaining variable was cry termination (i.e., escape).

Given the documented prevalence of abusive head trauma involving shaking, and not other behavior such as kicking or striking the infant with an object, it makes sense that some behavioral process is likely involved with the selection or induction of intense shaking during the abusive episode(s). Of course, survivorship bias could overrepresent these cases in the literature, and other behaviors like variations of singing, pleading, switching caregiver responsibilities with a partner, or walking away could be selected or induced through the same process(es). Regardless, the fact remains that there is likely some operant relation between crying and shaking that needs to be studied, even if not all caregivers a) engage in shaking as a response to infant crying and b) produce intense shaking under various conditions (e.g., extinction, cf. Alessandri & Lattal, 2021). As has been demonstrated with non-shaking behavior, parental behavior is modifiable by perceived changes in infant behavior (e.g., Gewirtz & Boyd, 1977). Though this brief overview says nothing of the conditions that might select caregiver responding and those that induce a single abusive episode (what might be labeled as "impulsive").

The logical place for researchers to find answers to the questions about the processes involved in abusive head trauma in infants is in ecological analyses of infant-caregiver interactions. These ecological studies would give equal attention to both infant and caregiver behavior, highlighting the reciprocal control of infant crying and caregiver responding under more natural conditions (see Etzel & Gewirtz, 1967 and Gewirtz, 1976 for examples of and arguments for reciprocal control).

However, these ecological studies will be costly and time-consuming. Preliminary analogue work can aid in preparing for ecological analyses by helping to identify potentially important variables and serve to test the recording devices and determine variables of interest. One possible process that could be of interest to ecological assessment is the role of negative reinforcement on refining and altering already-selected caregiver responses. The natural development of intense shaking could take several weeks or months, and any device used to study this development would need to capture both large and subtle changes to shake intensity.

Our working hypothesis is that variations in shake intensity might be negatively reinforcing for the infant as well as the adult (save for cry termination due to sleep induction). With repeated shakings, habituation or similar processes might dull the infant's reaction. Over longer periods, repeated shake-cry-termination interactions in which greater and greater shaking is required to achieve the same break from crying might enter into a shaping procedure. Interestingly, as would be expected from behavioral variability stemming from extinction during prolonged inconsolable crying episodes, formerly successful shaking might take the form of an alternative response. For example, throwing as an extension of the shaking form, where the sudden stop upon impact would produce head trauma, would fit the prediction well (see Adamsbaum et al., 2010 for examples of throwing after a shaking episode). At best,

[1] As Adamsbaum et al. (2010) noted, this is likely due to traumas to the infant's brain.

the literature suggests a link between infant crying and abusive shaking, possibly because some history of shaking—even of milder forms—tended to be effective. Alternatively, shaking might be selected, as other acts (e.g., fetching an object to strike the infant) allow more time for competing responses and competing stimulus control to develop.

To date, no study has examined cry termination as a negative reinforcer for shake variations, let alone subtle changes captured over time. However, several studies have established cry termination as a negative reinforcer for the selection of different caregiving responses. In these works, classes of responses, not qualities of responses within a class, were selected for study. For example, Bruzek et al. (2009) and Thompson et al. (2011) both used a simulated caregiving preparation to establish that caregiving behavior was sensitive to recorded infant cry termination as a negative reinforcer. Additionally, Chen et al. (2019) studied individuals in 60 couples either in the hospital laboratory (n = 6) or in their homes (n = 54) when their newborns were 2-4 months old with a Realityworks RealCare Baby II-*plus*. In brief, when these parents' attempts to calm the simulated infant failed, they were more likely to try an alternative response (47 of 59 participants) compared to others whose attempts succeeded (32 of 61 participants). Finally, in an unpublished master's thesis, Tye (2014) replicated and extended Thompson et al. (2011) across two studies. In Tye's first study, three participants interacted with a Realityworks Real Care Baby II-*plus* infant simulator. They were assessed on the frequency of simulator mishandling (shaking, rough handling, or failure to support the head) and cumulative time spent rocking or feeding. Feeding and rocking came under the control of escape and avoidance of crying. During extinction (feeding for all sessions), feeding eventually ceased for two participants, with relatively lower rates of feeding for the third, but not low enough to meet criteria for success. At the same time, most extinction sessions revealed cumulative durations of rocking similar to those in the rocking-as-escape from crying conditions. No mishandling was reported.

Left unanswered is whether shake intensity, or what might be called "aggression", can be functionally related to cry termination. Important to the ecological assessment will be establishing whether to look for interrelations between shaking and cry characteristics, particularly when changes to one or the other are subtle. The question thus asked in this study was to what extent we could get participants to alter their rate of shaking (i.e., "aggression") when shaking terminated crying in a simulated infant[2].

METHOD

Participants

Six self-identified female (M_{age} = 22.33, range = 19-32) and one self-identified male (22 years) college students participated. Age, gender, infant care experience, and beliefs about infant crying are outlined in Table 1. The only inclusion criterion was that the participant was at least 18 years old at the time of the study.

Setting

The study was conducted in a small nursery lab in a medium-sized Midwestern university (see supplemental files for photo). The lab was divided into two spaces separated by floor-to-ceiling curtains. One side of the lab housed bookcases, filing cabinets, and a desk with two computers. The other side (8.25 ft x 9 ft) contained nursery items, such as a rocker/glider, dresser, changing table, crib, lamps, and miscellaneous toys, feeding, and diapering materials. During the study, overhead lights were turned off and the two lamps, providing soft, dim light, were turned on.

Materials

Simulated Infant. The infant surrogate (hereafter "infant") was a male Caucasian

[2] See Bechtel et al., 2019 for an overview of arguments for analogue studies in infant caregiving research

Table 1. *Participant Characteristics*

	P1	P2	P3	P4	P5	P6	P7
Age	21	19	20	22	22	20	32
Gender	F	F	F	F	M	F	F
Current or former parent/guardian to a baby?	No	No	No	No	No	Yes	Yes
Experience calming crying infants	Yes	Yes	Yes	Yes	No	Yes	Yes
Estimated experience calming infants (in hours)	10-19	10-19	30-39	10-19	0	50-59	100+
I believe that a baby (newborn to 3 months of age) should never be left to cry	Somewhat untrue	Somewhat true	Somewhat true	Somewhat untrue	Neither true nor untrue	Somewhat true	Somewhat true
Parents and caretakers should do everything in their power to stop a baby from crying	Somewhat untrue	Somewhat untrue	Somewhat true	Neither true nor untrue	Somewhat true	Somewhat untrue	Somewhat true
Sometimes babies (newborn to 3 months of age) will cry for no reason	Somewhat true	Somewhat true	Neither true nor untrue	Very true	Somewhat true	Somewhat untrue	Neither true nor untrue
Sometimes babies (newborn to 3 months of age) will continue to cry if they don't get what they want	Very true	Very true	Somewhat true	Somewhat true	Neither true nor untrue	Very true	Very true

Realityworks RealCare Baby 3 Infant Simulator (a doll) and housed the customized Arduino-based sensors and control modules. The customized hardware was designed to fit inside the RealCare Baby 3's electronics housing compartment. The device controlled the cry type, duration, and communicated with sensors to determine when successful cry termination criteria had been met. Cries were played through a small 8 ohm, 0.5 watt speaker placed inside the doll. Customized software in the device allowed for modifications to the study parameters.

Cry Types. Two cry types were used in this study, and both were taken from audio tracks from internet videos of crying infants. The first cry, the *regular cry*, was a fussy baby cry more typical of normal crying, and the second, *intense cry*, was more paroxysmal (i.e., sudden) and intense. The cries were cut and looped to be as seamless as possible. An acoustical analysis documenting the differences between these cries, confirming the labels *regular cry* and *intense cry*, can be found in Koffi (2022).

Demographics and Exit Survey. Participants were given a demographics form (see Table 1) and an exit survey that asked them to respond to the question, "What do you think this study was about?" and the prompt, "Please note any thoughts, reactions, or feelings you had while completing this study".

Calibration. Prior to the study's start, the first author and the designer of the Arduino module and its related software calibrated the device to ensure accurate data collection and reporting. They calibrated the device by shaking the doll while simultaneously watching data on shake frequency in real-time on a computer monitor. By counting the number of shakes timed against a stopwatch and comparing it to the data output, we confirmed the device was 100% accurate in recording the number of shakes per s across a variety of rates. During piloting, test participants were observed holding the infant up to their chest and shoulder while patting the infant's back; therefore, tests of back-patting were conducted and found that light patting did not register as shaking, but more forceful slapping (which would be considered abusive) did[3]. In other words, some degree of change in distance would need to occur before the device would register the response. The component parts are also independently calibrated prior to production.

Procedure

The procedures were approved by the St. Cloud State University IRB. Participants were recruited by email from an electronic flyer distributed by three professors (each for one course in the same department) at the same university wherein the study was conducted. The flyer indicated the study requirements, location, duration, and days of availability. Due to COVID-19 concerns, sessions were conducted on Tuesday and Friday nights from 5:00 to 8:00 p.m. to allow for proper sanitization of the room and equipment between participants. Due to an error in advertising, participants were promised $100 in Amazon.com gift cards for participation[4]. While the amount might appear coercive, particularly given the task, the decision was made to honor the mistake and award the promised incentive.

Following consent, participants were read a script modeled after Bruzek et al. (2009) and Thompson et al. (2011):

> We are conducting this study to learn how adults will respond in a simulated caregiving situation. Do what comes naturally. After 60 min we will take a short break before resuming again. We ask that you leave your cell phone with us; we will place it in a storage cabinet. Let us know if you are expecting any emergency calls and we will be sure to hand you your phone should anyone call or text—just make sure if you are expecting a call that your phone is not on silent. We will also collect any additional materials you might have, like books or homework. Everything will be returned during the break and again at the end of the study. Please do not remove the infant's clothing, as it could damage the doll[5]. At this time, please ask any questions you have, as after this we will not be able to answer any study-related questions.

No phone calls or other outside distractions occurred during the course of participation.

Design. A latency-based range-bound changing criterion design with baseline and reversals was implemented (see Klein et al., 2017 for review). The range-bound element was selected to demonstrate control over responding, rather than an attempt to permit responding to go to extreme levels as can happen with typical criterion designs. The range-bound criterion was based on a moving per s average rate of responding, and advancement was further limited by an additional criterion of latency to meeting the range-bound criterion. The study was conducted in two phases, each 60 min and consisting of baseline followed by the range-bound changing criterion elements. The two phases were separated with a 10-min break during which participants could leave the lab if they needed to attend to personal needs. With successful movement through the phases, participants were to experience both cry types. Movement through the phases would require several reversals which would also include some phases with relatively high shake rates. However, no participant progressed far enough to take advantage of this counterbalancing measure, and therefore this design defaults to a within-participant analysis of each cry and between participant analysis of the two cries.

Independent and Dependent Variables. *Shake-Rate Criterion***.** Shakes were measured

---

[3] For this study, we were unable to measure in objective terms what force would be needed during the slap.

[4] Originally, participants were to report to the university via Uber at 11:00PM and leave, again via Uber, at 2:00AM, thus requiring $100 to compensate Uber drivers, tips, and participation incentive. With COVID-19, the requirement for an Uber would perhaps be too risky, so an earlier time was selected so there would be less risk of sleep-deprived participants who would need supports to get home safely after the study.

[5] This line was added after P1, who stripped the infant to change its diaper and, in doing so, nearly had the Arduino module break free of the infant (as viewed on the video camera).

first as any acceleration in one direction, with a change in direction on the accelerometer indicating the end of one shake and the start of the next. As an operant response, this change in direction could be accomplished through any behavior (e.g., rocking, bouncing, forcefully slapping the infant's back). Shakes were measured as a moving average across 2-s bins (e.g., seconds 1 and 2 comprise a bin, seconds 2 and 3 comprise a bin, seconds 3 and 4 comprise a bin).

Latency to shake-rate criteria were measured per trial, and a trial started when the infant began crying. Shake-rate criteria were set as ranges of average per-s shaking. Successful completion of a shake-rate criterion required that two out of three consecutive 2-s bin averages were within the criterion range in effect. For example, if the first 2-s bin averaged 1.5 shakes per s, the second averaged 2.5 shakes per s, and the third averaged 1.5 shakes per s, then successful completion would be registered if the criterion was set between 1 and 2 shakes per s averaged. Once the criterion was met, crying terminated for that trial and the inter-trial interval started. For data reporting purposes, success was recorded as the first bin of the successful 3-bin average. In the example above, if the first 2-s bin (the first to achieve 1.5 shakes per s) occurred 42 s after the crying started, then 42 s was recorded as when the criterion was met. The software also recorded the shake-rate criterion as being met for the first of the three bins for the purposes of determining criterion advancement.

Latency-Determined Criterion Advancement. Advancement from one range-bound shake-rate criterion to the next required that a) meeting the shake-rate criterion started within 10 s of the cry's initiation and that b) this requirement was satisfied across three consecutive trials, except as noted. The requirement for three consecutive completions within 10 s was selected to demonstrate control by showing that the shake-rate criterion was produced relatively quickly and consistently, and not by incidentally fulfilling the criterion. In other words, as the criteria changed, the participant's performance also needed to change in a manner that repeatedly demonstrated that when the infant began crying, they quickly defaulted to that particular rate of shaking. Second, while other work in cry termination through negative reinforcement makes use of sustained behavior, the restricted range of shaking rate at each criterion level might prove difficult to maintain over long periods (e.g., 5 consecutive min), and so consecutive repeat performances permitted us to deviate from past preparations.

Baseline. While participants were completing their demographics form and informed consent, the first author or his research assistant—a master's-level student trained to 100% fidelity—plugged in the Arduino's battery packs, inserted the Arduino module into the infant, closed and dressed the infant, and placed the infant in the crib.

For all participants, the infant remained silent for the first 5 min of the study (time measured from the moment the battery packs were plugged in to the Arduino module). When 5 min had elapsed, participants 1-4 experienced the same 12 baseline cries (i.e., inconsolable crying across three volumes) randomly ordered with an Excel random number generator prior to the study. Because Participants 1-4 never experienced the loudest version of the regular cry condition during the study, the remaining 9 quieter cry baseline trials were removed from the analyses (see Table 2 for baseline cry types, durations, and inter-cry intervals presented in this study). Participants 5-7 were provided with the same 6 baseline trials, randomly ordered with an Excel random number generator prior to the study.

The baseline condition, then, contained all elements of the intervention phase except that the contingency between shaking and cry termination was absent. We presented the putative negative reinforcer in a pre-determined random time in an effort to reduce the possibility of extinction-like effects during baseline while providing an appropriate comparison for intervention phases (see, e.g., Thompson & Iwata, 2005). After the 10 min break, the baseline phase was repeated.

Table 2. *Baseline conditions*

Baseline Sequence	Cry Type	Cry Duration	Inter-Cry Interval
Participants 1-4			
3	Regular	54 s	20 s
4	Regular	12 s	20 s
7	Regular	52 s	20 s
Participants 5-7			
1	Colic	12 s	10 s
2	Colic	6 s	30 s
3	Colic	26 s	10 s
4	Colic	34 s	30 s
5	Colic	46 s	10 s
6	Colic	22 s	20 s

Range-Bound Changing Criteria. Following baseline, the range-bound changing criteria were in effect, and advanced according to the advancement criteria described. All participants experienced the cry at either a naturally occurring range of 56.8 to 82.2 dB (Participants 1-4; regular cry; measured 6 inches from infant's chest with a Tenma 72-947) or a range of 79.2 to 91.2dB (Participants 5-7; intense cry). Cry termination occurred when either 120 s of crying elapsed without success or when the range-bound shake-rate criterion was met. Predetermined 10-, 20-, or 30-s intervals of silence were randomly distributed after cry terminations. A 120 s maximum cry bout and 10- to 30-s inter-cry-intervals were programmed based on pilot data suggesting that prolonged inescapable crying might prove to be too aversive and lead to restricted variability in shake rates and unacceptable rates of attrition. While this decision was antithetical to ecological validity, such validity was not the priority of this study and was therefore sacrificed to better understand the relation between escape from crying and shake intensity. Furthermore, while similar studies have extended uninterrupted crying for 30 minutes, our inability in pilot work to generate variability in shaking with similar protocols might be attributed to our operant response (variation within a response) as opposed to the response in other work (variation between responses).

Debriefing. After the second 60-min phase, participants completed the exit survey and then received their compensation. Participants were given the purposes of the study, the manipulations used, and then were thanked for their time and escorted to the building's exit. The debriefing emphasized the analogue nature of the study and the engineered attempts to induce more "aggressive" responding to the infant.

RESULTS

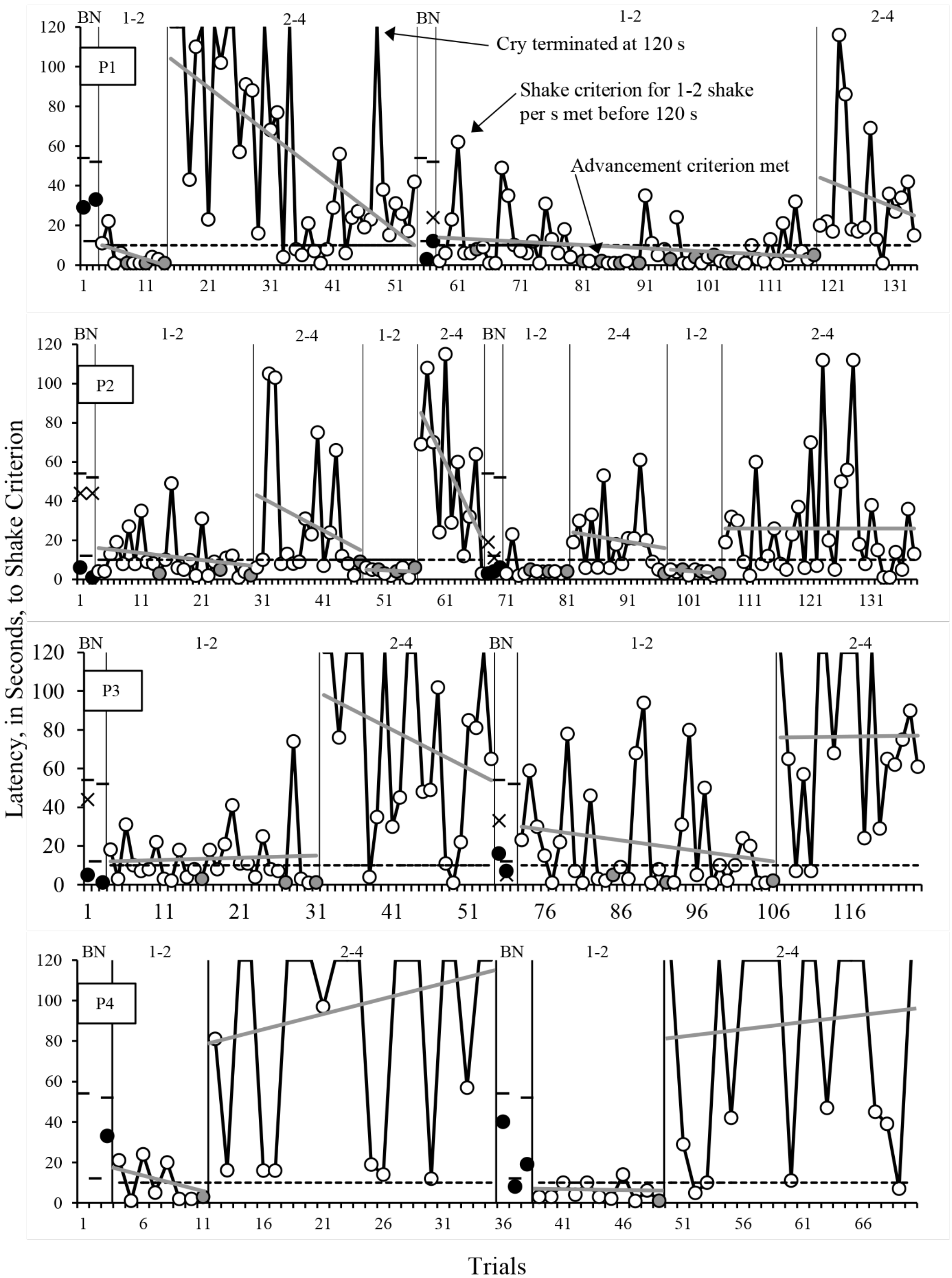
Data are reported as the first and second half of each participant's data, with the first half ending as the 10-min break before the second baseline condition began. P1-P4's data are represented in Figure 1, while P5-7 are represented in Figure 2. P1-P4 all experienced the regular cry. In the first half of the study, P1 met the advancement criterion for 1-2 shakes per s three times, but was unable to meet the criterion advancement of 2-4 shakes per s. Trend line analyses show developing control over shaking as the 2-4 shakes per s requirement progressed. In the second half, P1 was expected to meet the advancement criterion of 1-2 shakes per s 10 times before it changed to 2-4 shakes per s. A less pronounced trend toward criteria was present in the 2-4 shakes per s criteria. P1's baseline showed a decrease in latency to producing 1-2 shakes per s, and showed the emergence of one 2-4 shakes per s baseline trial in the second half.

In both halves of the study, P2 alternated between three successful completions of 1-2 and 2-4 shakes per s advancement criteria. The second 2-4 shakes per s criterion trended toward a demonstration of control, though this trend was absent in the second half. P2's baseline data showed two successful 1-2 and 2-4 shakes per s in the first baseline, with three 1-2 and two 2-4 shakes per s successes in the second baseline, all faster, on average, than in the first baseline.

P3 successfully completed two sets of advancement criteria of 1-2 shakes per s but failed to advance out of the 2-4 shakes per s requirement. In the first half of the experiment, P3's data showed some control over 2-4 shakes per s developing. In the second half, P3 completed one set of advancement criteria for 1-2 shakes per s, though in more time than it took in the first

Figure 1

Latencies to shake criterion by range-bound criteria for P1, P2, P3, and P4



Note. Trend lines provided only when more than one data point for the range-bound criterion exceeds 10 s. Jittered lines in baseline assessments indicate cry termination time. Closed circles represent meeting shake criterion at 1-2 shakes per s; X's for shake criterion at 2-4 shakes per s. No baseline shake criterion at 4-7 shakes per s occurred. Dashed stable line indicates the 10 s threshold to determine criterion advancement. Open circles indicate a successful shake criterion was met. Criterion advancement is marked by gray closed circles. Latency timed out at 120 s.

half. P3 ended the study working to succeed in 2-4 shakes per s with no demonstration of decreasing latencies to cry termination. Baselines tended toward faster completion of 1-2 and 2-4 shakes per s, with one of each criteria in the first baseline and two of each in the second.

P4 successfully met advancement criteria for 1-2 shakes per s in both halves of the experiment, but was unable to meet advancement criteria for 2-4 shakes per s in either half, instead showing longer latencies to cry termination as that criterion progressed. Baseline responding showed an increase from one successful 1-2 shakes per s in the first half to three 1-2 shakes per s and one 2-4 shakes per s in the second half.

P5-P7 all completed the study while interacting with an infant that produced the intense cry. P5 completed two sets of advancement criteria for 1-2 and 2-4 shakes per s in both halves of the experiment. In neither half could P5 complete the 4-7 shakes per s advancement criterion, though in the first half the last three cry terminations were completed at 10, 7, and 15 s, nearly completing the requirements for advancement from the 4-7 shakes per s range. Trend line analyses demonstrated developing control over 4-7 shakes per s in the first half, but growing latencies under the same conditions in the second. Latencies in the second 4-7 shakes per s criterion did not show a trend toward faster completion. During the first baseline, P5 produced four 1-2 and 2-4 shakes per s successes. During the second baseline, only one successful 2-4 shakes per s criterion was met, and it was slower than the 2-4 shakes per s successes in the first baseline.

P6 was unable to produce the performance needed to advance out of the 1-2 shakes per s criteria until near the end of the second half of the experiment. Latencies in the second half of the experiment routinely reached 120 s. Baseline in the first half produced five successful 1-2 shakes per s, all in under 10 s from the start of the cry. In addition, two successful 2-4 shakes per s were produced, one in under 10 s from cry start. No successful baseline performances were recorded in the second half of the experiment for P6.

In both halves of the study, P7 was unable to meet advancement criteria for 2-4 shakes per s, though some evidence of shorter latencies emerged in the second half for this criterion range. The first baseline saw five successful 1-2 shakes per s, four of which were under 10 s from the time of cry onset. In the second baseline, three 1-2 shakes per s criteria met, all under 10 s from cry onset.

In terms of what participants thought the study was about, six made statements regarding learning about how people soothe or respond to a crying infant. Two participants made mention of efforts to teach the infant to self-soothe. When asked about reactions to the study, two participants noted the realistic nature of the scenario, two noted that the second half was easier than the first, and five noted some form of frustration during the study.

Several unplanned anecdotal observations were noted. P4 was observed putting the infant in the crib and walking away to sit in the rocking chair. P4 was also observed, on occasion, sighing loudly during the start of a new crying episode. P5 placed the infant on the floor and told it stories. Three participants were observed singing to the infant while rocking it.

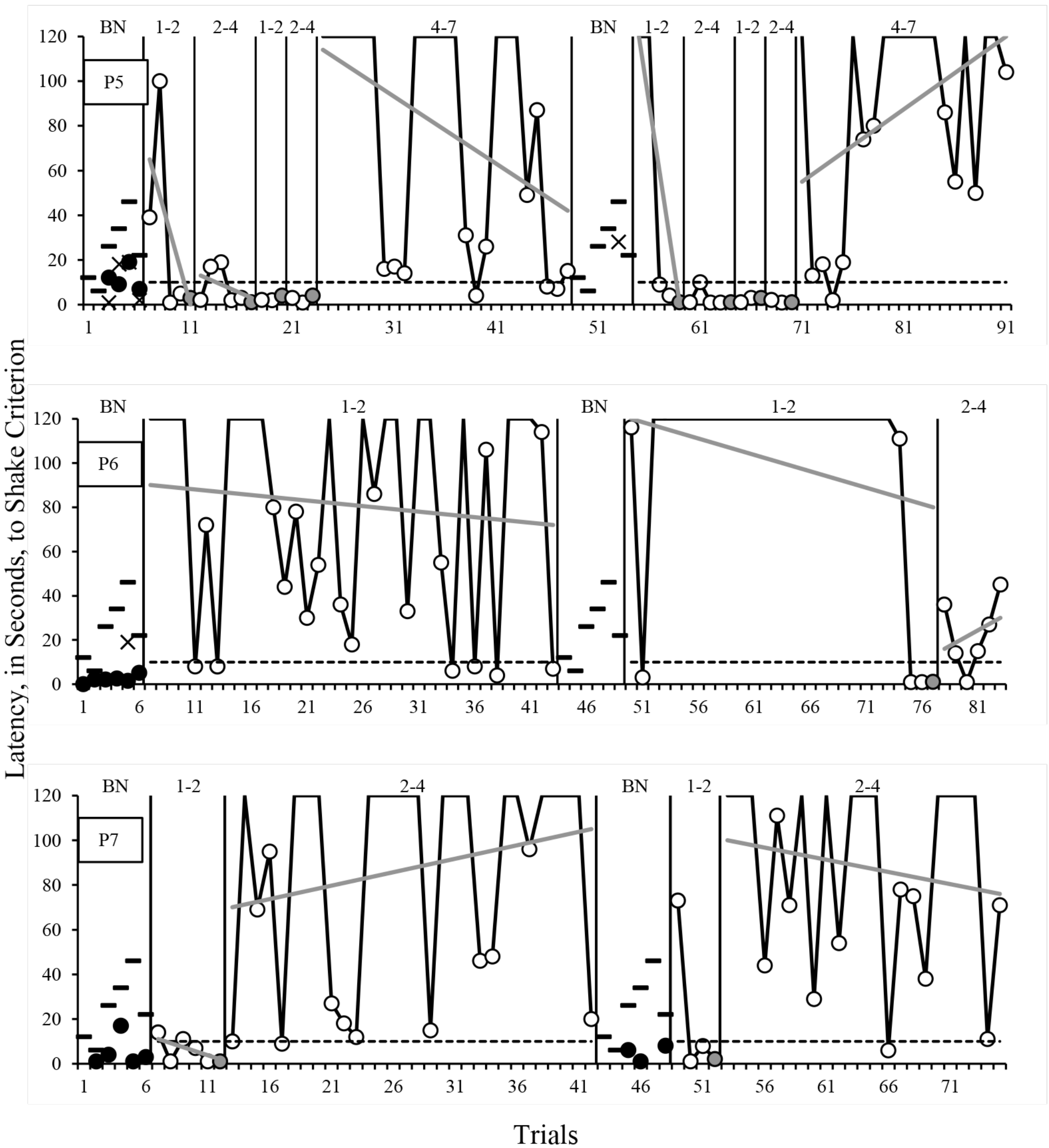
DISCUSSION

The goal of this study was not to demonstrate whether abusive shaking would occur, nor was it to emulate the naturalistic conditions found in caregiver interactions with an infant. Instead, the goal of this study was to demonstrate that rate of shaking with a simulated infant could come under the control of negative reinforcement, which in turn, could help inform ecological observational research. In this light, the study largely succeeded in its aims. Given the paucity of data showing control over intensity, let alone trends toward control, we believe this study can help to inform and justify ecological studies on the topic.

These results provide evidence of control of lower and relatively moderate shake rates, but not necessarily of higher shake rates. These results should not be surprising as variations of shaking are common reactions to infant crying (e.g., rocking, bouncing) while shaking that could cause harm is less

Figure 2

Latencies to shake criterion by range-bound criteria for P5, P6, and P7



likely to occur. The variables responsible for dangerous shaking might be developed over extended periods and involve several variables not accounted for in this study (e.g., hormonal changes, sleep deprivation, societal expectation, parental support options, previous crying patterns, prolonged cry bouts). Further, only some segment of the population might be more prone to abuse, and it is yet unclear if more variability would have been achieved in this study had we sought out this subpopulation. For example, Crouch et al. (2008) and Crouch et al. (2022) found that adults who might be more prone to abuse tended to view infant crying more negatively and feel more hostile toward that crying. Though Crouch et al. (2008) explored the use of a hand dynamometer to test grip strength—believed by some to be a surrogate measure for aggression—during videos of crying infants, differences between subgroups only approached significance. However, other researchers, like Compier-de Block et al. (2015) did find that mothers who had abused their infants used greater force on the handgrip as compared to mothers who had reportedly not abused their infant. Thus, it is plausible that while our data are promising in terms of control, clearer results would be achieved by targeting subpopulations of interest. As a final consideration, as part of a larger study, Crouch et al. (2022) rated 153 parents on their quality of caregiving with an inconsolable simulated infant. In that study, participants first watched a research assistant calm a crying *RealCare Baby-3* through actions like rocking, feeding, touching, and so on. The parent, unlike the assistant, was unable to calm the infant during a 30-minute interaction where the infant cried continually. Those parents identified as being at high risk for child abuse engaged in significantly lower quality interactions. Interestingly, regardless of risk factor, participants trended toward lower quality interactions compared to their initial attempts as the study progress; perhaps this last observation could help explain why our participants were more likely to succeed in calming the infant during the first half their time.

Interesting trends emerged that will need further exploration. In particular, P1, P2, and P3 were all required to complete multiple criterion advancement successes before the shake-rate range altered, whereas P4, P5, P6, and P7 were only required to complete one criterion advancement success before the shake-rate range changed. P1, P2, and P3 showed more variability in their data compared to P4, P5, P6, and P7. At this point, it is unclear if the variations are due to the number of criterion advancement successes or, perhaps, to differences in cry type, with P4 being the outlier in performance for the regular cry group. Similarly, P6 and P7 had the most experience with infant caretaking, and this extra-experimental history might have had some influence over their performance (see also Bruzek et al., 2009, P-7, p. 333). At this point, these data suggest that more history of success with a particular shake rate might help to induce more variability when conditions change.

Several limitations exist. First, while calibration tests were 100% accurate, coding decisions in the software undoubtedly had some influence on the data. For example, P1 produced successful shake-rate criteria for 2-4 shakes per s in 10 s or less across three consecutive cries in the first half of the study, but the program as designed did not count that as being successful in terms of moving to the next shake-rate criteria. The issue was related to the formula used to calculate the latency to average shakes-per-s and reporting whole numbers to the output spreadsheet. For example, the latency might have been rounded down (e.g., rounded from 10.25 s to 10.00 s), but the program itself, which did not round down, noted that the latency was greater than 10 s. However, given the thousands of data points and the relatively few incidences where errors like this were produced, we are confident that the effect on the data was minimal. Regardless, future work should aim for a more sensitive measure or more sensitive data output (e.g., reporting data to one or two decimal places).

Second, there is some evidence that baseline responding might have interfered with performance. For example, during the first baseline, P6 produced shaking that met criteria for 1-2 shakes per s in under 10 s across five consecutive cries. However, after baseline, P6 was unable to reproduce this performance until late into the second half of the study. It is possible that during baseline either this type of responding was extinguished (had it been in their repertoire prior to the study) or that early production of the behavior required negative reinforcement with escape from crying for that behavior to be selected. However, for other participants, baselines served as a redundant measure of control, as other parts of the study were under extinction-like conditions or at least conditions where cry termination was contingent on time. As baseline is used to judge performance of an intervention to some problem (e.g., too much or too little responding), its use during translational studies exploring functional relations for use in future applied work might not be necessary, and in some cases might be detrimental to the data.

Third, as St. James-Roberts et al. (1996) argued, the length of inconsolable crying might be more important than the type of crying (see also St. James-Roberts, 1999; cf. Koffi, 2022), suggesting that for laboratory work more effective control over shake variability might occur if crying is only terminated by participant responding and not the passage of time. Much of this medical literature, however, rests on parental diaries or short segments of recorded cries, and, thus, the validity of their conclusions is in question. Regardless, future research should attempt to manipulate the duration of crying and cry types to identify the extent to which these variables affect caregiver behavior.

Fourth, while shaking was the behavior of interest, it might prove to be a difficult one to study in laboratory conditions. Consider Tye's (2014) first experiment, in which participants continued rocking, arguably on the spectrum of shaking, during extinction conditions but did not exceed acceptable levels to where it developed into mishandling. The fact that none exceeded rocking on the shaking spectrum might suggest that developing variability in shaking can be difficult. In this light, the fact that for most participants shake rate either came under the control of escape from crying or trended toward shorter latencies, despite not advancing to other conditions of the study (e.g., alternating cry types), might actually suggest a strength of the design, at least in its potential to capture these relations.

Two final participant limitations are worth noting. The accidental advertising of $100 in compensation might have influenced participant decisions (e.g., persistence). We did not collect socioeconomic status data, and so we are unsure how influential that compensation might have been. While the compensation was more appropriate for the original study proposed, COVID-19 restrictions required a change in programming and a change in compensation to match. As compensation did not change along with new study parameters, threats from coercion are present. Put bluntly, the compensation we used might be inappropriate for future studies. Furthermore, participant characteristics are homogenous and not necessarily representative of those who are more likely to be directly involved in caregiver activities. However, as this study was designed to spur more work in the area, we anticipate demographic concerns will be addressed in subsequent work.

While research in behavior analysis has shown that different activities can be controlled through cry cessation, this is the first to demonstrate that variations of one particular response can be selected and maintained. Even more impressive is the fact that increased shaking rate is a response most would be resistant to producing, particularly given the reported realism of the analogue. We believe that continued work in this area will yield important returns in the understanding and, we hope eventual treatment of abusive head trauma in infants.

REFERENCES

Adamsbaum, C., Grabar, S, Mejean, N., & Rey-Salmon, C. (2010). Abusive head trauma: Judicial admissions highlight violent and repetitive shaking. *Pediatrics*, *126*(3), 546-555. https://doi.org/10.1542/peds.2009-3647

Alessandri, J., & Lattal, K. A. (2021). Early-extinction effects on a force-response of humans. *Journal of the Experimental Analysis of Behavior, 1-12.* https://doi.org/10.1002/jeab.683

Barr, R. G. (1990). The normal crying curve: What do we really know? *Developmental medicine & child neurology,* *32*(4), 356–362. https://doi.org/10.1111/j.1469-8749.1990.tb16949.x

Barr, R. G. (1993). Normality: A clinically useless concept: The case of infant crying and colic. *Journal of Developmental and Behavioral Pediatrics, 14*(4), 264–270. https://doi.org/10.1097/00004703-199308010-00011

Barr, R. G. (2001). Colic is something infants do, rather than a condition they "have": A developmental approach to crying phenomena, patterns, pacification and (patho)genesis. In R. G. Barr, I. St. James-Roberts, & M. R. Keefe (Eds.) *New evidence on unexplained early infant crying: Its origins, nature and management* (pp. 87-104). Johnson & Johnson.

Barr, R. G. (2014). Crying as a trigger for abusive head trauma: A key to prevention. *Pediatric Radiology, 44*, S559-S564. https://doi.org/10.1007/s00247-014-3100-3

Barr, R. G., Trent, R. B., & Cross J. (2006). Age-related incidence curve of hospitalized shaken baby syndrome cases: convergent evidence for crying as a trigger to shaking. *Child Abuse Neglect, 30*, 7-16. https://doi.org/10.1901/jaba.2011.44-295

Bechtel, K., Bhatnagar, A., & Auerbach, M. (2019). Simulation-based research to improve infant health outcomes: Using the infant simulator to prevent infant shaking. *Infant Behavior and Development, 56*, 1-7. https://doi.org/10.1016/j.infbeh.2018.05.006

Bruzek, J. L., Thompson, R. H., & Peters, L. C. (2009). Resurgence of infant caregiving responses. *Journal of the Experimental Analysis of Behavior*, *92*, 327–343. https://dx.doi.org/10.1901%2Fjeab.2009-92-327

Chen, M., Barr, R. G., Miller, E., Fairbrother, N., Antle, A., & Brant, R. (2019). When soothing succeeds: Simulating a risk for repeated shaking in abusive head trauma in infants. *Infant Behavior and Development, 56*, 1-8. https://doi.org/10.1016/j.infbeh.2018.02.004

Compier-de Block, L. H.C.G., Alink, L. R.A., Reijman, S., Werner, C. D., Maras, A., Rijnberk, C., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2015). Handgrip force of maltreating mothers in reaction to infant signals. *Child Abuse & Neglect, 40*, 124-131. http://dx.doi.org/10.1016/j.chiabu.2014.03.006

Crouch, J. L., Bridgett, D. J., Milner, J. S., Cote, K., Lelakowska, G., Davila, A., McKay, E., & Savoree, S. (2022). *Journal of Interpersonal Violence*, 1-23. https://doi.org/10.1177/08862605221106137

Crouch, J. L., Skowronski, J. J., Milner, J. S., & Harris, B. (2008). Parental responses to infant crying: The influence of child physical abuse risk and hostile priming. *Child Abuse & Neglect, 32*, 702-710. https://doi.org/10.1016/j.chiabu.2007.11.002

Etzel, B. C., & Gewirtz, J. L. (1967). Experimental modification of caretaker-maintained high-rate operant crying in a 6- and 20-week-old infant (*Infans tyrannotearus*): Extinction of crying with reinforcement of eye contact and smiling. *Journal of Experimental Child Psychology, 5*, 303-317. https://psycnet.apa.org/doi/10.1016/0022-0965(67)90058-6

Fujiwara, T., Barr, R. G., Brant, R., & Barr, M. (2011). Infant distress at five weeks of age and caregiver frustration. *The Journal of Pediatrics*, *159*, 425-430. https://doi.ord/10.1016/j.jpeds.2011.02.010

Gewirtz, J. L. (1976). The attachment acquisition process as evidenced in the maternal conditioning of cued infant responding (particularly crying). *Human Development*, *19*, 143-155. https://doi.org/10.1159/000271525

Gewirtz, J. L., & Boyd, E. F. (1977). Experiments on mother-infant interaction underlying mutual attachment acquisition: The infant conditions the mother. In T. Alloway, P. Pilner, & L. Krames (Eds.) *Attachment behavior* (pp. 109-143). Springer.

Klein, L. A., Houlihan, D., Vincent, J. L., & Panahon, C. J. (2017). Best practices in utilizing the changing criterion design. *Behavior Analysis in Practice, 10*, 52-61. https://doi.org/10.1007/s40617-014-0036-x

Koffi, E. (2022). Is intensity (decibel levels) the most aversive correlate in infant cry? Preliminary and exploratory results. *Linguistic Portfolios, 11*, 1-24. Downloaded from https://repository.stcloudstate.edu/stcloud_ling/vol11/iss1/2

Lee, C., Barr, R. G., Catherine, N., & Wicks, A. (2007). Age-related incidence of publicly reported shaken baby syndrome cases: Is crying a trigger for shaking? *Journal of Developmental & Behavioral Pediatrics, 28*, 288-293. https://doi.org/10.1097/dbp.0b013e3180327b55

Levitzky, S., & Cooper, R. (2000). Infant colic syndrome—Maternal fantasies of aggression and infanticide. *Clinical Pediatrics, 39*, 395-400. https://doi.org/10.1177/000992280003900703

Lucassen, P. L. B. J., Assendelft, W. J. J., van Eijk, J. Th. M., Gubbels, J. W., Douwes, A. C., & van Geldrop, W. J. (2001). Systematic review of the occurrence of infantile colic in the community. *Archives of Disease in Childhood, 84*, 398-403. https://dx.doi.org/10.1136%2Fadc.84.5.398

Maxted, A. E., Dickstein, S., Miller-Loncar, C., High, P., Spritz, B., Liu, J., & Lester, B. M. (2005). Infant colic and maternal depression. *Infant Mental Health Journal, 26*, 56-68. https://doi.org/10.1002/imhj.20035

Mayo Clinic Staff (2014, May). Colic. *Diseases and conditions*. Retrieved from http://www.mayoclinic.org/diseases-conditions/colic/basics/definition/con-20019091

Megel. M. E., Wilson, M. E., Bravo, K., McMahon, N., & Towne, A. (2011). Baby lost and found: Mothers' experiences of infants who cry persistently. *Journal of Pediatric Health Care, 25*, 144-152. https://doi.org/10.1016/j.pedhc.2009.10.005

Radesky, J. S., Zuckerman, B., Silverstein, M., Rivara, F. P., Barr, M., Taylor, J. A., Lengua, L. J., & Barr, R. G. (2013). Inconsolable infant crying and maternal postpartum depressive symptoms. *Pediatrics, 131*, e1857-e1864. https://doi.org/10.1542/peds.2012-3316

Rao, M. R., Brenner, R. A., Schisterman, E. F., Vik, T., & Mills, J. L. (2004). Long term cognitive development in children with prolonged crying. *Archives of Disease in Childhood, 89*, 989-992. https://dx.doi.org/10.1136%2Fadc.2003.039198

Reijneveld, S. A., van der Wal, M. R., Brugman, E., Sing, R. A. H., & Verloove-Vanhorick, S. P. (2004). Infant crying and abuse. *The Lancet*, *364*, 1340-1342. https://doi.org/10.1016/s0140-6736(04)17191-2

Sloman, J., Bellinger, D. C., Krentzel, P. (1990). Infantile colic and transient developmental lag in the first year of life. *Child Psychiatry and Human Development, 21*, 25-36. https://doi.org/10.1007/bf00709925

St. James-Roberts, I. (1999). What is distinct about infants' "colic" cries? *Archives of Disease in Childhood, 80*, 56-62. http://dx.doi.org/10.1136/adc.80.1.56

St. James-Roberts, I., Conroy, S., & Wilsher, K. (1996). Bases for maternal perceptions of infant crying and colic behaviour. *Archives of Disease in Childhood, 75*, 375-384. http://dx.doi.org/10.1136/adc.75.5.375

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

Thompson, R. H., & Iwata, B. A. (2005). A review of reinforcement control procedures. *Journal of Applied Behavior Analysis, 38*, 257-278. https://dx.doi.org/10.1901/jaba.2005.176-03

Tye, M. (2014). *Negative reinforcement in infant care simulation: Alternative caregiver responses to prevent child abuse*. [Master's thesis, University of South Florida]. ProQUest Dissertations and Theses Global. https://search.proquest.com/openview/40382fe6510c680e88d27ff2bd09ddde/1?pq-origsite=gscholar&cbl=18750