Spring 5-8-2021

Phase Duration and Resurgence

Sean Smith
University of Nebraska Medical Center

PHASE DURATION AND RESURGENCE

by

Sean Smith

A DISSERTATION

Presented to the Faculty of

the University of Nebraska Graduate College

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

Medical Sciences Interdepartmental Area Graduate Program

(Applied Behavior Analysis)

Under the Supervision of Professor Brian D. Greer

University of Nebraska Medical Center

Omaha, Nebraska

April, 2021

Supervisory Committee:

Kevin C. Luczynski, Ph.D.

Wayne W. Fisher, Ph.D.

Nicole M. Rodriguez, Ph.D.

Bethany A. Hansen, Ph.D.

ACKNOWLEDGEMENTS

"Think where man's glory begins and ends, And say my glory was I had such friends." -William Butler Yeats

Thank you to everyone who has made this possible for me. To Mr. Amatrucola, who piqued my interest in computer science and taught me the foundational skills necessary to build my experimental software. To Ms. Cupillari, who gave me the confidence I needed to never be afraid of math no matter the complexity. To Professor Birky, who taught me the basics of mathematical modeling—who knew it would prove so helpful! To Professor Harbert, who taught me about the functions of music, which has been integral to my understanding of the functions of behavior. To Dr. Vorndran, who opened my mind to the complexities of experimental design. To Dr. Kettering, who taught me to bring those experimental designs into all aspects of my clinical work. To Dr. Page, who is a constant reminder that behavior analysis can and should be fun. To Katie Hurlock, who is a constant reminder that I can pour my heart into the science of behavior, but I should never leave it there for too long.

To my brothers, Matt and Brendan, who always exuded greatness and are unsurpassed role-models. To my sister, Emily, who is the reason I entered this field and has always been there for me when I've needed her most. To my parents, Maria and Stephen, who have been setting me up for success since the day I was born and kept me in school despite my numerous attempts to stop. To my wife, Regina, I never could have achieved this without your love and support—you make me a better person in every way imaginable. Your mother gave you the sage advice, "Be giving," and I am blessed that you embody those words each and every day.

PHASE DURATION AND RESURGENCE

Sean W. Smith, Ph.D.

University of Nebraska, 2021

Supervisor: Brian D. Greer

Resurgence, the recurrence of responding due to a worsening of reinforcement conditions for current behavior, is a prevalent form of treatment relapse. Resurgence as Choice in Context predicts that increasing the duration of exposure to reinforcement for target responding during Phase 1 will increase resurgence magnitude, whereas increasing the duration of exposure to reinforcement for alternative responding and extinction for target responding during Phase 2 will decrease resurgence magnitude. We conducted an experiment evaluating these predictions with human participants recruited through Amazon's Mechanical Turk platform. We varied Phase 1 and Phase 2 durations across four experimental groups. Resurgence as Choice in Context successfully predicted the differences in resurgence magnitude across these groups, and fitting the quantitative model to the obtained data yielded an exceptional coefficient of determination. We discuss the implications of these results for using Resurgence as Choice in Context to inform experiments with human participants and the feasibility of using human-operant preparations to evaluate resurgence.

Keywords: human operant, phase duration, resurgence, translational research

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

FCT	Functional communication training
HIT	Human intelligence task
MTurk	Mechanical Turk
RaC^2	Resurgence as Choice in Context
Serial DRA	Serial differential reinforcement of alternative behavior
TWR	Temporal Weighting Rule
VI	Variable interval

INTRODUCTION

Treatment relapse, or the recurrence of problem behavior following effective treatment, is a highly prevalent phenomenon in clinical settings (Briggs et al., 2018; Falligant et al., 2020; Mitteer et al., under review; Muething, Call, et al., 2020; Muething, Pavlov, et al., 2020). One particular form of treatment relapse is resurgence, which is the recurrence of responding due to a worsening of the reinforcement conditions for current behavior (Lattal et al., 2017). Although differential reinforcement of alternative behavior is highly effective at producing rapid decreases in problem behavior and increases in appropriate alternative responses (Greer et al., 2016; Hagopian et al., 1998; Kurtz et al., 2011), disruptions to the reinforcement contingency for the alternative response, such as delays to reinforcement (e.g., Jarmolowicz & Lattal, 2014), decreases in reinforcer magnitude (e.g., Craig et al., 2017), periods of extinction (e.g., Leitenberg et al., 1970), and reinforcement schedule thinning (e.g., Volkert et al., 2009) may lead to the resurgence of problem behavior. In fact, recent evaluations of the prevalence of resurgence in clinic settings found that resurgence of problem behavior occurred in 76–90% of applications of reinforcement schedule thinning following initially effective treatment with FCT (Briggs et al., 2018; Mitteer et al., under review; Muething, Pavlov, et al., 2020).

To study resurgence, researchers typically employ a three-phase experimental preparation. In Phase 1, a target response is reinforced. In Phase 2, the target response is placed on extinction, and an alternative response is reinforced. In Phase 3, the reinforcement contingency for the alternative response worsens in some way. Typically, researchers place the alternative response on extinction, while the target response remains on extinction. Resurgence occurs if the participant engages in increased levels of target responding during Phase 3 relative to Phase 2. Experimenters use this

preparation because it is analogous to treating problem behavior with FCT. Phase 1 is comparable to a history of reinforcement for problem behavior prior to treatment, Phase 2 approximates the treatment of problem behavior with FCT, and Phase 3 often simulates treatment-integrity degradation (e.g., Volkert et al., 2009) or thinning of the reinforcement schedule maintaining the alternative response (e.g., Briggs et al., 2018; Mitteer et al., under review; Muething, Pavlov, et al., 2020), either of which can lead to the recurrence of problem behavior.

Although several hypotheses seek to explain resurgence, Resurgence as Choice in Context (RaC^2) posits that resurgence is a byproduct of the same basic processes that control choice across concurrent operants (Shahan, Browning, & Nall, 2020; Shahan & Craig, 2017; Greer & Shahan, 2019). The quantitative model proposed by RaC^2 is based on the concatenated matching law (Baum & Rachlin, 1969) and its associated models of stimulus detection (e.g., Davison & Nevin, 1999; Davison & Tustin, 1978). RaC^2 extends these previous models to predict target and alternative responding throughout resurgence evaluations. According to RaC^2, target responding is described by the equations below when alternative reinforcement is available,[1]

$$B_T(\text{on}) = \frac{kV_T}{V_T + d_1 (V_{Alt}) + \frac{1}{A}} \qquad A = a(V_T + V_{Alt}) \qquad d_1 = d_m(1 + e^{-x_{on}}) \qquad (1),$$

where B_T(on) is the absolute rate of target behavior, and k is the asymptotic rate of B_T. V_T and V_{Alt} represent the relative values of the consequences for the target and alternative responses, respectively, as described by the concatenated matching law. A represents arousal, and a is a free parameter that scales the invigorating effects of reinforcement. Further, d_1 serves as a bias term that varies based on x_{on}, which

[1] Different equations describe responding when alternative reinforcement is and is not available. Interested readers should review Shahan, Browning, and Nall (2020) for a more complete account of the quantitative model proposed by RaC^2.

represents the number of sessions experienced with alternative reinforcement available, and d_m, which is a free parameter representing the asymptotic value of the bias term. Importantly, V_T and V_{Alt} also account for how relative reinforcer values change across time via an adapted version of the temporal weighting rule (TWR; see Devenport & Devenport, 1994; Mazur, 1996, for reviews), described by the following equation,

$$w_x = \frac{1/t_x}{\sum_{i=1}^{n} 1/t_i} \quad (2),$$

where w_x is the weight applied to a particular past experience. Experiences are weighted according to the amount of time since that experience occurred, represented by t_x. More recent events are weighted more heavily than temporally distal events. Thus, the TWR modifies the concatenated matching law to make specific, quantitative predictions about how a history of reinforcement is carried through time to affect current behavior.

One important prediction that RaC^2 makes based on its incorporation of the TWR is that the duration of exposure to a reinforcing contingency for target behavior during Phase 1 (i.e., baseline) and for alternative behavior during Phase 2 (i.e., treatment) will affect the amount of resurgence observed in Phase 3 (i.e., when treatment is disrupted). Specifically, RaC^2 predicts that longer exposure to Phase 1 contingencies will cause greater resurgence, and longer exposure to Phase 2 contingencies will cause less resurgence. Figure 1 shows how these two predictions interact across various combinations of long (i.e., 20-session) and short (i.e., 5-session) exposures to Phase 1 and Phase 2 contingencies to produce different amounts of resurgence during a subsequent transition from Phase 2 to Phase 3. This is conceptually important because it relates specifically to the influence of the TWR and how a history of reinforcement carries forward in time to affect current behavior.

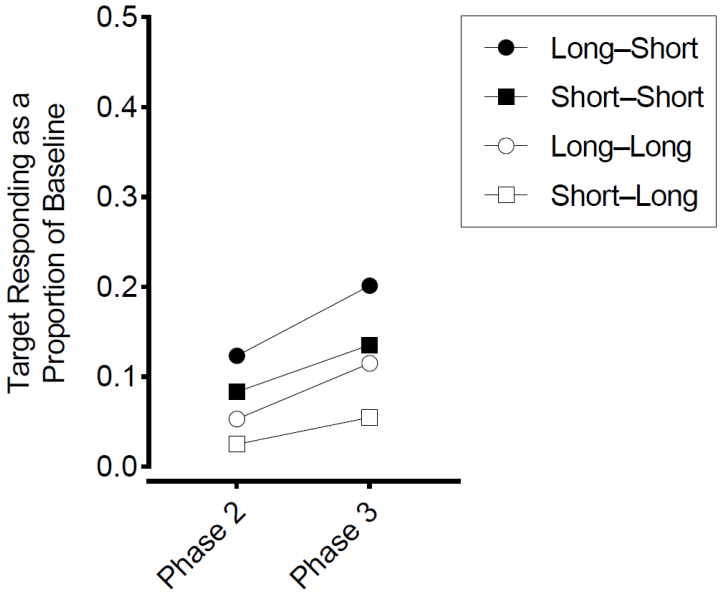

Figure 1: Predicted target responding. Predicted rate of target responding expressed as a proportion of baseline in the last minute of Phase 2 and the first minute of Phase 3 following different combinations of Phase 1 and Phase 2 durations with the following parameter values: k = 15.76, a = .0073, λ = .0033, d_m = .33. In the legend, long indicates a 20-session exposure, and short indicates a 5-session exposure; the first exposure listed indicates Phase 1.

Although no prospective study has explicitly evaluated the predictive validity of RaC[2] for human behavior, its predictions may have important clinical implications (Greer & Shahan, 2019; Shahan & Greer, under review). First, early intervention for problem behavior may be easier to justify if there is empirical support that shorter exposure to the contingencies maintaining problem behavior can significantly reduce the likelihood that the individuals' problem behavior will re-emerge following successful treatment. Second, it may increase clinicians' motivation to expedite the assessment and baseline phases of treatment for problem behavior to prevent any increased probability of resurgence. Third, longer clinical admissions may be easier to justify if longer treatment durations meaningfully decrease the likelihood that treatment relapse will occur following discharge. Taken together, evaluating the effects of these variables may lead to the development of more effective resurgence-mitigation techniques in applied settings.

There are mixed findings regarding the effects of baseline duration on resurgence. Winterbauer et al. (2013) exposed rats to four or 12 sessions of reinforcement for target responding during Phase 1 of a three-phase resurgence evaluation. During the resurgence test, the rats exposed to 12 sessions of Phase 1 engaged in significantly more target responding than those exposed to only four sessions of Phase 1, supporting RaC[2]'s predictions regarding the effects of Phase 1 duration. This result must be interpreted cautiously, however, because the rats experiencing longer Phase 1 durations also engaged in higher rates of target responding at the end of Phase 1. Previous research has demonstrated that higher Phase 1 response rates will result in more resurgence (Fisher et al., 2019; Podlesnik & Shahan, 2009), so it is possible the differences in Phase 1 response rates caused the differences in resurgence rather than Phase 1 duration *per se*.

Bruzek et al. (2009) conducted a translational experiment which demonstrated that caregiving responses with a longer history of reinforcement resulted in a greater magnitude of resurgence. In this study, two different caregiving responses were negatively reinforced by the cessation of simulated crying from a baby doll for two different Phase 1 durations. For five of the eight participants, the caregiving response with a longer history of reinforcement in Phase 1 occurred for longer durations during the resurgence test. Further, the caregiving response with a shorter history of reinforcement in Phase 1 never resurged more than the response with a longer history of reinforcement for any participant. However, these results must be interpreted cautiously because the longer Phase 1 exposure always preceded the shorter Phase 1 exposure, so it remains unclear whether the greater resurgence magnitude was related to the order in which responses were trained and primacy effects (e.g., Lambert et al., 2017) rather than the different durations of exposure to Phase 1 contingencies.

Although these studies provide preliminary evidence that baseline duration can affect resurgence magnitude, Lambert et al. (2020) provided evidence that Phase 1 duration might not affect resurgence. Lambert et al. (2020) had three participants experience Phase 1 contingencies in one component of a multiple schedule for over 400 sessions across 11 months, but in the second multiple schedule component, the participants experienced Phase 1 for a maximum of four sessions within a single day. Next, the experimenters implemented serial differential reinforcement of alternative behavior (serial DRA) with each participant. During a subsequent resurgence test, the target response associated with the longer Phase 1 duration recurred in more sessions for one participant but not for the other two participants. There are several reasons why this study may not have shown a consistent effect of Phase 1 duration. First, all participants experienced a serial-DRA intervention, which may have mitigated resurgence (see Fuhrman et al., 2021 for a brief review), limiting the ability to detect the effects of Phase 1 duration. Further, the participants may not have discriminated across the components of the multiple schedule, so it remains unclear whether the two components represented two separable reinforcement histories or a single, combined reinforcement history when interpreting resurgence within participant.

There are also mixed findings regarding the effects of Phase 2 duration on resurgence. For example, when Winterbauer et al. (2013) evaluated resurgence with rats following four, 12, or 36 sessions of Phase 2, the differences in resurgence were not significant. Similarly, Nall et al. (2018) failed to demonstrate significant differences in resurgence following either five or 20 sessions of Phase 2. On the other hand, when Leitenberg et al. (1975) evaluated resurgence with rats following three, nine, or 27 sessions of Phase 2, resurgence was significantly reduced for rats experiencing 27 sessions relative to those experiencing only three or nine sessions. Most recently,

Shahan, Browning, and Nall (2020) evaluated resurgence following four, eight, 16, 24, and 32 days of Phase 2, which resulted in a small but reliable decrease in resurgence as Phase 2 duration increased.

Research on Phase 2 duration using human subjects is also mixed. Wacker et al. (2011) evaluated whether increasing exposure to treatment reduces resurgence by repeatedly exposing participants to extinction at different points in their treatment (i.e., Phase 2), which showed that successive exposures to extinction produced decreasing levels of resurgence. These findings must be interpreted cautiously because the reinforcement schedule for alternative responses was thinned between resurgence tests, which may have contributed to the decrease in resurgence (cf. Sweeney & Shahan, 2013). Further, recent research has shown that repeated exposures to extinction may decrease resurgence (Shahan, Browning, & Nall, 2020), so the repeated exposure to extinction contingencies may be a plausible alternative explanation for the decreases in resurgence. In another applied study, Fisher et al. (2018) evaluated the effects of Phase 2 duration on resurgence; however, Phase 2 duration was manipulated in combination with two other resurgence-mitigation strategies (i.e., lower rates of reinforcement during baseline and treatment), making it difficult to draw conclusions regarding the effect of Phase 2 duration in isolation. Most recently, Greer et al. (2020) evaluated the effect of Phase 2 duration in isolation with six participants who engaged in destructive behavior using a multielement design but failed to demonstrate a reliable effect within or across participants. Greer et al. suggested that the lack of an effect may have been related to poor discrimination across the conditions in the multielement design, which is similar to one of the limitations of Lambert et al. (2020). Greer et al. suggested that future resurgence evaluations should consider using a group design to control for this potential confound.

Although RaC^2 predicts that phase duration will affect resurgence and this may have important conceptual and clinical implications, the existing research on these effects have often produced results that are difficult to interpret when evaluated with humans. Thus, the purpose of our translational experiment was to evaluate the effects of phase duration on resurgence using a human-operant preparation.

CHAPTER 1: METHOD

Recruitment and Setting

We recruited participants through Amazon's Mechanical Turk (MTurk) platform to complete the experiment evaluating the effects of phase duration on resurgence. We presented the experiment as an MTurk Human Intelligence Task (HIT) that participants could complete on a personal electronic device in any location where they were able to access the MTurk platform. All participants were at least 18 years old. We used MTurk's qualifications feature so only MTurk workers with >99% HIT approval rates could participate (i.e., they never had their work rejected previously), and we collected MTurk worker identification numbers to exclude workers trying to complete the experiment more than once. We informed the participants they could receive up to $6 for participating based on their performance during the experimental task.

Independent and Dependent Variables

The primary independent variables were Phase 1 and Phase 2 durations, which we varied across experimental groups. Phase 1 and Phase 2 durations were either 5 min or 20 min. The primary dependent variable was correct completion of the target response chain. The target response chain was defined as (a) clicking the target response button with the cursor, (b) typing a number that correctly solves an arithmetic problem, and (c) clicking the submit button. A secondary dependent variable was correct

completion of the alternative response chain, which was the same as the target response chain except that the initial link in the chain was clicking the alternative response button instead of the target response button. Target and alternative response chains were differentiated by button color (i.e., blue for target, yellow for alternative) and the side of the interface where the buttons semi-randomly appeared (i.e., left half of the interface for the target, right half of the interface for the alternative). We collected frequency data on both correct target and alternative response chain completion. We also collected frequency data on incorrect target and alternative response chain completion, both of which were defined the same as their respective correct response chain completion with the exception that the number typed during the second link of the chain did not correctly solve the arithmetic problem.

The dependent variables in this experiment were response chains and were designed to be more effortful than responses that are often used in human-operant relapse preparations. Specifically, the second link in the response chain, inputting the correct answer to the arithmetic problem (described below), increased the response effort to obtain reinforcement relative to a simple mouse click or button press, which are often the target responses in human-operant research. Increasing the response effort may help mitigate indiscriminate response allocation, a pattern of responding that has been observed in previous research on resurgence with human-operant preparations using computer-based tasks (e.g., Bolívar et al., 2017; Cox et al., 2019; Sweeney & Shahan, 2015). For example, Cox et al. (2019) attempted to increase response effort during a resurgence evaluation by requiring verbally competent adult humans to click the same button six times to produce reinforcers on a computer-based task (instead of requiring a single click). These authors continued to observe response allocation inconsistent with research conducted with other populations (i.e., nonhuman animals

and humans with minimal verbal repertoires), but it is possible that increasing the number of clicks did not increase the response effort to such an extent that it would produce significant changes in response allocation. Thus, the proposed research aimed to mitigate this possibility by increasing the response effort to a greater extent by requiring participants to complete arithmetic problems.

General Procedures

Participants independently accessed the experimental software described and validated by Smith and Greer (in preparation) through Amazon's MTurk platform. First, the experimental interface presented the participants with the informed-consent narrative. Next, participants had to complete a quiz based on the content of the informed-consent narrative and answer all questions correctly before proceeding to the experiment. Then, participants were instructed to adjust the zoom on their web browser to ensure they would be able to view the entire experimental interface, and they were required to check boxes in the outermost corners of the interface to increase the probability of compliance with this instruction before proceeding. The experimental software then presented the following instructions:

> Your task for the experiment will require you to use your cursor and keyboard to input the correct answers to math problems that will be presented on your computer screen. On your screen, one or two differently colored rectangles will appear and move around the screen randomly. First, you must click one of the rectangles. Next, a math problem will appear and you will need to type the correct number that answers the math question and click the 'SUBMIT' button. You might receive points when you submit the

> correct answer. The goal of the game is to get as many points as possible. Your compensation will be directly tied to how many points you earn. If you earn more points, you will be paid more money. The maximum amount that you can earn is $6. Do NOT press the back button on your browser or refresh your browser window during the experiment, it will start the experiment over again from the beginning.

Below these instructions there was a button with “Next” written in it. When the participant clicked this button, a short quiz was presented to ensure that the participant comprehended the instructions. The participant was asked: “How do you get points?” “What is your goal?” and “Will you get more money for scoring more points?” The participant had to select the correct response (i.e., “answering math questions correctly,” “get as many points as possible,” and “yes,” respectively) to each question from a drop-down menu. If the participant answered any question incorrectly, they were presented with the instructions again, followed by another opportunity to complete the quiz. There was no limit to the number of times a participant could take the quiz. The participant proceeded to the experiment once they answered all the quiz questions correctly.

We informed the participants that their compensation would be based on obtaining a high score to increase their motivation to respond throughout the experiment; however, all participants were paid $6 at the end of the experiment, regardless of performance on the experimental task. At the end of the experiment, we presented participants with a debriefing script informing them that compensation was not related to performance, all participants received $6, and they could withdraw their consent to have their data excluded. No participants who completed the experiment requested for their

data to be excluded. At the end of the experiment, we asked participants whether they experienced any problems or issues during the experiment. No participants reported issues other than the suspension of reinforcement at the end of the experiment, which was consistent with the programed contingencies in Phase 3.

Resurgence Evaluation

The experimental apparatus displayed a running total of the participant's points at the top of the display at all times. All participants experienced the same resurgence-evaluation conditions; the only difference across experimental groups was the duration of exposure to Phases 1 and 2.

Phase 1. A blue target response button was present and moved randomly around the left half of the display after each response. When the participant clicked the button, the background color of the display changed to the color of the response button (i.e., blue). Further, a single-digit two-term addition problem (e.g., 3 + 9) was presented along with a field to type the answer and a button to click to submit the answer. Clicking the button to submit the answer made the math problem and background color disappear. If the participant typed the correct answer prior to clicking the submit button, their point total increased by 10 points according to a variable interval (VI) 30-s schedule of reinforcement, which varied randomly between 15-s and 45-s intervals. When points were delivered, the color corresponding to the response button and background during the math problem presentation (i.e., blue for target responses) flashed behind the running point total to increase the discriminability of the source of point deliveries. All other submit button clicks did not produce points (i.e., extinction).

Phase 2. In addition to the target response button, a yellow alternative response button appeared and moved randomly around the right half of the display. Clicking the

alternative response button produced the same effects as clicking the target response in Phase 1, except the alternative response changed the background to yellow and produced points according to a VI 5-s schedule of reinforcement, which randomly varied between 0-s and 10-s intervals. The first alternative response in Phase 2 always produced reinforcement to help with the acquisition of the alternative response. In Phase 2, clicking the target button continued to produce addition problems; however, answering the math problems after clicking the target button no longer produced points (i.e., extinction). Additionally, a changeover was in place such that reinforcement was withheld for correct alternative responses if the previous response was a target response. This discouraged rapid alternation between target and alternative responses.

Phase 3. Both the target and alternative response buttons remained present. Clicking either response button produced addition problems, but correctly solving the addition problems no longer produced points (i.e., extinction).

Experimental Design

We randomly assigned each participant to one of four experimental groups until the total number of viable datasets for that group (n = 16) was obtained. We randomized participants to groups in blocks so that group totals remained relatively balanced throughout recruitment. Phase 3 lasted 5 min across all groups.

Short–Short. Phase 1 lasted 5 min, and Phase 2 lasted 5 min.

Short–Long. Phase 1 lasted 5 min, and Phase 2 lasted 20 min.

Long–Short. Phase 1 lasted 20 min, and Phase 2 lasted 5 min.

Long–Long. Phase 1 lasted 20 min, and Phase 2 lasted 20 min.

The experimental groups were designed to evaluate the effects of phase duration on resurgence when (a) the total time spent across Phases 1 and 2 were identical, but the proportional duration of each phase varied (Short–Long vs. Long–Short); (b) the duration of Phase 1 remained constant, but the duration of Phase 2 varied (Short–Short vs. Short–Long; Long–Short vs. Long–Long); (c) the duration of Phase 1 varied, but the duration of Phase 2 remained constant (Short–Short vs. Long–Short; Short–Long vs. Long–Long); and (d) the proportion of time in each phase remained constant, but the duration of each phase varied (Short–Short vs. Long–Long).

CHAPTER 2: RESULTS

We obtained a total of 68 datasets but excluded four participants' data: two participants who engaged in no responses for over 15 min, one participant who engaged in only two alternative responses during 20 min of Phase 2, and one participant who maximized reinforcement during the first five minutes of Phase 2 but obtained no reinforcers for the remaining 15 min. We included 64 of the 68 obtained datasets (i.e., 94% retention).

We evaluated the mean rate of target responding expressed as a proportion of baseline (i.e., Phase 1) response rates during the transition from Phase 2 to Phase 3 across experimental groups. To do this, we first calculated the mean Phase 1 response rate for each participant across the last 2 min of Phase 1. Next, we divided each participant's frequency of target responding in each minute of Phases 2 and 3 by their respective mean baseline response rate. We used these transformed data in the subsequent analyses.

Figure 2 displays the mean rate of target responses expressed as a proportion of baseline for each experimental group. The left panel shows the mean for each

experimental group for the last minute of Phase 2 and the first minute of Phase 3. The right panel shows the group means averaged across the last 5 min of Phase 2 and all of Phase 3 (i.e., 5 min). At both time points (i.e., end of Phase 2, beginning of Phase 3) and at both levels of analysis (i.e., 1 min, 5 min), target responding was highest in the Long–Short group, followed in order by the Short–Short, the Long–Long, and the Short–Long groups. These group means match the ordinal predictions made by RaC[2], as depicted in Figure 1.

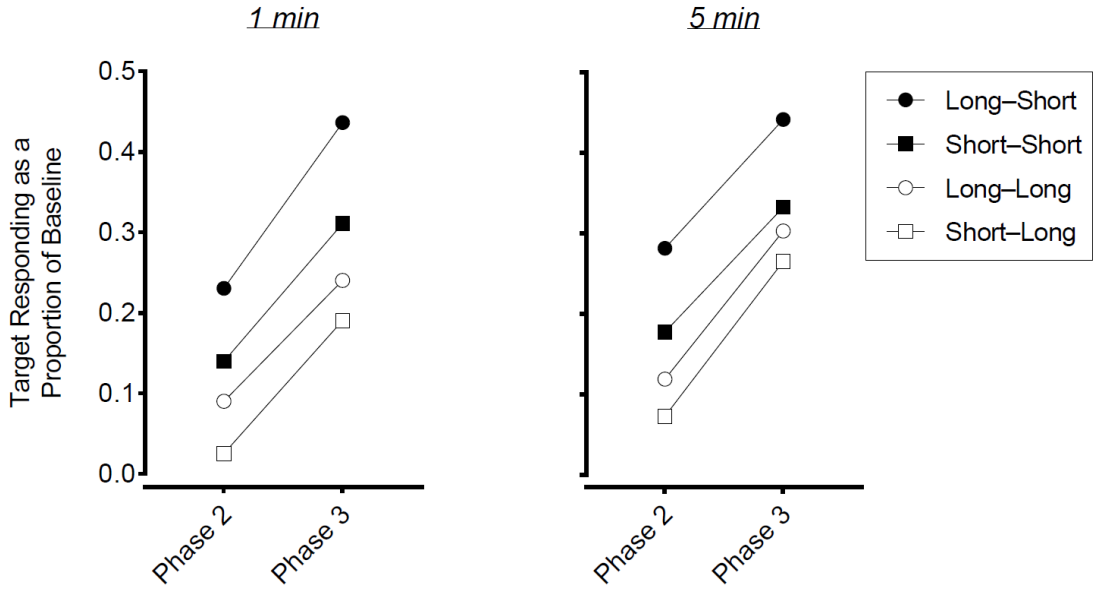

Figure 2: Obtained target responding. Mean rate of target responding expressed as a proportion of baseline. The left panel represents data obtained in the last 1 min of Phase 2 and the first 1 min of Phase 3. The right panel represents the mean obtained target response rate across the last 5 min of Phase 2 and all 5 min of Phase 3.

Figure 3 displays the frequency of target responses expressed as a proportion of baseline for each participant during the first 1 and 2 min of Phase 3, separated by experimental group. Target responding was greatest in the Long–Short group (*M* = .42, *SEM* = .08), followed in order by the Short–Short group (*M* = .37, *SEM* = .04), the Long–Long group (*M* = .29, *SEM* = .08), and the Short–Long group (*M* = .19, *SEM* = .04).

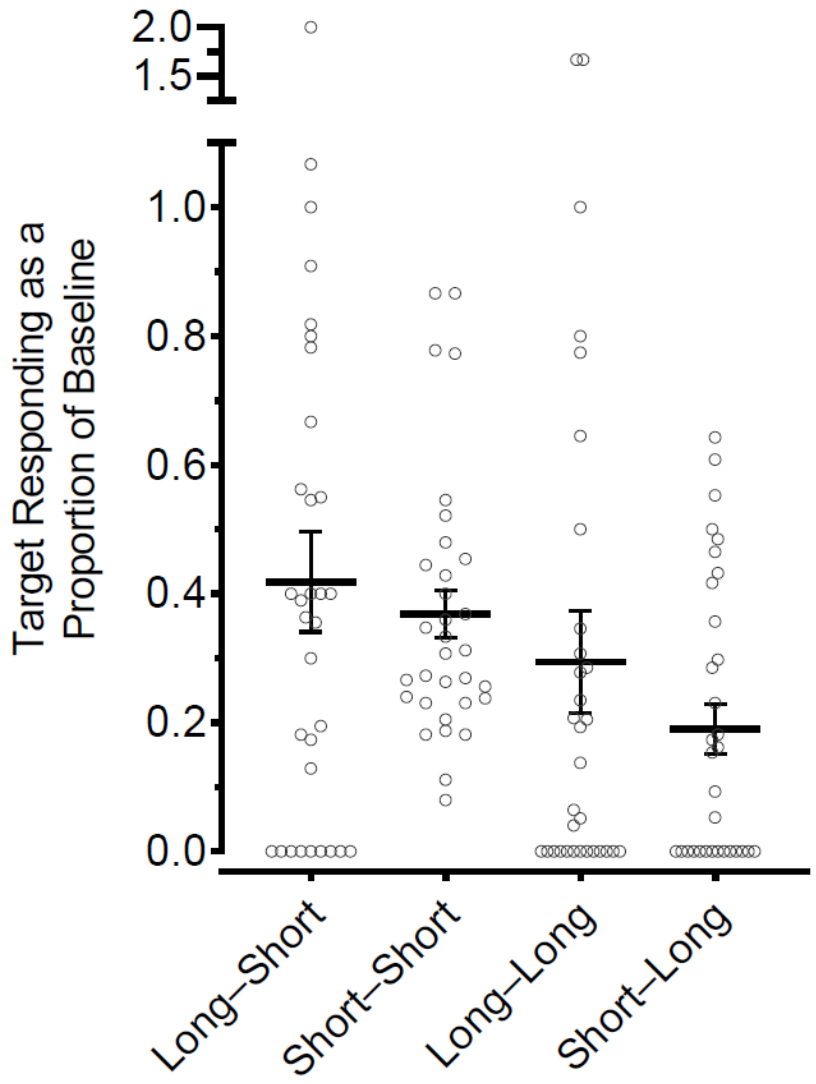

Figure 3: Target responding as a proportion of baseline. Target responses expressed as a proportion of baseline in the first 1 and 2 min of Phase 3. Horizontal lines represent the mean for the experimental group. Error bars represent standard error of the mean. Note the break and change of scale in the y-axis.

The amount of resurgence across experimental groups during the first 2 min of Phase 3 depicted in Figure 3 is consistent with the ordinal predictions made by RaC[2]. We used these data along with the frequency of target responses expressed as a proportion of baseline during the last 2 min of Phase 2 to conduct a 2 X 4 (Phase X Group) repeated-measures ANOVA, which showed a significant main effect of Phase,

$F(1, 124) = 46.05$, $p < .01$, and a significant main effect of Group, $F(3, 124) = 3.02$, $p < .05$. The Phase X Group interaction effect was not significant, $F(3, 124) = .67$, $p > .05$. A follow-up Tukey's multiple-comparisons test was conducted to evaluate which differences between experimental groups were significant. This analysis revealed a significant difference between the Long–Short and Short–Long groups in Phase 3 ($p = .03$), which RaC^2 predicted to be the largest between-group difference. No other between-group comparisons were significant (i.e., $p > .05$).

We conducted additional ANOVAs to evaluate whether the significant difference in resurgence between the Long–Short and Short–Long groups may have been related to between-group differences in response or reinforcement rates at other times during the experiment. We conducted a one-way ANOVA to evaluate differences in target response rates in the last 2 min of Phase 1, $F(3, 124) = 0.67$, $p = .57$, and differences in reinforcement rates in the last 2 min of Phase 1, $F(3, 124) = 0.42$, $p = .74$. Neither test revealed significant differences, suggesting it is unlikely that differences in Phase 1 response or reinforcement rates caused the between-group differences in Phase 3.

To evaluate the predictive validity of RaC^2 for human data, we fit the quantitative model to the obtained data. First, we calculated the mean target response rate for each experimental group across each minute of the experiment. Then, we used Microsoft Excel's Solver add-in to minimize the squared residuals between the obtained target and alternative response rates and the rates predicted by the quantitative model for all experimental groups simultaneously. We fit the quantitative model to the obtained data allowing all relevant free parameters (i.e., k, a, λ, and d_m) in RaC^2 to vary. Solver identified that the squared residuals were minimized with the following parameter values: $k = 15.76$, $a = .0073$, $\lambda = .0033$, $d_m = .33$, resulting in an $r^2 = .96$ across all experimental groups based on responding across all phases (i.e., 205 data points). We also

calculated the coefficient of determination for each group separately (Short–Long r^2 = .95; Long–Short r^2 = .96; Short–Short r^2 = .90; Long–Long r^2 = .97). Figure 4 displays the results of this model fit along with the obtained mean target and alternative responses for each experimental group across all experimental phases. Visual inspection of the data suggests the model does not systematically deviate from obtained response rates across phases or experimental groups, with the exception of slight underpredictions of target responding in Phase 3 for the Short–Short and Short–Long groups.

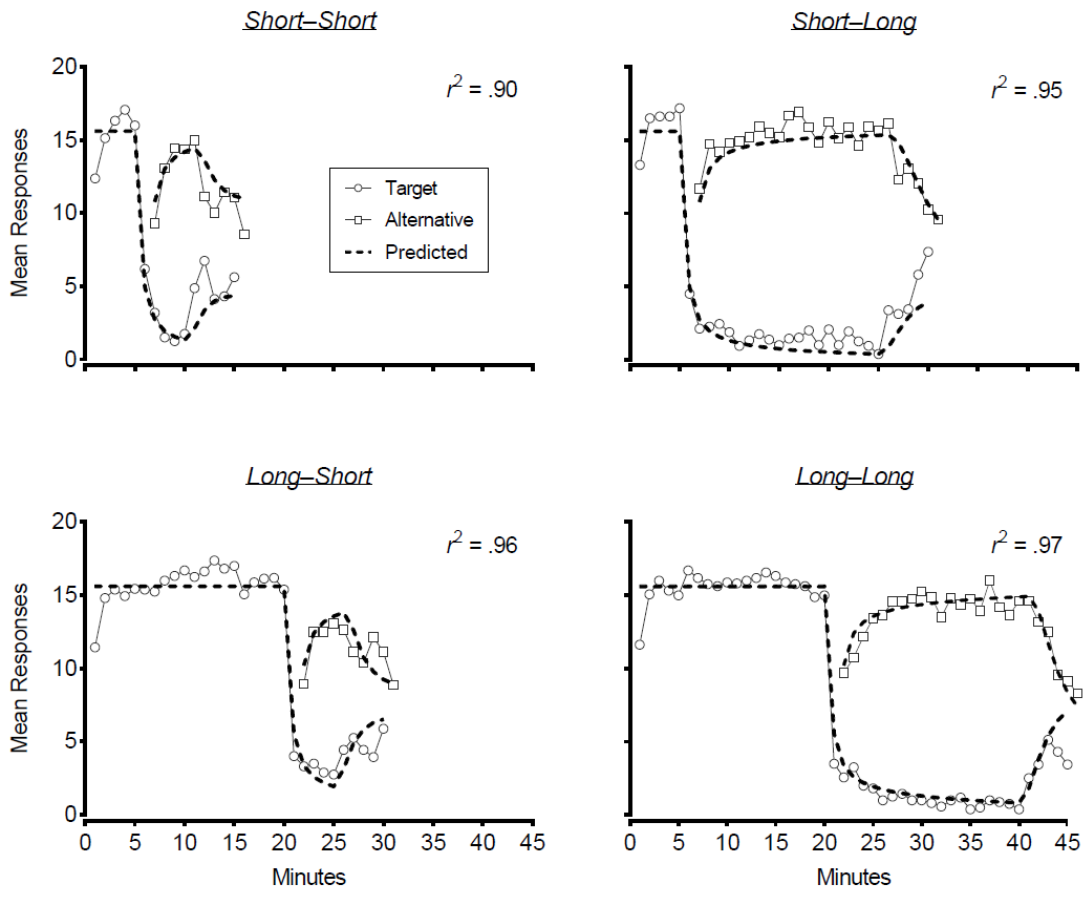

Figure 4: Model fits of RaC² to obtained data. Each panel represents the quantitative model provided by RaC² fit to the obtained mean target and alternative responses of each group throughout all three phases of the experiment. The dashed lines show RaC² fits with the following parameter values: k = 15.76, a = .0073, λ = .0033, d_m = .33.

DISCUSSION

Two findings generally support the notion that phase duration affects resurgence as predicted by RaC^2. First, the mean rate of target responding during the first 2 min of Phase 3 for the Long–Short group was significantly greater than for the Short–Long group. Although the differences between other experimental groups did not reach statistical significance, this is likely due to the relatively small predicted differences across groups, not the absence of an effect. For example, Shahan, Browning, and Nall (2020) observed small decreases in resurgence with increasing Phase 2 durations and noted that previous evaluations of Phase 2 duration with nonhuman animals have tended to demonstrate numerically more resurgence with shorter Phase 2 durations even if the observed differences did not reach statistical significance (e.g., Nall et al., 2018; Winterbauer et al., 2013). This tendency is consistent with the increased resurgence we observed with shorter Phase 2 durations when we held Phase 1 duration constant across groups (i.e., Short–Short vs. Short–Long; Long–Short vs. Long–Long). A complementary finding was the tendency for longer Phase 1 durations to increase resurgence when holding Phase 2 duration constant (i.e., Short–Short vs. Long–Short; Short–Long vs. Long–Long). Thus, by combining these relatively small effects with a phase-duration manipulation designed to maximize and minimize resurgence (i.e., Long–Short and Short–Long, respectively), we were able to demonstrate a significant effect.

A second finding supporting this combined effect of phase duration was the group means of target responding at the end of Phase 2 and the beginning of Phase 3 (Figure 2). The rates of target responding matched the order predicted by RaC^2 (Figure 1) and are consistent with the prediction that longer Phase 1 durations increase resurgence, as well as the prediction that longer Phase 2 durations decrease

resurgence. Notably, the level of target responding at the end of Phase 2 and the beginning of Phase 3 matched RaC^2's ordinal predictions at both the 1-min and 5-min level of analysis, which is highly unlikely to occur by chance (i.e., the simple probability of all 8 datapoints occurring in the exact order predicted by RaC^2 is .0017). Beyond providing support that phase duration had a predictable effect, this also suggests that there may be additive effects of phase durations. Although research in applied settings has not clearly demonstrated an effect of phase duration by manipulating the duration of a single phase in isolation (e.g., Greer et al., 2020; Lambert et al., 2020), our method of combining manipulations in both phases may prove helpful for research in more applied settings and may have important clinical implications. For example, this research may support the need for combining phase duration manipulations (i.e., earlier intervention for problem behavior, longer clinical admissions) to mitigate the possibility of treatment relapse.

This experiment may have important conceptual implications. The high coefficient of determination across all four experimental groups when we fit RaC^2 to target response rates provides promising preliminary support that RaC^2 can accurately predict human behavior. With a single set of parameter values, r^2 was at least .90 for each experimental group and equal to .96 across groups. These values are comparable to those obtained in relevant basic research with nonhuman animals (e.g., r^2 = .92 in Shahan, Browning, & Nall, 2020; r^2 = .90 and .91 in Shahan, Browning, Nist, et al., 2020). Further, our values for *k, a,* and λ (15.76, .0073, and .0033, respectively) are similar to those used to fit animal data (e.g., k = 20 in Shahan & Craig's [2017] reanalysis of Sweeney & Shahan [2013]; a = .0015 in Shahan & Craig's [2017] reanalysis of Podlesnik & Shahan [2010]; λ = .0028 and .004 in Shahan, Browning, & Nall [2020] and Shahan, Browning, Nist, et al. [2020], respectively). Our value for d_m (d_m

= .33) is smaller than other research (i.e., d_m = 8.61 and 22.25 in Shahan, Browning, & Nall, 2020; d_m = 7.67 and 5.56 in Shahan, Browning, Nist, et al., 2020). The d_m term represents the asymptotic value of the bias term, which is influenced by the discriminative properties of reinforcement deliveries (Shahan, Browning, & Nall, 2020), and larger values indicate greater biasing effects of discriminating the availability of reinforcement. One reason our d_m value may be considerably smaller than values from previous research is that our human participants may have developed rule-governed behavior that interfered with detecting shifts in contingencies. Perhaps a more compelling reason could be the large difference in the timescale across experiments. Rats in both Shahan, Browning, and Nall (2020) and Shahan, Browning, Nist, et al., (2020) experienced numerous 30-min experimental sessions, whereas our participants experienced a single experimental session for a maximum of 45 min, and this relatively short exposure may have interfered with our participants developing a larger bias based on the discriminative properties of reinforcement delivery. To our knowledge, this is the first prospective study to fit RaC^2 to human data. It is promising that our model fits resulted in parameter values that were mostly similar to those obtained with nonhuman animals; however, further research is needed to establish reasonable ranges for these parameters with humans.

This experiment provides additional support that relapse phenomena can be evaluated using human-operant preparations presented on the internet via crowdsourcing websites. Robinson and Kelley (2020) demonstrated that response patterns consistent with resurgence and ABA renewal could be obtained using an internet-based software and participants recruited through MTurk. The current experiment extends their initial finding by showing that these methods can also produce differentiated response patterns across independent variables, even when the predicted

difference is relatively small. Future experiments should continue to evaluate whether other independent variables (e.g., reinforcement rate, magnitude, quality) evaluated with nonhuman animals are replicable in a human-operant preparation. If successful, this general methodology (i.e., human-operant evaluations conducted via crowdsourcing websites) could increase the efficiency of future translational research by increasing recruitment rates and decreasing the resources needed for individual experiments.

A limitation of this experiment is that the number of reinforcer deliveries varied across phase durations. Previous research has demonstrated that the number of reinforcer deliveries across phases can impact relapse phenomena regardless of their contingencies (e.g., Podlesnik & Shahan, 2009), suggesting that phase duration was confounded by the number of reinforcer deliveries in the present study. An experimenter could control for this confound by adjusting reinforcement schedule densities such that they deliver approximately equal numbers of reinforcers in each phase across different durations (e.g., longer phases with leaner schedules, shorter phases with denser schedules). However, previous research has demonstrated that reinforcement schedule density in Phase 1 (e.g., Fisher et al., 2019) and Phase 2 (e.g., Pritchard et al. 2014; Sweeney & Shahan, 2013) both have significant effects on resurgence, so this control technique would likely introduce a different confounding variable. Instead, we held reinforcement schedule densities constant across varied phase durations, which is consistent with previous research on phase duration (e.g., Shahan, Browning, & Nall, 2020). Future research could evaluate the utility of a trial-based experimental structure to disentangle the interrelated effects of phase duration, number of reinforcer deliveries, and reinforcement schedule density.

The results of this experiment suggest that Phase 1 and Phase 2 durations impact resurgence and that their effects may be additive. These results may have

applied implications because they suggest support for early intensive treatments and for lengthened clinical admissions to minimize resurgence. Further, this experiment provides support for the utility of conducting human-operant experiments via crowdsourcing websites, demonstrates how researchers can effectively use RaC^2 to develop hypotheses that can be empirically evaluated, and provides preliminary support for the ability of RaC^2 to accurately predict human behavior. Future research along these lines holds the potential for more rapidly translating findings from the laboratory to clinical populations.

BIBLIOGRAPHY

Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, *12*(6), 861–874. https://doi.org/10.1901/jeab.1969.12-861

Bolívar, H. A., Cox, D. J., Barlow, M. A., & Dallery, J. (2017). Evaluating resurgence procedures in a human operant laboratory. *Behavioural Processes*, *140*, 150–160. https://doi.org/10.1016/j.beproc.2017.05.004

Briggs, A. M., Fisher, W. W., Greer, B. D., & Kimball, R. T. (2018). Prevalence of resurgence of destructive behavior when thinning reinforcement schedules during functional communication training. *Journal of Applied Behavior Analysis*, *51*(3), 620–633. https://doi.org/10.1002/jaba.472

Bruzek, J. L., Thompson, R. H., & Peters, L. C. (2009). Resurgence of infant caregiving responses. *Journal of the Experimental Analysis of Behavior*, *92*(3), 327–343. https://doi.org/10.1901/jeab.2009-92-327

Bullock, C. E., Fisher, W. W., & Hagopian, L. P. (2017). Description and validation of a computerized behavioral data program: “BDataPro”. *The Behavior Analyst, 40*(1), 275–285. https://doi.org/10.1007/s40614-016-0079-0

Cox, D. J., Bolívar, H. A., & Barlow, M. A. (2019). Multiple control responses and resurgence of human behavior. *Behavioural Processes*, *159*, 93–99. https://doi.org/10.1016/j.beproc.2018.12.003

Craig, A. R., Browning, K. O., Nall, R. W., Marshall, C. M., & Shahan, T. A. (2017). Resurgence and alternative-reinforcer magnitude. *Journal of the Experimental Analysis of Behavior*, *107*(2), 218–233. https://doi.org/10.1002/jeab.245

Devenport, L. D., & Devenport, J. A. (1994). Time-dependent averaging of foraging information in least chipmunks and golden-mantled ground squirrels. *Animal Behaviour*, *47*(4), 787–802. https://doi.org/10.1006/anbe.1994.1111

Falligant, J. M., Kranak, M. P., McNulty, M. K., Schmidt, J. D., Hausman, N. L., & Rooker, G. W. (2020). Prevalence of renewal of problem behavior: Replication and extension to an inpatient setting. *Journal of Applied Behavior Analysis*. https://doi.org/10.1002/jaba.740

Fisher, W. W., Greer, B. D., Fuhrman, A. M., Saini, V., & Simmons, C. A. (2018). Minimizing resurgence of destructive behavior using behavioral momentum theory. *Journal of Applied Behavior Analysis*, *51*(4), 831–853. https://doi.org/10.1002/jaba.499

Fisher, W. W., Saini, V., Greer, B. D., Sullivan, W. E., Roane, H. S., Fuhrman, A. M., Craig, A. R., & Kimball, R. T. (2019). Baseline reinforcement rate and resurgence of destructive behavior. *Journal of the Experimental Analysis of Behavior*, *111*(1), 75–93. https://doi.org/10.1002/jeab.488

Fuhrman, A. M., Lambert, J. M., & Greer, B. D. (2021). A brief review of expanded-operant treatments for mitigating resurgence. *The Psychological Record*, 1–5. https://doi.org/10.1007/s40732-021-00456-z

Greer, B. D., Fisher, W. W., Retzlaff, B. J., & Fuhrman, A. M. (2020). A preliminary evaluation of treatment duration on the resurgence of destructive behavior. *Journal of the Experimental Analysis of Behavior*, *113*(1), 251–262. https://doi.org/10.1002/jeab.567

Greer, B. D., Fisher, W. W., Saini, V., Owen, T. M., & Jones, J. K. (2016). Functional communication training during reinforcement schedule thinning: An analysis of 25

applications. *Journal of Applied Behavior Analysis*, *49*(1), 105–121. https://doi.org/10.1002/jaba.265

Greer, B. D., & Shahan, T. A. (2019). Resurgence as Choice: Implications for promoting durable behavior change. *Journal of Applied Behavior Analysis, 52*(3), 816–846. https://doi.org/10.1002/jaba.573

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

Jarmolowicz, D. P., & Lattal, K. A. (2014). Resurgence under delayed reinforcement. *The Psychological Record*, *64*(2), 189–193. https://doi.org/10.1007/s40732-014-0040-0

Kurtz, P. F., Boelter, E. W., Jarmolowicz, D. P., Chin, M. D., & Hagopian, L. P. (2011). An analysis of functional communication training as an empirically supported treatment for problem behavior displayed by individuals with intellectual disabilities. *Research in Developmental Disabilities*, *32*(6), 2935–2942. https://doi.org/10.1016/j.ridd.2011.05.009

Lambert, J. M., Bloom, S. E., Samaha, A. L., & Dayton, E. (2017). Serial functional communication training: Extending serial DRA to mands and problem behavior. *Behavioral Interventions, 32*(4), 311–325. https://doi.org/10.1002/bin.1493

Lambert, J. M., Pericozzi, H. G., Bailey, K. M., Standish, C. M., & Perry, E. C. (2020). Evaluating duration of baseline as a moderator of resurgence following serial training. *Behavioral Development, 25*(2), 52–65. https://doi.org/10.1037/bdb0000096

Lattal, K. A., Cançado, C. R., Cook, J. E., Kincaid, S. L., Nighbor, T. D., & Oliver, A. C. (2017). On defining resurgence. *Behavioural Processes*, *141*, 85–91. https://doi.org/10.1016/j.beproc.2017.04.018

Leitenberg, H., Rawson, R. A., & Bath, K. (1970). Reinforcement of competing behavior during extinction. *Science*, *169*(3942), 301–303. https://doi.org/10.1126/science.169.3942.301

Leitenberg, H., Rawson, R. A., & Mulick, J. A. (1975). Extinction and reinforcement of alternative behavior. *Journal of Comparative and Physiological Psychology, 88*(2), 640–652. https://doi.org/10.1037/h0076418

Mazur, J. E. (1996). Past experience, recency, and spontaneous recovery in choice behavior. *Animal Learning & Behavior*, *24*(1), 1–10. https://doi.org/10.3758/BF03198948

Mitteer, D. R., Greer, B. D., Randall, K. R., & Haney, S. D. (under review). On the scope and characteristics of treatment relapse when treating destructive behavior.

Muething, C., Call, N., Pavlov, A., Ringdahl, J., Gillespie, S., Clark, S., & Mevers, J. L. (2020). Prevalence of renewal of problem behavior during context changes. *Journal of Applied Behavior Analysis*, *53*(3), 1485–1493. https://doi.org/10.1002/jaba.672

Muething, C., Pavlov, A., Call, N., Ringdahl, J., & Gillespie, S. (2020). Prevalence of resurgence during thinning of multiple schedules of reinforcement following functional communication training. *Journal of Applied Behavior Analysis*. Advance online publication. https://doi.org/10.1002/jaba.791

Nall, R. W., Craig, A. R., Browning, K. O., & Shahan, T. A. (2018). Longer treatment with alternative non-drug reinforcement fails to reduce resurgence of cocaine or

alcohol seeking in rats. *Behavioural Brain Research*, *341*, 54–62. https://doi.org/10.1016/j.bbr.2017.12.020

Podlesnik, C. A., & Shahan, T. A. (2009). Behavioral momentum and relapse of extinguished operant responding. *Learning & Behavior, 37*(4), 357–364. https://doi.org/10.3758/LB.37.4.357

Podlesnik, C. A., & Shahan, T. A. (2010). Extinction, relapse, and behavioral momentum. *Behavioural Processes*, *84*(1), 400–411. https://doi.org/10.1016/j.beproc.2010.02.001

Pritchard, D., Hoerger, M., Mace, F. C., Penney, H., & Harris, B. (2014). Clinical translation of animal models of treatment relapse. *Journal of the Experimental Analysis of Behavior, 101*(3), 442–449. https://doi.org/10.1002/jeab.87

Robinson, T. P., & Kelley, M. E. (2020). Renewal and resurgence phenomena generalize to Amazon's Mechanical Turk. *Journal of the Experimental Analysis of Behavior*, *113*(1), 206–213. https://doi.org/10.1002/jeab.576

Shahan, T. A., Browning, K. O., & Nall, R. W. (2020). Resurgence as Choice in Context: Treatment duration and on/off alternative reinforcement. *Journal of the Experimental Analysis of Behavior*, *113*(1), 57–76. https://doi.org/10.1002/jeab.563

Shahan, T. A., Browning, K. O., Nist, A. N., & Sutton, G. M. (2020). Resurgence and downshifts in alternative reinforcement rate. *Journal of the Experimental Analysis of Behavior*, *114*(2), 163–178. https://doi.org/10.1002/jeab.625

Shahan, T. A., & Craig, A. R. (2017). Resurgence as Choice. *Behavioural Processes*, *141*, 100–127. https://doi.org/10.1016/j.beproc.2016.10.006

Shahan, T. A., & Greer, B. D. (under review). Destructive behavior increases as a function of reductions in alternative reinforcement during schedule thinning: A retrospective quantitative analysis.

Smith, S. W., & Greer, B. D. (in preparation). Description and validation of an internet-based human decision-making software.

Sweeney, M. M., & Shahan, T. A. (2013). Effects of high, low, and thinning rates of alternative reinforcement on response elimination and resurgence. *Journal of the Experimental Analysis of Behavior, 100*(1), 102–116. https://doi.org/10.1002/jeab.26

Sweeney, M. M., & Shahan, T. A. (2015). Resurgence of target responding does not exceed increases in inactive responding in a forced-choice alternative reinforcement procedure in human. *Behavioural Processes, 124*, 80–92. https://doi.org/10.1016/j.beproc.2015.12.007

Volkert, V. M., Lerman, D. C., Call, N. A., & Trosclair-Lasserre, N. (2009). An evaluation of resurgence during treatment with functional communication training. *Journal of Applied Behavior Analysis, 42*(1), 145–160. https://doi.org/10.1901/jaba.2009.42-145

Wacker, D. P., Harding, J. W., Morgan, T. A., Berg, W. K., Schieltz, K. M., Lee, J. F., & Padilla, Y. C. (2013). An evaluation of resurgence during functional communication training. *The Psychological Record, 63*(1), 3–20. https://doi.org/10.11133/j.tpr.2013.63.1.001

Winterbauer, N. E., Lucke, S., & Bouton, M. E. (2013). Some factors modulating the strength of resurgence after extinction of an instrumental behavior. *Learning and Motivation, 44*(1), 60–71. https://doi.org/10.1016/j.lmot.2012.03.003