REVISTA MEXICANA DE ANÁLISIS DE LA CONDUCTA 2015 NÚMERO 2 (SEPTIEMBRE)
MEXICAN JOURNAL OF BEHAVIOR ANALYSIS VOL. 41, 63-83 NUMBER 2 (SEPTEMBER)

EFFECTS OF REPETITION FREQUENCY ON OPERANT STRENGTH AND RESURGENCE OF NON-CRITERIAL FEATURES OF OPERANTS

EFECTOS DE LA FRECUENCIA DE REPETICIÓN SOBRE LA FUERZA OPERANTE Y EL RESURGIMIENTO DE LAS CARACTERÍSTICAS DISCRECIONALES DE LAS OPERANTES

FRANCIS MECHNER
THE MECHNER FOUNDATION

LAURILYN D. JONES
THE MECHNER FOUNDATION AND
OSLO AND AKERSHUS UNIVERSITY COLLEGE

Abstract

The present research is based on the conceptualization of resurgence as reappearance of behavior that occurred earlier in the individual's history but not recently, without restoration of the conditions under which the earlier behavior occurred. In a series of five experiments, human participants typed nonword sequences of letters on a computer keyboard. Each sequence was initiated with the spacebar and ended with the enter key, and was treated as a "revealed operant." Each operant was composed of criterial (mandated) and noncriterial (discretionary) keystrokes. Participants learned several unique operants, each defined by a different set of criterial keystrokes. The objective was to study the effect of varying the number of repetitions required for each operant during the learning sessions on the relative frequency with which those operants were performed during a test session. The operants that had previously been performed most frequently were chosen for performance most often. Noncriterial resurgence was measured by "antiquity"—how far back one has to go in the partici-

Corresponding Author: Francis Mechner, The Mechner Foundation, 200 Central Park South, 18E, New York, NY 10019. E-mail: fmechner@panix.com

pant's prior history to find previous instances of that same noncriterial keystroke sequence. Criterial and noncriterial components of the operants were affected by the independent variable in different ways. There were significantly higher-than-average levels of resurgence in invalid operants, as well as higher numbers of never previously used noncriterial patterns, suggesting that many types of performance errors in operant behavior may be instances of resurgence. The research implications of the present conceptualization of resurgence are explored and discussed.

Keywords: resurgence, history variables, variability, criterial and noncriterial features, repetition frequency, revealed operant, performance mistakes, practice effects.

Resumen

La presente investigación está basada en la conceptualización de resurgimiento como la reaparición de la conducta que ocurrió antes en la historia del individuo pero no recientemente, sin la restauración de las condiciones bajo las cuales la conducta anterior ocurrió. En una serie de cinco experimentos, participantes humanos escribían en un teclado de computadora secuencias de letras que no conformaban palabras. Cada secuencia se iniciaba con la barra espaciadora y terminaba con la tecla de *enter* y se trató como una "operante revelada". Cada operante estuvo compuesta de presiones a las teclas que cumplían con un criterio establecido (obligatorias) y presiones que no cumplían con el criterio (discrecionales). Los participantes aprendieron varias operantes únicas, cada una definida por un conjunto diferente de presiones obligatorias de las teclas. El objetivo fue estudiar el efecto de variar el número de repeticiones requerido para cada operante durante las sesiones de aprendizaje sobre la frecuencia relativa con la cual estas operantes se ejecutaban durante una sesión de prueba. Las operantes que se habían ejecutado más frecuentemente se eligieron para su ejecución la mayor parte de las veces. El resurgimiento discrecional se midió mediante su "antigüedad"—qué tan atrás uno debe de ir en la historia del participante para encontrar instancias previas de la misma secuencia discrecional de presiones a las teclas. La variable independiente afectó de diferentes maneras a los componentes obligatorios y discrecionales de las operantes. Hubo niveles de resurgimiento de operantes inválidas significativamente más altos que el promedio, así como un número más grande de patrones discrecionales que nunca se habían usado, lo cual sugiere que muchos tipos de errores de ejecución en la conducta operante pueden ser instancias de resurgimiento. Se exploran y se discuten las implicaciones para la investigación de la presente conceptualización de resurgimiento.

Palabras clave: resurgimiento, variables de historia, variablidad, características obligatorias y discrecionales, frecuencia de repetición, operante revelada, errores de ejecución, efectos de práctica

An organism's operant behavior repertoire changes continuously throughout its lifetime. It is therefore hardly surprising that the evolution of behavioral repertoires has generated the heuristic that "if somethings works, keep doing it"—the susceptibility to reinforcement. Nor is it surprising that a mirror heuristic evolved in parallel, one that can be stated as "if something isn't working or no longer works, try something that worked previously, or worked in a similar situation."

To be useful, a definition and conceptualization of resurgence should describe the phenomenon without presumptions as to its causation. The present conceptualization of resurgence as the reappearance of behavior that occurred earlier in the individual's history but not recently, without restoration of the conditions under which the behavior occurred previously, leaves the door open wide to an experimental analysis of the variables that may cause it. The definition of resurgence that references only extinction (e.g., Epstein, 1985), on the other hand, is conceptually as well as empirically unsatisfactory—conceptually because it invokes a presumptive cause, and empirically because extinction is clearly not the only cause of resurgence.

This perspective raises the question of how much of an organism's normal behavior involves resurgence. Is substantially *all* operant behavior composed of pieces or variants of earlier forms? Is the normal variability in any behavior stream due to resurgence events? The answers to these questions require an experimental analysis program that would address a wide spectrum of such topics.

The five experiments presented in this paper can be viewed as a small beginning for such a research program, one that focuses on the effects of learning history variables. The experiments examined the effects of the frequency with which operants were previously repeated on the frequency with which they are subsequently chosen for performance. The same methods were used in a series of five experiments on the effect of the *sequence* in which operants were previously learned and practiced, with the finding that the first learned and the most recently learned operants were the ones most often chosen for performance under forced–choice "test" conditions similar to those used in the present experiments (Mechner & Jones, 2011).

All of these experiments also examined resurgence of the operants' noncriterial (optional) features under various conditions, and the differential effects of various experimental conditions on the criterial (mandated) features of operants —those that must be present for the operant to be considered as having occurred (for example, the number of degrees a rat must depress a lever)— and some of the noncriterial aspects of operants, using the revealed operant technique (Mechner, 1992; Mechner, Hyten, Field & Madden, 1997), which permits tracking of both the criterial and certain noncriterial attributes of each individual occurrence of an operant. The revealed operant unit is a research preparation consisting of a sequence of recorded actions whose beginning and end is marked by distinct behavioral events. Some of these actions (sub–operants) are mandated by the definition of the operant unit, and some are not.

All operants have both criterial and noncriterial features. The noncriterial features are the nondefining characteristics of an operant, including topographic ones. These features of the operant are usually not identified and recorded, but the revealed operant preparation allows some of them to be identified and recorded conveniently. Variability in noncriterial features of an operant increases during extinction in successive conditioning–extinction cycles (Antonitis, 1951). Variability levels during extinction, in turn, are affected by the topography chosen for the response (Morgan & Neuringer, 1990).

Prior studies have shown that resurgence can occur when extinction is instituted (Epstein, 1985; Mechner et al., 1997). However, resurgence is not induced only by extinction. It is also observed when a participant's work requirement is abruptly raised (Mechner et al, 1997), and resurgence of derived relations has been documented under restricted choice conditions in research on equivalence classes (Wilson & Hayes, 1996). It has also been shown that differences in contingencies or in the nature and amount of reinforcement during acquisition can affect the amount and nature of resurgence during extinction (Dixon & Hayes, 1998; Pittenger, Pavlik, Flora, & Kontos 1988).

The present experiments examined the relation between the number of times an operant was previously performed and its subsequent strength, by using different but roughly equivalent operants repeated different numbers of times. This topic has been addressed previously via studies that used operants defined by single switch closures (such as bar presses or key pecks) and that examined resistance to extinction (Dyal & Holland, 1963; Kass & Wilson, 1966; Lewis & Duncan, 1958; Perin, 1942; Senkowski, 1978; Tombaugh, 1967) or responding in the presence of free food, known as contrafreeloading (Bilbrey, Patterson & Winokur, 1973; Jensen, 1963; Jensen, Leung & Hess, 1970; Lentz & Cohen, 1980; Stolz & Lott, 1964). These studies produced conflicting results.

Many studies have measured the effect of amount of prior training on the degree of control that a stimulus acquires for discrimination or generalization tasks (Farthing & Hearst, 1968; Hearst & Koresko, 1968; Rilling & Budnik, 1975; Sewell & Nickel, 1979; Thomas & Williams, 1963) and reversal (Lovejoy, 1966; Mackintosh, 1965, 1969; McAllister, Capehart & Rogers, 1970; Reid, 1953). In general, results from this body of research show that additional training (or "overtraining", as it is sometimes called) improves stimulus control and leads to faster reversal, although this effect is far from simple and, again, there are conflicting results.

The main independent variable in the five experiments presented in this paper is the number of times each of several different but equivalent operants were repeated in the participant's previous history. The dependent variable is the number of times a participant chose each of those operants for performance in a final test session in which the experimental conditions were made "stressful" by the imposition of time pressure and accuracy requirements. In examining the performed operants, both the criterial and the noncriterial features of those operants were recorded and analyzed.

All of the experiments followed the same general procedure: Participants learned several different but presumably equivalent operants, and then performed these different operants different numbers of times in a series of sessions according to the experimental design, followed by a final test session in which they chose which operants to perform, under time pressure and an accuracy requirement.

Method

Participants

The participants were mostly university students, both male and female, ranging in age from 18 to 57 years and recruited through flyers posted on local campuses. Each experiment consisted of either eight or nine "learning sessions" approximately an hour in duration in which the participants learned and practiced the required operants, followed by a final "test session." Participants completed one session per day, with all of each participant's sessions scheduled at the same time each day. They were told they would be paid $10 in cash per session completed and, in addition, could earn a bonus of up to $200 during the test session, with the amount earned depending on their performance.

Participants received a briefing on all requirements of the study in writing at the beginning of the first session; this document also informed them of their right to withdraw from the experiment at any time. If they did so, they would keep the $10 already paid per session completed, but would not be eligible to earn the bonus during the final session. They signed consent forms attesting that they were at least 18 years old, indicating that they understood their right to withdraw, and agreeing to keep such behavioral factors as caffeine consumption, eating habits and amount of sleep consistent from day to day during the course of the experiment.

Seven participants each completed Experiments 1, 2, 3, and 4, and nine completed Experiment 5, for a total of 37 participants. Ten participants dropped out before completing the final session. After the final session of the study, all participants were debriefed by the experimenter regarding the purpose of this research.

Apparatus

The experimental room contained four Dell 486 desktop computers at workstations separated by screens. Each had a 14–inch CRT monitor and a standard computer keyboard. Each of the keyboards was fitted with a particleboard "mask" —seen in Figure 1— that covered all the keys except for those used in the experiment: 12 char-

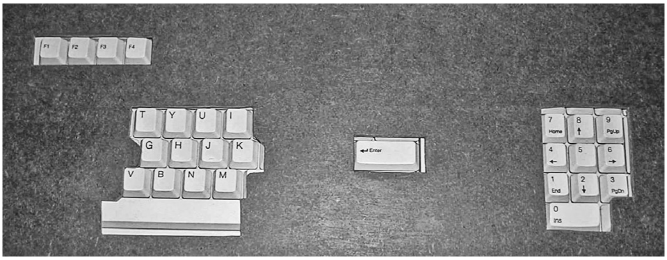

Figure 1. Picture of the keyboard and mask used in the experiments, showing the subset of 12 letter keys in use.

acter keys (tyuighjkvbnm), the space bar, the enter key, the number keypad, and four function keys. The computers operated from custom–written software that controlled the experiment and recorded the results.

General Procedure

The particular revealed operant used in this series of experiments consisted of a nonword sequence of 14 or more keystrokes. Each operant was initiated by pressing the space bar and ended by pressing the enter key, with 12 or more letter keypresses required in between. The first three and last three letters of each sequence were mandated (criterial), with each specific pattern of letters defining a unique operant. In all experiments, participants learned and practiced six different operants, defined by the first and last three–letter sequences, as follows: 1. BHJ–HMK, 2. HUB–THI, 3. VYN–KUB, 4. GTM–KMT, 5. UGK–JIB, and 6. TMY–IBY.

To complete any given operant, participants were required to type at least six noncriterial keystrokes between the mandated first and last three letters of the sequence. For these six or more noncriterial keystrokes, participants could type any letters from the set of twelve available character keys, as long as they did not repeat either the immediately preceding keystroke, or the one before that —a requirement imposed to prevent excessive keystroke repetition, like typing the same letter six or more times.

For example, to complete the operant VYN-KUB, participants were required to perform the following sequence of keystrokes: space bar, V, Y, N, six or more of any of the 12 letters available without repeating the same letter twice in a row or with only one other letter separating them, K, U, B, enter key.

During the learning sessions, only one operant was acceptable at any given time; the software was programmed to switch the required operant after every completed "block" consisting of either 13 or 26 valid operants of that type. The block size varied, as did the number of learning sessions, according to the total number of operant repetitions programmed for the particular experiment. The order in which blocks of different operants were required was varied so as to be unpredictable to the participants. In all of the experiments, all six operants were practiced during each learning session.

Operants were considered valid if they fulfilled the definitional criteria, i.e. began with the space bar press, ended with the enter key press and contained at least 12 letter keystrokes with the first and last three matching the specific criterial pattern in use at that time, and the middle six or more fulfilling the variability requirement explained above. Invalid operants were not counted toward the total required in order to finish a block or the session, thus ensuring that each participant typed each of the operants correctly the number of times required by the experimental design.

The computer monitor displayed a visual cue (the monitor screen turned from black to blue) at the instant the participant initiated an operant (i.e., pressed the space bar), with the screen turning back to black when the enter key was pressed. During the learning sessions the computer also displayed visual feedback for half a second—a 3–inch green square flashed in the middle of the screen—every time a valid operant was completed. At no point in any of the sessions did the monitor display the characters typed by the participants. At any time during the learning sessions, participants who had forgotten the individual letters mandated for a given operant could press a key on the number keypad to see the six letters making up that numbered operant displayed briefly on the monitor.

Before the first session began, the experimenter instructed the participants regarding the procedure for completing a valid operant, and allowed each of them to try a demonstration version of the software on the computer, while remaining nearby to answer any questions. After starting each participant on the first session, the experimenter left the room. Participants then worked at their own pace.

Throughout the first session, participants were also given explicit instruction by the software, in the form of a message periodically displayed on the screen, as to which of the different operants was in use at any given time during the session. At the beginning of the second session, the experimenter instructed the participants that they would no longer receive these messages and from this point forward must try each of the six different operants until they found the one that produced the green square, continue typing that operant until it stopped producing the green square, then try each of the other five until they found the next operant that would produce the green square. Participants who asked in advance about the nature of the test session were told only that they would be typing the same operants they had learned, and that the amount of money earned on the final day would depend on their accuracy, speed, and how well they remembered the patterns of letters.

In all five experiments, some of the operants were required more often than others during the learning sessions (without the participants being explicitly informed of this fact). In each experiment, the six operants were grouped into three categories of repetition frequency, those required least often ("lowest–repetition"), those required an intermediate number of times ("medium–repetition"), and those required most often ("highest–repetition"), with two different operants per category (to help balance out possible bias for or against the letters making up the individual operants). The independent variables were the different absolute numbers of repetitions required during the learning sessions for each category and also the different ratios (1:2:4 and 1:3:9) of relative number of repetitions among the three categories (see Table 1 in the section on individual experiment procedural differences).

There were two different types of test session, described below in the section on differences between experiments. In both, participants were allowed to choose which of the six operants to perform, within certain limits depending on the type of test session, and subject to time constraints. The green square did not appear, and each valid operant (correctly–typed and conforming to the limits on choice) instead produced 65 cents, provided it was executed quickly enough. Incomplete, invalid or too–slow operants resulted in a loss of 35 cents each. In addition, if the participant pressed one of the numeric keys on the number keypad to see one of the sequences of letters during the test session, they also lost 35 cents.

At the beginning of the test session the participants were instructed by the experimenter that each monetary award must be "rung up" by typing the amount on the number keypad, followed by pressing the enter key. Every money presentation was signaled by a high–pitched "beep" tone with the following message appearing on the screen: "You just earned 65 cents. Ring it up." Completion of this "consummatory response" was required before the participant could continue with the next operant. Continuously displayed in the upper left corner of the monitor was the net amount that had been earned by the participant up to that point. Whenever money was lost, the amount was deducted from this total automatically, with an accompanying low tone distinct from the "ring it up" beep. The test session for all 5 experiments was programmed to end after 460 operant attempts regardless of whether they were valid or invalid by the experiment's criteria.

Both types of test session imposed the same time constraints, penalizing too–long pauses between keystrokes to force a steady pace of work. The initial time limits were set at 1.25 s between keystrokes during an operant, and 5 s between operants. If the participant paused for longer than that, 35 cents was deducted from his or her total earnings for the session, as displayed on the screen, and the operant in progress (if any) ended and was recorded as incomplete. As the test session continued, those time limits adjusted to each participant's own particular speed, with the time limit for pausing between keystrokes programmed to be 1.25 times that participant's average time between keystrokes over the last 20 operant attempts, and the time limit for pausing between operants programmed to be 5 times the individual participant's average time over the preceding 20 operant attempts.

Individual Experiments

The experiments used two different formats for the test session to determine whether, and if so how, the results would depend on the type of test session used. In Experiments 1 and 2, participants could use any of the six operants during the final session, provided they did not type the same one twice in a row (Page & Neuringer, 1985). They were thus forced to vary the operants they used rather than type the same one over and over, as they were required to do during the learning sessions.

In Experiments 3, 4 and 5, the ban on repetition of an operant during the final session was lifted, but participants were limited to a subgroup of only three operants at any given time—one from each of the three repetition–frequency categories (lowest, medium and highest)—from which to choose the one to perform. At the beginning of the test session three of the six possible operants were displayed on the computer screen, one from each of the three frequency categories. After every block of 20 operant attempts the monitor displayed a new set of three acceptable operants, each set again consisting of one operant from each of the three categories, with the exception of a few "control blocks" which presented more than one operant from a category, just to check whether a participant showed any bias for or against any of the individual operants. No systematic bias of this kind was found, and the results from these control blocks were not counted when calculating the repetition data presented in the Results section.

The experiments also used two different ratios of required repetitions, to determine whether such ratios affect the result. Table 1 shows the ratio of numbers of repetitions for categories of operants during learning sessions, and type of test session, for each of the five experiments in the series. In Experiments 1, 4 and 5 the three categories of operants were required in a 1:2:4 ratio of repetitions during the learning sessions, while in Experiments 2 and 3 they were required in a 1:3:9 ratio, thus ensuring that if there were differences in the test session results between experiments it would be clear whether they were due to the ratio of repetitions, or the type of test session. The only difference between Experiments 4 and 5 was that in Experiment 5 each of the six operants used in all the studies was assigned to a different repetition–frequency category, again to control for possible individual operant bias.

Each of these experiments had two objectives:

1. To study the effects of number of prior repetitions of operants during learning sessions on the frequency of performance of those operants during the test session.
2. To study the occurrence and characteristics of resurgence of noncriterial patterns, and to determine whether criterial and noncriterial features of operants can respond to independent variables in different ways.

Table 1
Individual Experiments

Experiment Number	Repetitions Required per Operant in a Group	Number of Learning Sessions	Type of Final Test Session
1	234 / 468 / 936 (1:2:4 ratio)	9 (14 blocks of 26 operants each)	Free choice w/diversity required
2	104 / 312 / 936 (1:3:9 ratio)	8 (26 blocks of 13 operants each)	Free choice w/diversity required
3	104 / 312 / 936 (1:3:9 ratio)	8 (26 blocks of 13 operants each)	Rotating 1 of 3 choice
4	234 / 468 / 936 (1:2:4 ratio)	9 (14 blocks of 26 operants each)	Rotating 1 of 3 choice
5	234 / 468 / 936 (1:2:4 ratio)	9 (14 blocks of 26 operants each)	Rotating 1 of 3 choice

Results

Results for each of the objectives listed above are presented separately:

1. Impact of Prior Repetition on Subsequent Frequency of Performance

The median percent of times operants were chosen for performance from each group during the test session are presented in Figure 2; the error bars show the interquartile ranges. For complete individual data, please see Table 2.

In 31 out of the 37 participants in this series of experiments, the operants that were performed most often during the learning sessions (the highest–repetition category) were performed more often during the test session than either the ones performed least often (from the lowest–repetition category) or those from the medium–repetition category. This finding is statistically significant at the .001 level, and was consistent across both types of test session, $F(2, 108) = 66.63$, $p<.001$. In addition, the relative ratio of repetitions (1:2:4 vs. 1:3:9) had no statistically significant effect on the magnitude of the relative frequency of performance in the test session for the highest–repetition operants. Switching the specific operants required for the highest–repetition category in Experiment 5 also made no difference, thus ruling out the explanation of bias for individual operants.

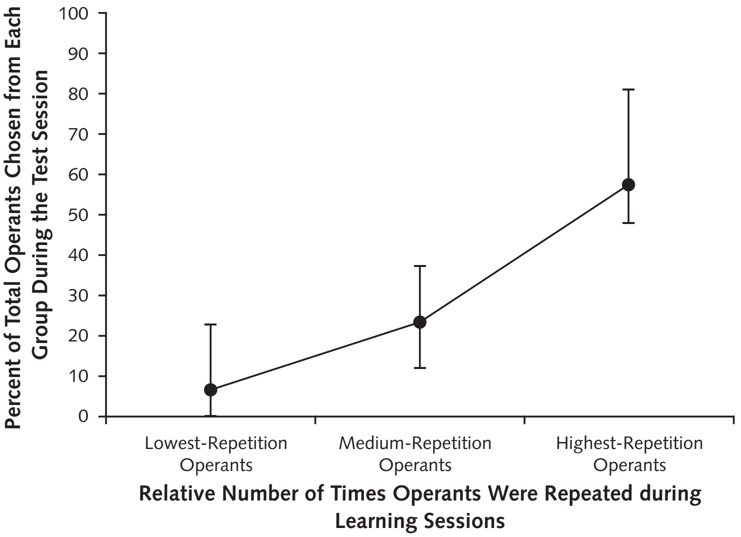

Figure 2. Frequency with which the three categories of operants occurred in the test session, as a function of the relative number of times they were repeated during the learning sessions; median of all participants with interquartile ranges.

The frequency of medium–repetition operants during the test session, however, showed considerable variability among participants. Although all of those who showed the repetition effect discussed above performed highest–repetition operants most often in the test session, some performed the medium–repetition operants less often than the lowest–repetition ones (creating a V–shaped function), while others (the largest group) produced a monotonically increasing function in the test session, with the medium–repetition category performed more often than the lowest–repetition one but less often than the highest–repetition one. Some participants performed the medium–repetition category almost as often as the highest–repetition category. Such variation for the medium–repetition operants is reflected in the significant overlap seen in the error bars for the lowest– and medium–repetition points in Figure 2, although the difference between those two categories is still statistically significant at the .01 level, $t(36) = 2.91$, $p < .01$.

In none of the five experiments was there any statistically significant difference in the accuracy and validity with which the different repetition–frequency categories of operants were executed during the test session.

Table 2
Individual Test Session Repetition Data, All Experiments

Experiment & Participant Number	Low–Repetition Group	Medium–Repetition Group	High–Repetition Group
1 - 221	0.00	49.64	50.36
1 - 222	40.05	12.09	47.87
1 - 223	0.00	21.50	78.50
1 - 224	13.48	0.25	86.27
1 - 225	0.00	46.75	53.25
1 - 226	0.96	30.38	68.66
1 - 227	0.00	48.99	51.01
2 - 241	13.87	28.71	57.42
2 - 242	20.62	18.47	60.91
2 - 243	0.00	0.00	100.00
2 - 244	18.14	20.91	60.96
2 - 245	32.84	38.56	28.61
2 - 246	0.00	45.20	54.80
2 - 247	32.85	34.05	33.09
3 - 251	0.00	0.00	100.00
3 - 252	49.36	1.28	49.36
3 - 253	0.00	11.36	88.64
3 - 254	0.00	6.09	93.91
3 - 255	3.18	16.88	79.94
3 - 256	20.58	31.19	48.23
3 - 257	10.71	33.12	56.17
4 - 281	50.65	36.60	12.75
4 - 282	0.00	37.42	62.58
4 - 283	6.33	0.00	93.67
4 - 284	6.49	23.38	70.13
4 - 285	25.08	6.44	68.47
4 - 286	19.05	43.54	37.41
4 - 287	12.99	4.55	82.47
5 - 291	32.39	22.01	45.60
5 - 292	0.00	15.31	84.69
5 - 293	38.24	31.03	30.72
5 - 294	6.39	37.38	56.23
5 - 295	0.00	57.32	42.68
5 - 296	0.00	12.22	87.78
5 - 297	30.48	16.19	53.33
5 - 298	18.30	60.88	20.82
5 - 299	0.00	25.16	74.84
Median	6.49	23.38	57.42
1st Quartile	0.00	11.73	48.05
3rd Quartile	22.85	37.40	81.21

2. Resurgence of Noncriterial Operant Dimensions

All noncriterial keystrokes typed between the mandated first and last three letters of each operant were recorded. For the purposes of data analysis, all noncriterial keystrokes in a given operant were grouped into overlapping 3–letter patterns that formed unique units. For example, if a participant had typed the letters GHJVYN in the middle of a specific operant, the groupings would be broken down into four 3–letter patterns: GHJ, HJV, JVY and VYN. Each of these 3–letter noncriterial patterns was then tracked to wherever in the experiment it had occurred previously in that participant's history of noncriterial keystrokes (if it had occurred at all).

Generally, noncriterial patterns used by participants in these experiments tended to be strongly associated with the specific criterial features of the operants in which they occurred, and became stereotyped relatively quickly. For example, when typing the operant that began with the letters BHJ and ended with the letters HMK, participants showed a strong tendency to type the same six letters in between (in the same order) every time they performed that operant, even though they could have used different letters. The total number of unique 3–letter noncriterial patterns used by a given participant during a single session was consistently highest during the first session, then dropped dramatically in the second session and quickly stabilized, suggesting that specific noncriterial patterns were becoming automatized as part of the complete chain of keystrokes that made up each operant.

As in Mechner & Jones (2011), resurgence was measured by calculating the "antiquity" of each three–letter noncriterial pattern. Antiquity is defined as the number of sessions in an individual participant's learning history one must count back to find prior instances of that same pattern (specifically, the average of the last three times the pattern was used— a refinement that helped clarify the effect). A higher average antiquity level (greater than 1) thus indicates an older noncriterial pattern. Noncriterial resurgence, as measured by average antiquity of noncriterial patterns, occurred in this series of experiments during the final test session at levels roughly comparable to those observed in a group of related unpublished experiments.

Interestingly, all of the increases in noncriterial resurgence observed during the test sessions occurred within invalid operants. (Note that noncriterial patterns *themselves* cannot be invalid, by definition. Operants are termed "invalid" only when the *criterial* requirements of the operant are not satisfied.) In all of the experiments, noncriterial patterns within invalid operants during the test session showed significantly higher average antiquity than those typed within valid operants. This effect was also observed, to a lesser degree, during the participants' learning sessions. Figure 3 shows the median for the average noncriterial pattern antiquity for all 37 participants, both during the test session and during the last learning session for comparison. The error bars show the interquartile ranges. For all individual data, see Table 3.

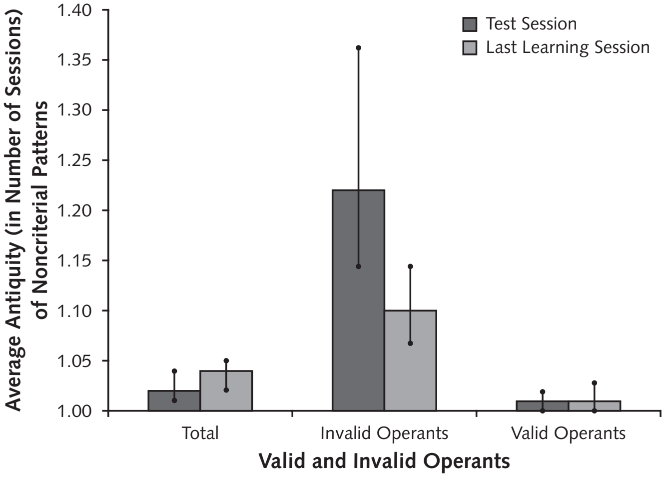

Figure 3. Comparison of average antiquity levels of noncriterial patterns (consisting of 3-keystroke sequences) within valid and invalid operants emitted during the test session and during the last learning session; median of all participants with interquartile ranges.

This effect—much greater resurgence in three–letter noncriterial patterns typed within invalid operants than within valid ones—occurs uniformly across the board for all 37 participants, and is statistically significant at the .001 level, $t(72) = 7.01$, $p < .001$. The increase in the level of resurgence in invalid operants in the test session over that in invalid operants in the last learning session is also significant at that level, $t(72) = 4.95$, $p < .001$.

Accompanying this increase in noncriterial pattern antiquity associated with invalid operants is a significantly higher percentage of brand–new noncriterial patterns (ones that have never previously appeared in that participant's history) observed in invalid operants, when compared with those that occur in valid operants. Again, this effect was seen consistently in all experiments; Figure 4 shows the median percentage of new patterns for all participants in all experiments during the test session and during the last learning session. The error bars show the interquartile ranges; there is considerably more variation among individual participants for this data than for the noncriterial antiquity data—and considerably more individual variation during the test session when compared with the learning sessions, but the difference is still significant at the .001 level, $t(72) = 4.88$, $p < .001$. For the individual data, see Table 4.

It should be noted that the occurrence of greater numbers of both old and brand–new noncriterial patterns in invalid operants is not merely the result of a general increase in variability. Although participants do use a greater variety of noncriterial patterns in invalid operants, overall they are still using only a small fraction of the 1320 possible unique three–letter patterns that could be formed using the twelve available letters.

Table 3
Individual Noncriterial Pattern Resurgence Data, All Experiments

Experiment & Participant Number	Final Test Session			Last Learning Session		
	Total	Incorrect Operants	Correct Operants	Total	Incorrect Operants	Correct Operants
1 - 221	1.02	1.31	1.00	1.01	1.04	1.01
1 - 222	1.24	1.40	1.23	1.04	1.11	1.01
1 - 223	1.00	1.11	1.00	1.02	1.11	1.00
1 - 224	1.02	1.20	1.00	1.03	1.13	1.00
1 - 225	1.03	1.39	1.01	1.04	1.15	1.01
1 - 226	1.04	1.26	1.02	1.02	1.01	1.02
1 - 227	1.36	1.69	1.32	1.08	1.15	1.07
2 - 241	1.11	1.35	1.07	1.20	1.22	1.19
2 - 242	1.02	1.11	1.00	1.02	1.04	1.00
2 - 243	1.00	1.10	1.00	1.03	1.06	1.00
2 - 244	1.19	1.67	1.10	1.11	1.10	1.11
2 - 245	1.06	1.16	1.02	1.05	1.08	1.00
2 - 246	1.00	1.00	1.00	1.29	1.16	1.40
2 - 247	1.04	1.46	1.00	1.01	1.01	1.01
3 - 251	1.02	1.13	1.01	1.01	1.02	1.01
3 - 252	1.04	1.16	1.03	1.03	1.06	1.01
3 - 253	1.03	1.21	1.01	1.04	1.05	1.04
3 - 254	1.02	1.25	1.01	1.01	1.03	1.01
3 - 255	1.01	1.13	1.00	1.03	1.05	1.00
3 - 256	1.04	1.38	1.02	1.11	1.16	1.08
3 - 257	1.02	1.25	1.00	1.01	1.02	1.00
4 - 281	1.02	1.22	1.01	1.04	1.11	1.01
4 - 282	1.05	1.52	1.02	1.04	1.13	1.00
4 - 283	1.00	1.09	1.00	1.03	1.08	1.01
4 - 284	1.04	1.30	1.02	1.13	1.24	1.10
4 - 285	1.02	1.26	1.01	1.04	1.10	1.03
4 - 286	1.11	1.36	1.09	1.17	1.33	1.10
4 - 287	1.02	1.22	1.01	1.05	1.11	1.01
5 - 291	1.29	1.67	1.26	1.05	1.09	1.02
5 - 292	1.06	1.29	1.00	1.05	1.11	1.02
5 - 293	1.00	1.00	1.00	1.02	1.03	1.01
5 - 294	1.01	1.18	1.00	1.03	1.12	1.00
5 - 295	1.02	1.06	1.01	1.04	1.11	1.02
5 - 296	1.01	1.14	1.00	1.02	1.06	1.01
5 - 297	1.01	1.15	1.00	1.08	1.22	1.04
5 - 298	1.00	1.00	1.00	1.01	1.04	1.00
5 - 299	1.03	1.44	1.00	1.03	1.14	1.00
Median	1.02	1.22	1.01	1.04	1.10	1.01
1st Quartile	1.01	1.13	1.00	1.02	1.05	1.00
3rd Quartile	1.04	1.36	1.02	1.05	1.13	1.03

Table 4
Individual New Noncriterial Pattern Data, All Experiments

Experiment & Participant Number	Final "Test" Session			Last Learning Session		
	Total	Incorrect Operants	Correct Operants	Total	Incorrect Operants	Correct Operants
1 - 221	1.32	11.96	0.72	1.10	3.82	0.27
1 - 222	5.77	24.58	2.91	1.07	3.03	0.27
1 - 223	0.50	10.84	0.00	0.86	3.96	0.00
1 - 224	0.34	3.38	0.06	0.37	1.65	0.00
1 - 225	1.16	14.17	0.24	1.46	6.05	0.27
1 - 226	1.71	15.38	0.54	1.87	5.10	0.75
1 - 227	1.26	3.47	0.97	0.79	0.71	0.82
2 - 241	1.73	1.28	1.79	2.82	3.07	2.06
2 - 242	0.23	1.48	0.00	0.47	0.86	0.07
2 - 243	0.12	2.33	0.00	1.61	3.39	0.07
2 - 244	4.21	16.94	1.17	1.80	2.29	1.42
2 - 245	1.16	4.47	0.00	1.82	2.63	0.52
2 - 246	0.06	0.00	0.06	29.00	24.12	32.56
2 - 247	2.40	19.90	0.13	1.02	3.72	0.00
3 - 251	0.53	3.39	0.23	1.17	1.66	0.37
3 - 252	1.63	13.27	0.24	0.93	1.91	0.22
3 - 253	3.25	16.89	1.34	4.50	4.74	4.38
3 - 254	0.40	4.49	0.18	0.67	2.02	0.07
3 - 255	0.00	0.00	0.00	0.69	1.01	0.28
3 - 256	0.96	2.06	0.90	5.23	7.12	4.38
3 - 257	0.17	1.92	0.00	0.32	0.85	0.00
4 - 281	0.35	4.55	0.18	1.87	5.67	0.27
4 - 282	1.52	14.16	0.66	2.71	6.31	0.99
4 - 283	0.00	0.00	0.00	1.18	2.33	0.74
4 - 284	0.36	3.42	0.11	1.50	4.11	0.68
4 - 285	0.52	5.56	0.13	0.61	2.19	0.07
4 - 286	2.59	4.32	2.46	2.42	3.16	2.08
4 - 287	20.92	38.93	19.28	2.90	6.69	0.14
5 - 291	1.83	16.56	0.49	2.16	4.86	0.14
5 - 292	1.78	7.91	0.07	1.72	4.39	0.27
5 - 293	0.00	0.00	0.00	0.51	2.32	0.14
5 - 294	0.23	5.80	0.00	0.72	2.87	0.00
5 - 295	5.82	21.19	4.80	5.82	12.91	3.87
5 - 296	0.74	11.61	0.00	1.87	5.78	0.21
5 - 297	0.23	5.88	0.06	1.37	4.29	0.34
5 - 298	0.00	0.00	0.00	0.76	2.07	0.26
5 - 299	0.11	0.98	0.06	0.69	3.09	0.00
Median	0.74	4.55	0.13	1.37	3.16	0.27
1st Quartile	0.23	2.06	0.00	0.76	2.19	0.07
3rd Quartile	1.73	14.16	0.72	1.87	4.86	0.75

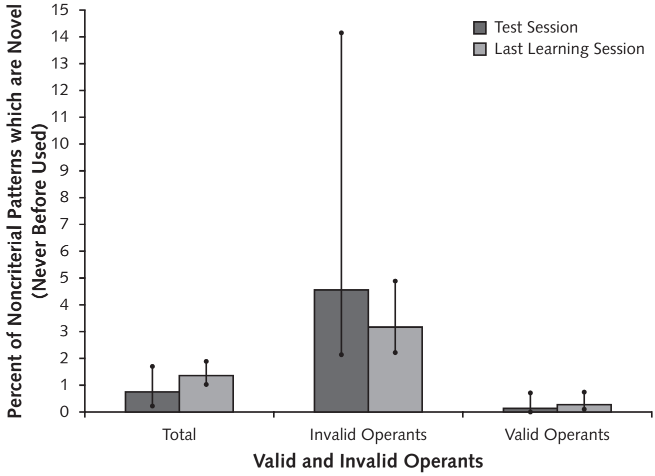

Figure 4. Comparison of percentages of brand-new (never previously seen) 3-letter noncriterial patterns within valid and invalid operants emitted during the test session and during the last learning session; median of all participants with interquartile ranges.

Discussion

These experiments show that the criterial and noncriterial features of operants can be affected in different ways by certain prevailing contingencies and other variables, and that noncriterial features of operants can show effects, such as resurgence, that are separate from changes in the criterial features. These kinds of differences between criterial and noncriterial features cannot be observed when the operant is recorded only as an all–or–none digital event. It is this differentiation potentiality that makes the revealed operant preparation a useful tool for studying the properties of individual occurrences of operants.

Regarding the effects of number of prior repetitions on later performance, operants that were performed most frequently during the learning sessions were also performed far more often in the test session, regardless of the ratio of their number of prior repetitions to those of the less frequently repeated ones. For this particular type of operant there may be a minimum number of prior repetitions (somewhere between 468 and 936) that the most often repeated operants must reach before they gain a significant advantage in the test session. It is to be assumed that the minimum number of repetitions would be different for every type of operant and experimental procedure; for instance, simpler operants may have lower repetition thresholds than more complex ones.

The threshold may also depend on the strength of possible pre–experimental biases for or against specific operants, with the threshold number being the number of repetitions needed to overcome such biases. Operant bias has been observed in other studies that required equivalent operants (Jones & Mechner, 2013). Such biases could also explain the large differences among participants in the frequency of use of the medium–repetition category. It should be noted that the issue of pre–experimental bias is an important methodological one in research on history variables, because if comparisons among operants that reflect different experimentally installed histories are to be valid, the operants must be pre–experimentally equivalent.

The motor program literature suggests an additional explanation—one based on automatization: when certain behavioral sequences are repeated often enough, they become more "automatized" and less susceptible to the effects of other variables (Mechner, 1995; Mechner et al., 1997; Schneider, 1985).

As stated earlier, resurgence, as measured by average noncriterial antiquity levels, was much greater in invalid operants than in valid ones. Invalid operants also contained greater numbers of novel (never–before–seen) noncriterial patterns. Both of these effects were also observed by Mechner & Jones (2011), and are also in line with experiments that have found that extinction increases not only variability but also the number of new responses (Neuringer, Kornell, & Olufs, 2001).

Although these effects were present to some degree in all sessions, they were most pronounced in the final test session. A plausible explanation for this is that the test session entailed an elevated level of stress (the time pressure and accuracy contingencies), suggesting that it was this increased stress that caused the increases in the frequency of invalid operants, which are, in effect, mistakes. When viewed in the context of the well–established observation that one of the features of increased automatization (usually the result of multiple repetitions of an operant), is a diminished susceptibility to the effects of stress and other variables, and diminished mistakes, this finding suggests the possibility that many types of mistakes, as in the present data and perhaps in general, are instances of resurgence, as discussed in Mechner (1995).

Due to the nature of the procedure used in the five experiments, most invalid operants were clustered at the beginning of a new "block" of required operants. In the learning sessions, this was the point at which the required operant changed and the participant had to try all the options before finding the new required operant. Invalid operants also occurred frequently in the "choose 1 of 3" type of test session whenever a given test block ended and the participant was presented with a new set of three operants from which to choose the one to perform. This indicates that resurgence—associated with disruptions in automatized behavior—can occur as a result of any change in the type or amount of work required, whether that change is unsignaled and unanticipated as in the case of extinction, or explicitly signaled and anticipated, as it is here. The observed resurgence may be due to the stress, even if minor, that the inherent challenges of such changes in the work routines may entail.

The experiments show that the operant an individual chooses to perform can be a function of the frequency with which the operant had previously been repeated.

They also show that resurgence of operants' noncriterial features is associated with stress conditions. The methodology demonstrates how the revealed operant preparation can be used to study criterial and noncriterial features of operants, as well as other issues involving the characteristics of individual occurrences of operants. The results provide several examples of the ways criterial and noncriterial features of operants can respond differentially to environmental influences.

The present research conceptualizes resurgence as the reappearance of earlier, nonrecent behavior, whose reappearance is due to factors other than restoration of conditions under which the behavior was previously effective. The experimental analysis program that this conceptualization suggests would have to address the following research questions:

- Does resurgence occur only when recently effective behavior no longer produces its effect, or also when the required effort, time, or the motivational state has changed?
- Does it occur when behavior has aversive consequences?
- Is resurgence augmented by various types of stress?
- Does it occur when automatized behavior is interrupted or disrupted?
- What are the similarities and differences between resurged behavior and its antecedent forms?
- What variables determine *which* earlier behavior resurges?
- Do these variables include prior frequency of occurrence, recency of occurrence, motivational modality, automaticity, or other variables?
- When successive occurrences of earlier behavior don't work, does resurgence tap into behavior of ever greater antiquity—a kind of regression effect?
- Are the normal minor deviations from practiced routines and noticeable errors made during performance of skilled behavior due to resurgence events (Mechner, 1995)?
- A conceptual issue: is it productive to categorize history variables as ones that *shape* future behavior (e.g., learning history factors), and those that *modulate the frequencies and circumstances of future occurrences* of the behavior?

These and many related questions would then define a widened frontier of resurgence research.

References

Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology, 42*, 273–281.

Bilbrey, J. L., Patterson, D. D., & Winokur, S. (1973). Maintenance and autoshaping of keypecking in undeprived pigeons. *Bulletin of the Psychonomic Society, 2*, 394–396.

Dixon, M. R., & Hayes, L. J. (1998). Effects of differing instructional histories on the resurgence of rule–following. *Psychological Record, 48*, 275–292.

Dyal, J. S., & Holland, T. A. (1963). Resistance to extinction as a function of the number of reinforcements. *American Journal of Psychology, 76*, 332–333.

Epstein, R. (1985). Extinction–induced resurgence: Preliminary investigations and possible applications. *Psychological Record, 35*, 143–153.

Farthing, G. W., & Hearst, E. (1968). Generalization gradients of inhibition after different amounts of training. *Journal of the Experimental Analysis of Behavior, 11*, 743–752.

Hearst, E., & Koresko, M. B. (1968). Stimulus generalization and the amount of prior training on variable–interval reinforcement. *Journal of Comparative and Physiological Psychology, 85*, 171–178.

Jensen, G. D. (1963). Preference for bar pressing over "freeloading" as a function of number of rewarded presses. *Journal of Experimental Psychology, 65*, 451–454.

Jensen, G. D., Leung, C. M., & Hess, D. T. (1970). "Freeloading" in the Skinner box contrasted with freeloading in the runway. *Psychological Reports, 27*, 67–73.

Jones, L. D., & Mechner, F. (2013). Systematic operant bias observed in human participants during research on choice. *European Journal of Behavior Analysis, 14,* 295–311.

Kass, N., & Wilson, H. (1966). Resistance to extinction as a function of percentage of reinforcement, number of training trials, and conditioned reinforcement. *Journal of Experimental Psychology, 71*, 355–357.

Lentz, B. E., & Cohen, S. L. (1980). The effect of prior training on the contrafreeloading phenomenon. *Bulletin of the Psychonomic Society, 15*, 48–50.

Lewis, D. J., & Duncan, C. P. (1958). Expectation and resistance to extinction of a lever–pulling response as a function of percentage of reinforcement and number of acquisition trials. *Journal of Experimental Psychology, 55*, 121–128.

Lovejoy, E. (1966). Analysis of the overlearning reversal effect. *Psychological Review, 73*, 87–103.

Mackintosh, N. J. (1965). Overtraining, reversal, and extinction in rats and chicks. *Journal of Comparative and Physiological Psychology, 59*, 31–36.

Mackintosh, N. J. (1969). Further analysis of the overtraining reversal effect. *Journal of Comparative and Physiological Psychology, 67*, 1–18.

McAllister, M. J., Capehart, J. E., & Rogers, C. A. (1970). A transfer approach to the ORE. *Perceptual and Motor Skills, 30*, 971–977.

Mechner, F. (1992). The revealed operant: A way to study the characteristics of individual occurrences of operant responses. S. Glenn (Ed.) *Monograph Series*. Cambridge, MA: The Cambridge Center for Behavioral Studies.

Mechner, F. (1995). *Learning and practicing skilled performance*. New York, NY: The Mechner Foundation (http//www.mechnerfoundation.org/).

Mechner, F., Hyten, C., Field, D. P., & Madden, G. (1997). Using revealed operants to study the structure and properties of human operant behavior. *Psychological Record, 47*, 45–68.

Mechner, F., & Jones, L. D. (2011). Effects of sequential aspects of learning history. *Mexican Journal of Behavior Analysis, 37,* 109–138.

Morgan, L., & Neuringer, A. (1990). Behavioral variability as a function of response topography and reinforcement contingency. *Animal Learning and Behavior, 18*, 257–263.

Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior Processes, 27*, 79–94.

Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 429–452.

Perin, C. T. (1942). Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *Journal of Experimental Psychology, 30*, 93–113.

Pittenger, D. J., Pavlik, W. B., Flora, S. R., & Kontos, J. M. (1988). The persistence of learned behavior in humans as a function of changes in reinforcement schedule and response. *Learning and Motivation, 19*, 300–316.

Reid, L. S. (1953). The development of noncontinuity behavior through continuity learning. *Journal of Experimental Psychology, 46*, 107–112.

Rilling, M., & Budnik, J. E. (1975). Generalization of excitation and inhibition after different amounts of training of an avoidance baseline. *Journal of the Experimental Analysis of Behavior, 23*, 207–215.

Schneider, W. (1985). Toward a model of attention and the development of automatic processing. In M. I. Posner, & O. S. M. Marin (Eds.), *Attention and Performance XI* (pp.475–492). Hillsdale, NJ: Lawrence Erlbaum Associates.

Senkowski, P. C. (1978). Variables affecting the overtraining extinction effect in discrete–trial lever pressing. *Journal of Experimental Psychology: Animal Behavior Processes, 4*, 131–143.

Sewell, W. R., & Nickel, R. (1979). Effect of length of training on stimulus control using a low density reinforcement schedule. *The Psychological Record, 29*, 401–408.

Stolz, S. B., & Lott, D. F. (1964). Establishment in rats of a persistent response producing a net loss of reinforcement. *Journal of the Experimental Analysis of Behavior, 24*, 17–21.

Thomas, D. R., & Williams, J. L. (1963). A further study of stimulus generalization following three–stimulus discrimination training. *Journal of the Experimental Analysis of Behavior, 6*, 171–178.

Tombaugh, T. N. (1967). The overtraining extinction effect with a discrete–trial bar press procedure. *Journal of Experimental Psychology, 73*, 632–634.

Wilson, K. G., & Hayes, S. C. (1996). Resurgence of derived stimulus relations. *Journal of the Experimental Analysis of Behavior, 66*, 267–281.