EFFECTS OF SEQUENTIAL ASPECTS OF LEARNING HISTORY

EFECTOS DE LOS ASPECTOS SECUENCIALES DE LA HISTORIA DE APRENDIZAJE

FRANCIS MECHNER AND LAURILYN JONES

THE MECHNER FOUNDATION, NEW YORK, USA

Abstract

In a series of five experiments, a number of similar operant classes, consisting of keystroke sequences on a computer keyboard, were learned and practiced in succession by human subjects. Each experiment consisted of learning sessions spread over several days, separated by either elapsed time or interpolated sessions in which unrelated but similar operant classes were performed. The learning sessions were followed by a final test session in which the subjects were required to choose and perform one from presented sets of three operant classes. The test was designed to be stressful by the imposition of time pressure and certain other contingencies. In the test session, preference was commonly shown for operant classes from the first- and/or last-learned groups—termed primacy and recency effects respectively—with minimal preference for the middle groups. Most subjects showed either primacy or recency effects, and relatively few showed both; the subjects that showed mainly recency effects also made the largest number of errors during initial learning of the last set of operant classes. In addition, certain noncriterial characteristics of these operants were measured. These revealed other effects, in particular the association of performance errors with both greater resurgence of older behavior patterns and greater numbers of new behavior patterns.
Key words: Learning history, resurgence, sequential effects, primacy, recency, criterial and noncriterial features, operant classes, humans.

110

Resumen

En una serie de cinco experimentos, los participantes aprendieron y practicaron en sucesión un número de clases operantes equivalentes, consistentes en secuencias de presiones de teclas en un teclado de computadora. Cada experimento consistió en sesiones de aprendizaje distribuidas a lo largo de varios días, separadas por tiempo o por sesiones interpoladas en las que se ejecutaron clases operantes no relacionadas, pero similares. Las sesiones de aprendizaje fueron seguidas por una sesión final de prueba en la que los participantes debían escoger y ejecutar una de tres conjuntos de clases operantes presentadas. La prueba se diseñó para generar estrés vía la imposición de presión temporal y de otro tipo de contingencias. En la sesión de prueba, con frecuencia se exhibió preferencia hacia las clases operantes de los grupos iniciales o finales aprendidos – efectos de primacía y de recencia, respectivamente- con una preferencia mínima hacia los grupos medios. La mayoría de los participantes mostró efectos de primacía o de recencia, y relativamente menos personas mostraron ambos; las personas que mostraron efectos de recencia también mostraron el mayor número de errores al aprender el primer conjunto de clases operantes. Adicionalmente se midieron ciertas características que no formaron parte de criterio de esas operantes. Esto reveló otros efectos, en particular la asociación de los errores con una mayor resurgencia de patrones viejos, así como con un mayor número de patrones de comportamiento novedosos.
Palabras clave: Historia de aprendizaje, resurgencia, efectos secuenciales, primacía, recencia, aspectos criterio y no criterio, clases operantes, humanos.

A general issue in the study of learning is how history variables affect current behavior. Such variables are of interest to researchers because effects of a subject's prior history can unexpectedly re-appear during an experiment (Tatham & Wanchisen, 1998; Wanchisen, 1990), an effect that has sometimes been termed "resurgence." In general, much of any organism's current behavior may have roots in the details of the learning history. For example, much of the variability in any behavior stream, including deviations from practiced routines and errors made during performance of skilled behavior, may be due to the precise way the behavior was learned and practiced (Mechner, 1995).

The five studies reported in this paper were designed to examine the effects of one basic history variable -- the order in which different classes of operant responses are learned -- on those operants' subsequent properties, particularly the relative frequencies with which they were chosen during a forced-choice test session which followed the learning sessions. The rationale for these experiments was to study the effects of learning and practice variables on later performance under more stressful conditions.

These experiments can be thought of as analogous to the large body of work on serial-position effects (though the current experiments use a very different depen-

dent variable than those normally used in the serial-position literature). The serial-position effect is the name commonly given to the general observation that when items are learned in a particular order, there is better recall or recognition of those at the beginning and/or end of the series than of those in the middle. The left peak of the resulting U-shaped curve has often been called the primacy effect and the right peak the recency effect (Crowder, 1976; McGeoch & Irion, 1952).

The present experiments, though they also involve serial learning, differ from those reported in the literature in that they examined the relative frequency with which previously learned classes of operant responses are chosen and physically performed, whereas the vast majority of the traditional studies deal with simple recall and/or recognition of stimuli. The serial learning literature is invoked here not for its comparability to the present line of research but for its prominence. The experiments described in this paper looked at a different kind of primacy and recency effect.

Serial-position effects are easily altered through methodological manipulations (see Wright, 1998, for a comprehensive review). The specific design elements that differentiate the five experiments reported in this paper – specifically, varying the amount of time allowed to elapse between sessions spent learning the different operant classes, and/or inserting sessions spent learning a different, but similar, task between those sessions – are similar in concept to those frequently used in the traditional serial-learning literature. In addition, some researchers contend that the two parts of the serial-position effect represent two different processes (Atkinson & Shiffrin, 1968; Glanzer & Cunitz, 1966; Jensen, 1962; Rundus & Atkinson, 1970; Waugh & Norman, 1965), while others posit a single process underlying both primacy and recency (Bjork & Whitten, 1974; Crowder & Neath, 1990; Hull, 1935; Lepley, 1934; Murdock, 1960; Wright, 1998).

The dependent variable used in the current experiments is also related to the literature on resurgence, an important topic in the study of the effects of learning history. The descriptive term resurgence is used here to refer to the reappearance of earlier behavior under certain conditions (although experimental resurgence has often been defined in the literature as a phenomenon that occurs during extinction, e.g., Epstein, 1985). But Epstein also cited punishment, satiation, and increased response requirements as conditions that may induce resurgence. Mechner (1994) discusses resurgence in the context of the reappearance of previously punished behavior, while resurgence of derived relations has been demonstrated under restricted choice conditions in research on equivalence classes (Wilson & Hayes, 1996). Bachá-Mendéz, Reid and Mendoza-Soylovna (2007) have shown extinction-induced resurgence of response sequences, rather than just single lever presses. Mechner, Hyten, Field and Madden (1997) showed that resurgence occurs in extinction and also when the performance requirement is abruptly increased. The stress contingencies used during the test sessions of the present experiments were intended to induce or magnify resurgence effects.

A full analysis of resurgence, as well as the effects of serial operant learning and performance variables, requires the separate tracking and examination of the criterial and noncriterial features of every occurrence of the operants involved. All operant occurrences have both criterial and noncriterial features. The criterial ones are those that must be present for the operant to be considered as having occurred (for example, the distance a rat must depress a lever), while the noncriterial aspects are all the other characteristics of that operant, including topographic ones and ones involved in "superstitious" behavior (Herrnstein, 1966; Mechner, 1994). Guthrie (1959) called them sub-responses; Schoenfeld (1961) referred to sub-categories making up the generic form of the operant. Criterial and noncriterial features of operants exhibit different properties. For example, variability in noncriterial dimensions of an operant has been shown to decrease after successive conditioning-extinction cycles (Antonitis, 1951). Di Lollo, Ensminger and Notterman (1965) showed that criterial and noncriterial aspects of an operant were affected differentially by amount of reinforcement.

In order to address the question of how the present independent variables affect not only the subsequent strength of operant classes, but also their noncriterial features, several distinct but equivalent operant classes are needed. Such operant classes were defined by keystroke sequences whose beginning and end is behaviorally marked, using Mechner's (1994) revealed operant procedure. Even though the operant as a whole is comprised of more than a single switch closure, the entire sequence of actions, usually executed, after practice, in less than three seconds, can be regarded as a single operant. Because all features of each occurrence -- criterial as well as noncriterial -- are recorded and readily accessible for analysis, the type of operant used here can be regarded as an operant viewed under magnification, a functional unit of behavior composed of subunits, some of which can vary to some extent. Variations in different individual occurrences are distinguishable so that instances are recognizable and separately identifiable. While the present studies could also have been conducted with other kinds of response sequences or units, including ones that do not meet the criteria for the definition of an operant, the advantage of using an operant as the unit is that the results obtained are then generalizable to other types of operant classes.

The present independent variables are applicable only to the criterial features of the operants, as the noncriterial features are, by definition, not under experimental control. Also, different approaches are needed to measure and examine data pertaining to criterial and noncriterial features.

A word on terminology: The term "operant" will be used for individual occurrences of the described response sequences, and the term "keystrokes" will be used for the sub-operants that comprise the operants. The term "operant class" will be used for the various generic operants as these are defined by their specific criterial attributes.

Experiment 1

Method

Subjects
Thirteen subjects -- mostly university students of varying ages, male and female -- were recruited through flyers posted on local university campuses.

Apparatus
The experimental room contained four computer workstations separated by screens. Each of the computer keyboards was fitted with a particleboard "mask" that covered all the keys except for those used in the experiment: twelve character keys (tyuighjkvbnm), the space bar, the enter key, the number keypad, and four function keys (See Figure 1 below).

Figure 1. Apparatus covering the computer keyboard during the experiments

Procedure

Nine different operant classes were learned and practiced by the subjects during nine learning sessions, followed by a final test session. Subjects completed one session per day on ten scheduled days within a 20-day period, with all of each subject's sessions taking place at the same time each day. They were paid $10 per session completed and, in addition, could earn an additional amount during the test session (of up to $200, depending on performance). They signed consent forms agreeing to keep caffeine consumption, meal schedules and amount of sleep consistent from day to day during the course of the experiment, to refrain from alcohol or drugs, and to arrive at the specified time.

At the beginning of the first session, the computer monitor displayed the message "Patterns 1, 2, 3. Start with 2." and the experimenter instructed the subjects that they would be learning nine different "patterns of letters" during the study. Subjects were shown how to press a given pattern's identification number on the number keypad in order to display the letters for that pattern on the computer screen. The experimenter then demonstrated how each of the operant classes (or "patterns) were to be performed: by pressing the space bar, followed by the first three letters of the pattern, followed by six or more letters of the subject's choice, followed by the last three letters of the pattern, followed by the enter key. A green square approximately 4 by 6 inches appeared on the screen for 500 milliseconds after the completion of a correct pattern. They were further instructed to repeat the same pattern until a message on the screen told them to switch patterns.

Thus the nine operant classes used each consisted of 14 or more keystrokes, counting the space bar at the beginning and the enter key at the end. The first three and last three character keys required for each operant were mandated (criterial), and defined a unique operant class. Keystrokes that repeated either of the two immediately preceding ones did not count toward the six or more additional discretionary keystrokes required between the first and last three mandated ones.

The monitor screen color turned from black to blue at the instant the subject initiated an execution of the operant by pressing the space bar, and remained blue until the operant was completed with the press of the enter key. While the screen was blue, acceptable keystrokes produced a subtle "click" feedback noise 100 milliseconds in duration. At no point did the monitor display the characters typed by the subjects.

During the learning sessions, operants from only one of the nine operant classes were followed by the green square within any given "block" of 35 consecutive occurrences of the scheduled operant class. Subjects were free to type anything, but any keystrokes other than the specific ones that comprised the operant class in use would be followed only by the black screen.

The required operant class was switched after every block of 35 valid operants completed by the subject. An operant was considered valid if it fulfilled the definitional criteria for the operant class called for by the program at the time. Invalid

operants were not counted toward the 35 required to complete a block. The sequence of the blocks of different operant classes during a session was programmed to be unpredictable to the subjects.

After the first session the identification numbers on the screen were phased out, although the computer was still programmed to switch the required operant class after every block of 35 valid operants. Subjects were then instructed to try each of the three patterns to find the one that produced the green square each day, and then continue with it until a switch was prompted.. At the beginning of each successive session, a message displaying the identification numbers of the three operant classes to be practiced that day continued to appear on the computer screen.

The nine operant classes were divided into three groups of three each, with the only difference between the groups being the order with which they were learned. Only one group was required during any given learning session. Operant classes 1, 2 and 3 (the first group learned) were programmed to produce the green square during sessions 1, 2 and 3. Likewise, operant classes 4, 5 and 6 (the middle group) were programmed during sessions 4, 5 and 6; and operant classes 7, 8 and 9 (the last group) were programmed for sessions 7, 8 and 9. Each learning session thus consisted of four blocks for each of the three operant classes in use during that session, with 35 required valid operants per block, for a total of 12 blocks and 420 valid operants per session.

Subjects completed one session on each of 10 specified days within a 20-day period. Sessions 1, 2 and 3 took place on days 1, 2 and 3, respectively. There followed a break of nine days without a session. Sessions 4, 5 and 6 were scheduled for days 13, 14 and 15. A single day without a session separated sessions 4, 5, and 6 from sessions 7, 8 and 9, which took place on days 17, 18 and 19. The 10th or test session then followed immediately on day 20. This schedule was chosen so that, on the final day, the number of days that had elapsed since the middle of the three sessions in which each group of operant classes was practiced followed a geometric 1:3:9 ratio (2, 6, and 18 days respectively). By the final session, each of the nine operant classes had been required precisely 420 times. The only difference between the learning histories for the three groups of operant classes was the recency with which they had last been used.

At the beginning of the test session the subjects were instructed that they would be tested on all nine patterns, which would be displayed three at a time. They could type any of the three patterns available at any given time, earning 65 cents for each valid operant while losing 35 cents every time they made a mistake, paused for too long between keystrokes or used the look-up function.

The test session was divided into 24 blocks of 20 operants each. Each set of three reinforceable operant classes (one from each recency group) was displayed on the computer screen at the beginning of each block. After every block of 20 performed operants (regardless of whether or not they were valid, or from these three classes) the computer displayed a new set of three reinforceable operant

classes, again always one from each of the three recency groups.

In the test session, the green square was never presented. Instead, the message "You just earned 65 cents. Ring it up." appeared on the monitor screen after each valid operant, accompanied by a 500-millisecond high-pitched tone, and the subject was required to type that amount on the number keypad and press the enter key, whereupon 65 cents was added to the subject's total earnings, which were displayed in the upper left corner of the screen at all times.

If a keystroke sequence did not match any of the three operant classes currently allowed, the computer emitted a 500-millisecond low-pitched tone and subtracted 35 cents from the subject's total (without requiring any action on the part of the subject). If the subject paused too long between keystrokes at any time, the same tone and loss of 35 cents occurred. The length of the pause allowed without penalty was programmed to be five times each subject's average inter-keystroke time over his or her 20 preceding operants (an average of 1 to 2.5 seconds). The pause allowed was thus reset every 20 operants. If the subject pressed any of the identifi-
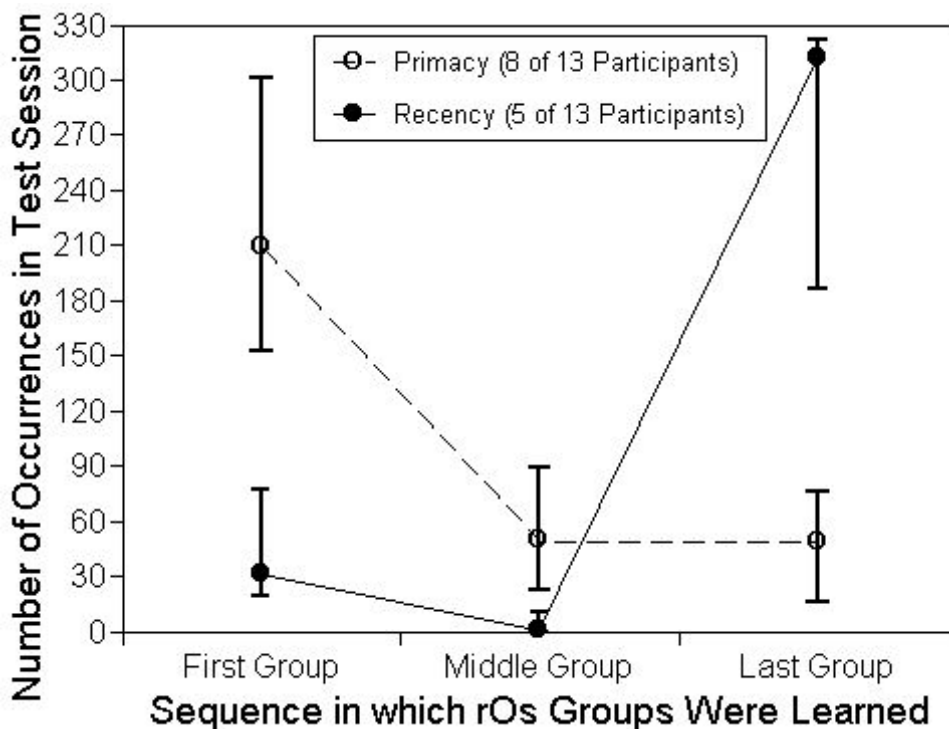

Figure 2. Operants from the given sequential group performed by subjects during the test session, Experiment 1

cation numbers on the number keypad, the computer displayed the keystrokes for the corresponding operant class on the monitor screen but simultaneously emitted the low-pitched tone and subtracted 35 cents from the subject's total. The test session ended after 480 operants, including both valid and invalid ones.

In the test session, each set of operant classes was presented as an option the same number of times, in a sequence designed to be unpredictable to the subjects.

Resutls

During the test session, 7 of the 13 subjects performed operants from the first-learned group more frequently than those from the middle- and last-learned groups, while five other subjects performed operants predominantly from the last-learned group. The remaining subject performed operants from the middle group most often.

Figure 2 shows the number of test session occurrences of operant classes from each of the three groups. Separate medians and interquartile ranges are displayed for the subjects whose results tended to favor the first group (primacy) and those who tended to favor the last group (recency), in order to show that the subjects fell into two distinct groups. (The subject who preferred the middle group overall -- subject AE -- is included in the primacy median value because he also favored the first group over the last group.) Individual data is shown below in Table 1.

The large separation and non-overlap of the primacy and recency groups' interquartile ranges make the point. Regardless of which of the three groups of operants was performed most often, operants of the other two groups generally occurred a much smaller number of times, resulting in an almost uniformly low level for the middle group.

The various effects seen in the subjects' noncriterial keystrokes were remarkably consistent across all five experiments. Therefore those results are presented in the "Combined Results" section later in this document, after the individual experiments' sections.

Experiment 2

The purpose of this experiment was to determine how the effect of interposing another, similar task between groups of operant classes learned would compare to the effect of interposing elapsed time, as in Experiment 1.

Method

Seven new subjects were recruited. The recruitment method, equipment, and procedure were the same as in Experiment 1, but instead of using elapsed time (days without a session) to separate the sessions spent learning the three groups of oper-

Table 1.
Number of operants performed from given sequential group during test session, Experiment 1 (individual data

Subject	First Group	Middle Group	Last Group
Showing	*Predominantly*	*Primacy*	
EW	153	81	76
JJA	321	1	19
JJY	226	58	55
SP	337	0	0
SSW	193	30	112
TB	294	42	7
TRD	151	116	77
Showing	*Predominantly*	*Recency*	
BTH	77	90	173
IAA	19	0	322
IR	127	11	186
SJB	31	1	312
SJM	19	0	324
Showing Neither	*Showing Neither*	*Showing Neither*	*Showing Neither*
AE	75	169	42

ant classes, a single session in which subjects learned a new, different group of operant classes (defined by different sets of mandated keystrokes) was interposed between sessions 3 and 4, and another between sessions 6 and 7. The additional operant classes were not presented as choices during the test session (although subjects were not informed of this fact in advance).

Experiment 2 thus consisted of 12 sessions on 12 consecutive days. 15 operant classes were learned. During sessions 1 through 3, the same operant classes 1, 2 and 3 were learned and practiced as in Experiment 1; during session 4, new operant classes A, B and C were required; during sessions 5 through 7, operant classes 4, 5 and 6, again as in Experiment 1; during session 8 new operant classes D, E,

and F; and during sessions 9 through 11, operant classes 7, 8 and 9. Session 12 was the test session, which was exactly the same as that of Experiment 1. On days 4 and 8 a piece of paper with the required keystrokes for either operant classes A, B and C or D, E and F was taped to the bottom of the computer monitor.

Results

In the test session, five of the seven subjects performed operants from the first-learned group far more often than those from the other groups, and one subject performed operants from the last-learned, most recent, group most often.
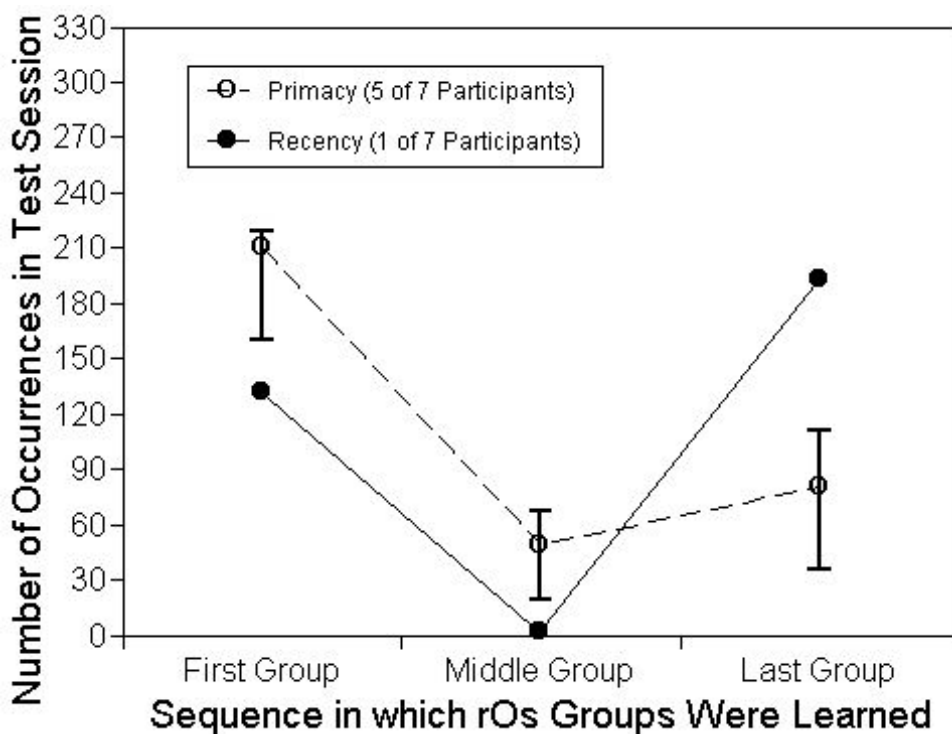

Figure 3. Operants from the given sequential group performed by subjects during the test session, Experiment 2.

Figure 3 shows the number of test session occurrences of operants from each of the three groups when learning was separated by interpolated learning sessions involving different patterns, rather than elapsed time. The median and interquartile ranges for the five subjects demonstrating primacy are shown, as well as the individual data points for the one subject demonstrating recency. The seventh subject performed operants from all three groups almost equally often. His data, along with all the individual data, is in Table 2 below.

Table 2.
Number of operants performed from given sequential group during test session, Experiment 2 (individual data)

Subject	First Group	Middle Group	Last Group
Showing	*Predominantly*	*Primacy*	
APB	210	20	111
FAD	153	78	114
MR	219	49	36
RAS	160	67	81
TDD	346	0	0
Showing	*Predominantly*	*Recency /*	*Both*
SYL	132	2	193
Showing	*Neither*		
JPV	100	94	108

Thus primacy was again the predominant effect, to a much greater degree than in Experiment 1. In fact, in Experiment 2 the preference for operants from the first group over those from the last group is fairly close to statistical significance, with $p = .1065$ ($t(6) = 1.90$, $p < .25$, two-tailed). As in Experiment 1, the middle group almost always occurred the least often during the test session.

Experiment 3

The purpose of this experiment was to determine whether the insertion of additional sessions spent performing another, similar task at the beginning and end of

the study would weaken the primacy and/or recency effects observed for the first- and last-learned operant classes in Experiments 1 and 2, thus potentially flattening the curve.

Method

Ten new subjects were recruited. The recruiting method, equipment, and procedure were the same as in Experiment 2, except that two additional sessions were inserted, one at the very beginning, before the first group of operant classes (1, 2 and 3) was introduced, and the other at the end of the learning sessions, after the last group of three operant classes (7, 8 and 9) but before the test session. In each of the four interpolated sessions, 6 new and different operant classes were learned rather than 3, for a total of 24 new operant classes in addition to the 9 learned in Experiments 1, 2, and 3. Experiment 3 therefore consisted of 14 sessions on 14 consecutive days, with sessions 2 through 4, 6 through 8, and 10 through 12 replicating sessions 1 through 9 of Experiment 1, and with session 14 being identical to the test session for Experiment 1. A total of 33 operant classes were learned. None of the 24 new operant classes learned in Sessions 1, 5, 9 and 13 were required or presented as choices during the test session, although once again, subjects were not told this in advance.

Results

During the test session, operants of classes 1, 2 and 3 (the "first group") were performed most often by 9 of the 10 subjects, even though these were no longer actually the first operants learned in the sequence. The tenth subject performed the last group of operant classes most often.

Figure 4 displays the number of test session occurrences of operants from each of the three groups when additional learning sessions devoted to different patterns were inserted at the beginning and end of the experiment. Medians and interquartile ranges are shown for the 9 of 10 subjects who demonstrated recency. Individual data is in Table 3 below.

Once again, primacy prevailed, in this case with operant classes from the first group showing a statistically significant advantage over those from the last group (t(9) = 3.69, p < .01, two-tailed). The middle group again was almost always the lowest.

In Experiments 1 through 3, the accuracy with which operant classes from each of the groups were performed during the final session was not significantly different for the three groups. However, there is a small but statistically significant difference between the accuracy level of the operants performed most often by each subject during the test session (M = 92.11%, SD = 6.10) when compared with the accuracy level of that same subject's less-used group, regardless of operant class
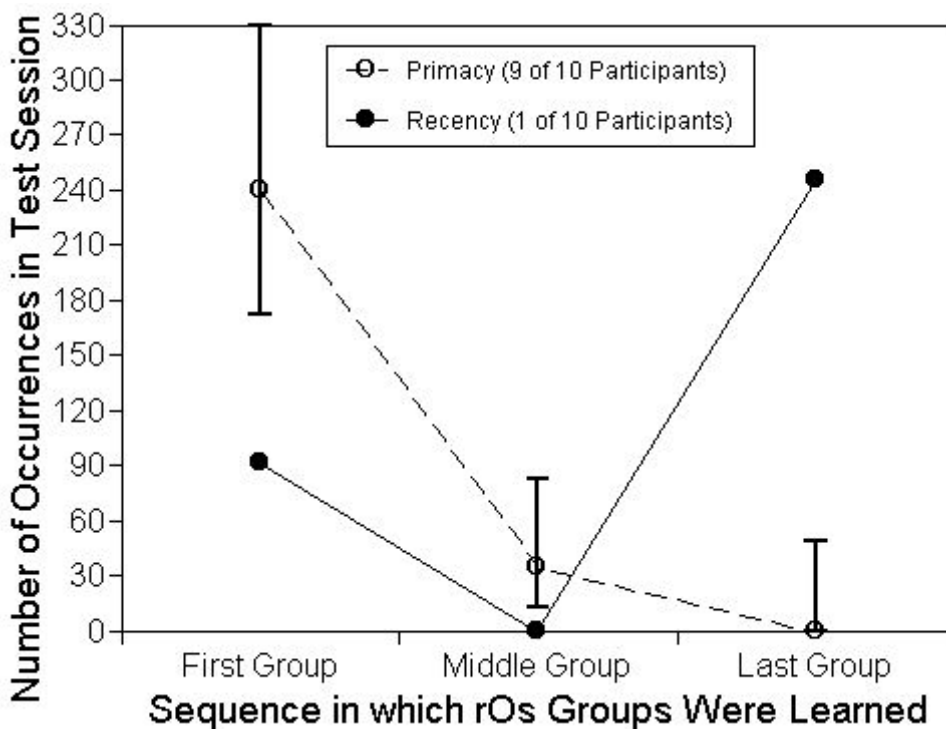
Figure 4. Operants from the given sequential group performed by subjects during the test session, Experiment 3

(M = 86.77%, SD = 10.91). This difference is significant, t(58) = 2.34, p < .05, two-tailed.

Experiment 4

Since the insertion of additional operant classes at the beginning and end failed to produce any flattening of the function in Experiment 3, Experiment 4 was conducted to determine whether giving equal weight during practice to the operant classes learned first and last, and then offering them as test session choices, would flatten the function for the three middle groups. A further purpose of both this experiment and the one that follows was to provide data for a five-point function.

Method

Eight new subjects were recruited. The recruitment method, equipment, and procedure were again the same, except that five groups of three different operant

Table 3.
Number of operants performed from given sequential group during test session, Experiment 3 (individual data).

Subject	First Group	Middle Group	Last Group
Showing	*Predominantly*	*Primacy*	
ADM	172	83	49
JJY	333	13	0
JSW	330	0	0
LA	167	155	16
LBN	319	0	0
LEB	240	40	80
LMA	233	118	0
LMS	336	19	0
MMG	166	35	133
Showing	*Predominantly*	*Recency*	*/ Both*
NVF	91	0	245

classes each were learned, all practiced the same number of times, and all tested during the final session.

This experiment consisted of 11 sessions on 11 consecutive days. A total of 15 operant classes were learned. Operant classes 1, 2 and 3 (the first group) were learned and practiced during sessions 1 and 2; operant classes 4, 5, and 6 (second group) in sessions 3 and 4; operant classes 7, 8 and 9 (third group) in sessions 5 and 6; operant classes 10, 11 and 12 (fourth group) in sessions 7 and 8; and operant classes 13, 14 and 15 (last group) in sessions 9 and 10. By the final session, each of the 15 operant classes had been required exactly 280 times. Session 11 was the test session, at the beginning of which the computer displayed the identification numbers of five different operant classes that would be accepted, always one operant class from each group. For every block of 20 consecutive operants, five different operant classes, one from each group, were accepted, and the numbers of those five were displayed at the start of each block. All other aspects of both learning sessions and the test session were the same as in previous experiments.

Results

In the final test session, five of the eight subjects chose and performed operants from the first-learned group most often, one chose operants from the last-learned group most often, one chose operants from the first and last groups equally often, and one chose operants from the second group most often. With the exception of this last subject, preference levels for the three middle groups were very low.

Figure 5 shows the number of test session occurrences of operants from each of the five groups learned in succession, with medians and interquartile ranges displayed for the 5 out of 8 subjects who exhibited predominantly primacy, and the data from last three subjects mentioned combined to form the recency median. Individual data is in Table 4 below.
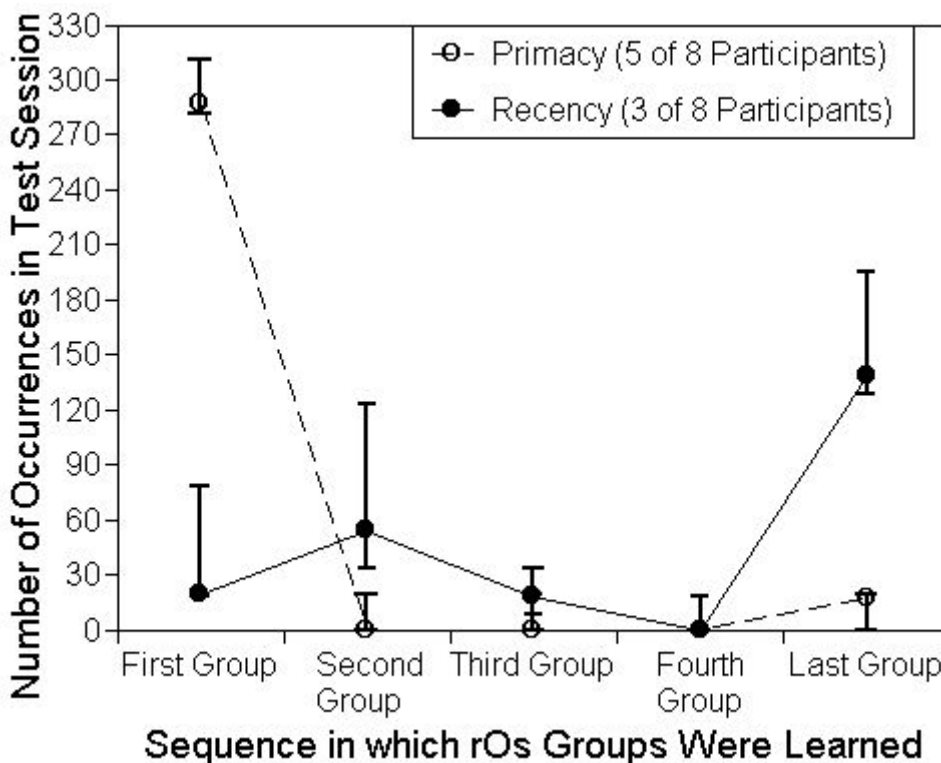

Figure 5. Operants from the given sequential group performed by subjects during the test session, Experiment 4.

Once again, primacy prevailed, however not as strongly as in Experiment 3, with preference for operant classes from the first group learned again only approaching statistical significance: p = .1457, t(7) = 1.64, p < .25, two-tailed.

Experiment 5

Table 4.
Number of operants performed from given sequential group during test session, Experiment 4 (individual data).

Subject	First Group	Second Group	Third Group	Fourth Group	Last Group
Showing	*Predominantly*	*Primacy*			
AJM	189	0	37	54	60
ESS	311	20	0	0	20
LAS	349	0	0	0	0
MA	281	0	34	0	17
SR	287	47	0	0	0
Showing	*Predominantly*	*Recency*	*/ Both*		
ACB	18	193	18	0	118
JLM	137	14	20	36	139
KPW	19	54	0	0	251

The finding in the previous experiments, that the interposition of additional learning tasks between learning sessions, rather than periods of time, seems to favor primacy over recency effects, raises the question of whether this effect was truly due to the time passage variable. The present experiment attempted to address this question by adding long time periods between groups of operant classes in a five-group design, rather than only three.

Method

Ten new subjects were recruited. The recruitment method, equipment, and procedure were the same as for Experiment 4, except that different numbers of days without a session (elapsed time) were inserted between the sessions for each group of operant classes. Experiment 5 thus consisted of 11 sessions that were exactly the

same as the 11 sessions in Experiment 4, but spread out over 25 days.

Subjects learned operant classes 1, 2 and 3 during sessions 1 and 2 as before, followed by an eight-day break without a session before returning for sessions 3 and 4, in which they learned operant classes 4, 5, and 6. Four more days then elapsed without a session before operant classes 7, 8 and 9 were introduced in sessions 5 and 6. There were then two days without a session, followed by sessions 7 and 8, when operant classes 10, 11 and 12 were introduced. Beginning the very next day came sessions 9 and 10 in which operant classes 13, 14 and 15 were introduced. Session 11, as in Experiment 4, was the test session, taking place the day after session 10 and identical in design to that in Experiment 4.

This schedule was chosen so that the number of days between the sessions spent practicing the different groups of operant classes formed a roughly geometric progression: 8, 4, 2 and 0 days.
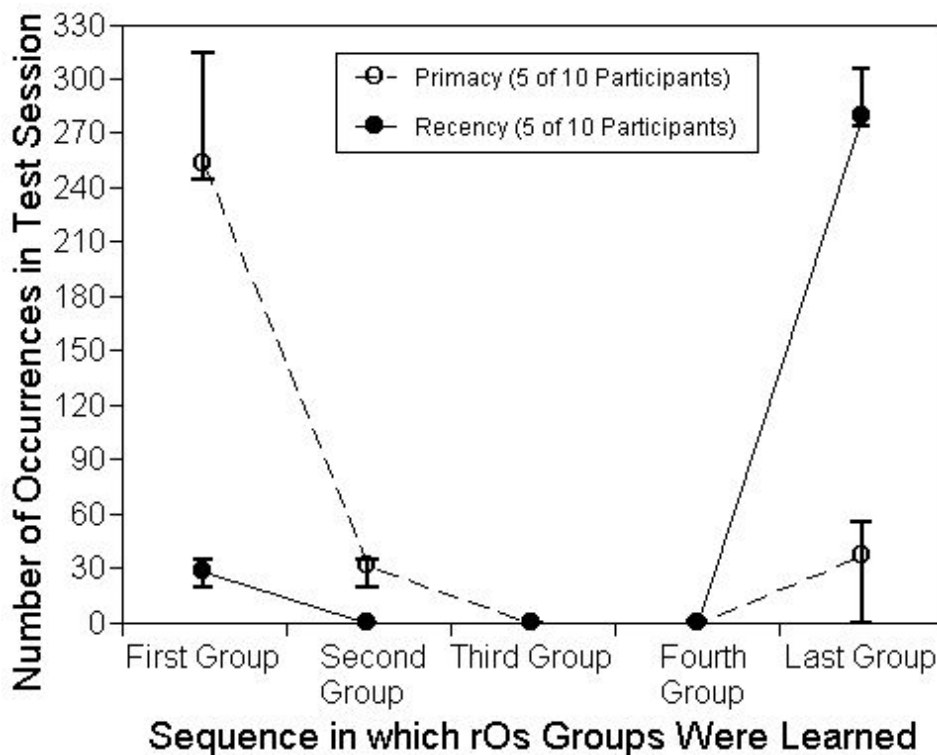

Figure 6.
Operants from the given sequential group performed by subjects during the test session, Experiment 5

Results

In the test session, 4 of the 10 subjects still chose operants from the first group learned far more often than others, even though it had been over three weeks since they had last used them. Another four subjects overwhelmingly chose operants from the last group learned most often. Of the two remaining subjects, one chose operants from the first, second and last groups learned almost equally often (while ignoring those from the third and fourth groups completely) and the other chose operants from all groups almost equally often.

Figure 6 shows these results, with one of these two anomalous subjects placed in the recency median category (RXN) and the other in the primacy median category (NAT). Individual data is in Table 5 below.

Median levels for the three middle groups were extremely low, even lower than those seen in Experiment 4.

In Experiments 4 and 5 there is a difference in the percent accuracy during the test session for operant classes from the first and last groups learned (M = 92.27%,

Table 5.
Number of operants performed from given sequential group during test session, Experiment 5 (individual data)

Subject	First Group	Second Group	Third Group	Fourth Group	Last Group
Showing	*Predominantly*	*Primacy*			
ALB	244	31	0	1	55
DJG	253	35	0	0	37
HJS	314	20	0	0	0
MTF	327	18	0	0	0
Showing	*Predominantly*	*Recency*			
KYC	20	0	0	0	320
LMR	35	0	0	11	279
REF	28	0	0	0	305
RJR	20	0	1	0	274
Showing	*Neither / Both*				
NAT	74	37	89	52	74
RXN	110	117	0	0	116

SD = 6.93) compared to the accuracy level for the second, third and fourth groups -- the middle three (M = 78.47%, SD = 24.22). This difference is significant, t(66) = 3.15, p < .01, two-tailed.

Combined Results

Noncriterial Pattern Analysis
For the purposes of data analysis, all of each subject's noncriterial keystrokes (the six or more discretionary keystrokes between the three mandated initial and the three mandated final keystrokes of each operant) were grouped into patterns of three consecutive keystrokes ("triplets"), using a sliding window with consecutive triplets overlapping (i.e., a single sequence of six noncriterial keystrokes would be broken into four distinct triplets). The triplets did not include any criterial keystrokes. The elimination of pairs of letters that repeated in noncriterial keystrokes meant that each triplet would always consist of three different letters. It is important to note that triplets are not operant units; they are patterns of noncriterial behavior within the operant units.

The total number of mathematically possible triplets that can be formed from the twelve character keys is 1,320 (12x11x10). The median value for the total number of unique (non-duplicated) triplets performed prior to the test session for all subjects in the five experiments is 518. However, the median value for the number of a subject's unique triplets that made up half of all their triplet occurrences is only 6, and the median value for the number that comprised 75% of all triplet occurrences is only 12. The majority of the other 500 or so triplets performed by the median subject occurred only once during the experiment. Furthermore, a significant portion of them occurred only during invalid operants – the median value for the total number of unique triplets found within at least one correctly-typed operant is 352.5, meaning that 165.5 of the median subject's 518 unique operants were performed only during mistakes. The vast majority of a subject's noncriterial keystrokes, although free to vary, are thus actually very highly stereotyped.

Noncriterial triplets tended to be stereotyped in different ways as well: almost half of the triplets used with a given operant class are only found within operants from that class. Of the 352.5 noncriterial triplets associated with at least one valid operant performed by the median subject, 161.5 of these only occur within operants of that class. This implies that a significant number of the noncriterial triplets performed within operants of a given class become tied to the criterial keystrokes for that class.

Finally, the level of variability seen in the use of the noncriterial keystrokes drops quite quickly after learning begins. The average for the total number of unique triplets performed during session 1 (for all subjects in all experiments) is 152.09 (SD 127.12). By the very next session that number has dropped to 104.53 (and the standard deviation has dropped, to 79.71). This difference is significant despite the high variance (t(92) = 2.17, p < .05, two-tailed). The average number

of unique triplets occurring per session then remain essentially the same for the rest of the experiments; in the last learning session before the test, the average is 103.83 (SD 80.24), while in the test session itself it is 111.28 (SD 80.36).

In addition to counting numbers of unique noncriterial triplets, the experiment's software tracked and recorded the occurrence of each one throughout all of each subject's sessions. Tracking the triplets permits them to be identified as old (having occurred previously) or as new -- never having occurred previously. Resurgence of old triplets during the test session could thus be measured in terms of their "antiquity," defined as the number of blocks in an individual subject's learning history one must count back to find a prior instance of that same triplet. A higher antiquity level would thus represent an older noncriterial pattern.

Another significant difference in noncriterial behavior was thus revealed between noncriterial triplets performed during valid operants and those performed during invalid ones. Noncriterial differences between valid and invalid operants were compared, prior to and during the test session, for rates of occurrence of new triplets, and for the antiquities of the old ones.

Figure 7 is a comparison of average antiquity levels of non-criterial patterns within valid operants performed during the test session vs. those within invalid operants. The bar groups show the median values of the average antiquity levels
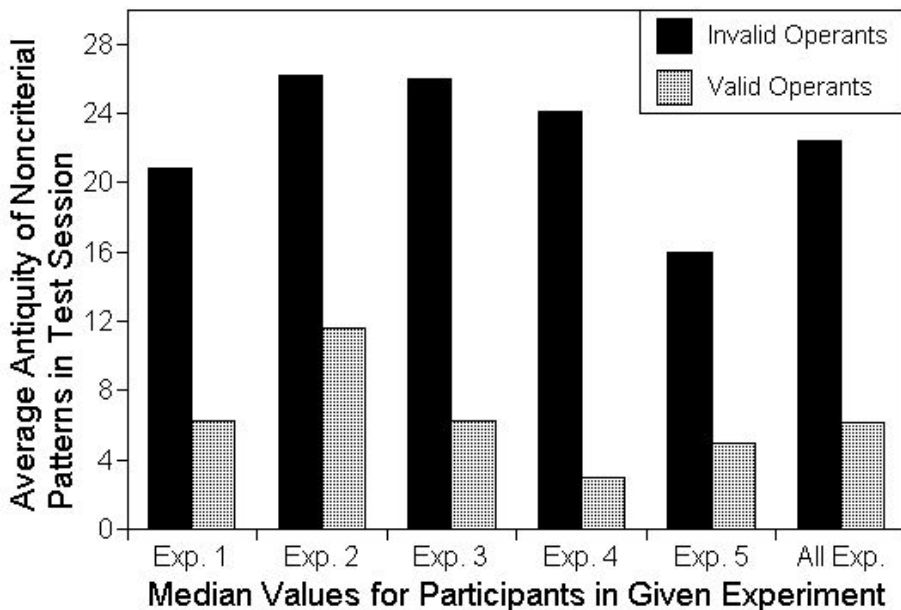

Figure 7. Average antiquity of noncriterial patterns in invalid vs. valid operants performed during test session, all experiments.

for subjects in each individual experiment, and (at the far right) for all subjects in all experiments combined. (Note that triplets themselves cannot be invalid or incorrect, by definition.) This effect is remarkably robust across all five experiments. (For individual subjects' noncriterial data for all experiments, please see the supplementary document, "Effects of Sequential Learning - Additional Tables," available at http://mechnerfoundation.org/newsite/downloads.html.) This antiquity difference was observed for invalid vs. valid operants during the subjects' learning sessions as well, although the difference there is smaller.

In addition to having higher average antiquity levels, a considerably higher percentage of the triplets occurring in invalid operants during the test session were new (never before used), when compared to those within valid operants during the test session – 7.99% vs. 0.36%.

Figure 8 compares percentages of new (never previously recorded) noncriterial patterns within valid and invalid operants during the test session. Again, the median values for subjects in each experiment individually are shown, as well as the median across all experiments. This difference was observed for invalid vs. valid operants during learning sessions as well, but was more pronounced in the test
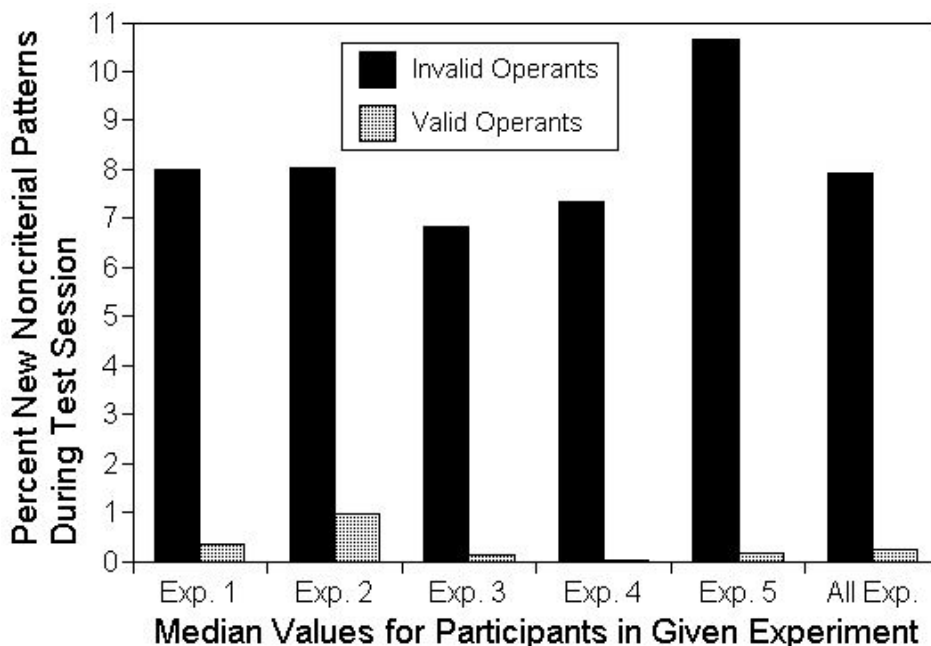

Figure 8. Percent new noncriterial patterns in invalid vs. valid operants performed during test session, all experiments.

session. Note that since each of a subject's unique triplets was, by definition, new the first time it occurred, the rate at which new triplets occurred was, of necessity, higher in the initial sessions than in later sessions.

Taking the median for all subjects in all experiments gives an average antiquity level of 22.42 blocks for triplets in invalid operants and 6.13 blocks for triplets in valid ones during the test session (this difference is shown graphically at the right end of Figure 3). The median (for all subjects in all experiments) average antiquity level in invalid operants during the last learning session is 3.40 blocks, while for valid operants it is 1.53 blocks.

The median value for the subjects' percent of triplets that are new during the test session, for all subjects in all experiments (shown at the right end of Figure 4) was 7.92% within invalid operants and 0.25% within valid ones. In the last learning session of all five experiments, the median (for all subjects) average percent of triplets within invalid operants that were new was 5.32%; within valid operants it was 0.25%.

In addition to average antiquity levels, each subject's individual median antiquity levels for invalid vs. valid operants in the test session were tracked. These levels are much lower in general; for all subjects in all experiments, the median antiquity level is 4 blocks for triplets in invalid operants and 1 block for triplets in valid ones during the test session, not counting blocks within the test session itself. The reason for the large gap between an individual's average and median antiquity is that invalid operants contain a larger number of individual triplets with very high antiquity levels, what one might call "spikes" in antiquity. Defining a "spike" as a triplet with an antiquity value equal to or greater than half of the available history blocks, the median percent of triplets within invalid operants during the test session which are spikes is 14.06%, versus only 0.14% spikes occurring in valid operants.

If an increase in the number of different triplets in use (as was consistently observed in the test sessions) were the result of a general increase in variability, one might expect new triplets and previously used triplets to occur in roughly equal proportion, drawn from the population of (median) 518 previously used ones and, therefore, the 802 (1,320 total possible triplets minus 518) new (not-yet-used) ones. The actual median value for the percentage of triplets occurring in the test session that were previously-used is 92.15% for incorrect operants and 99.78% for correct operants, for all five experiments combined. This means that the vast majority of the additional triplets used during the test session were reoccurrences of ones that had appeared previously in a given subject's learning history, rather than novel ones.

Recency/Primacy Effects
In the test sessions of these five experiments, in which the independent variable was the sequence in which the operant classes had been learned, 42 of the 48

subjects showed recency or primacy or both, i.e. performed the first-learned and/ or last-learned operant classes more often than those learned in between (with the individual differences being statistically significant to the .05 level using z-scores). Of those 42, 18 showed both recency and primacy (i.e., a V- or U-shaped function with statistically significant differences between both the first and middle point, and the last and middle point), and 24 showed either recency or primacy, but not both. In 28 of the 42, primacy predominated, meaning that the first-learned operant classes have a statistically significant advantage over all of the other groups, and in 12, recency predominated, by an analogous definition (again, using z-scores with an alpha of .05 to determine statistical significance within each individual's data).

Far more subjects showed primacy rather than recency in the two studies that used interpolated learning to separate three groups of operants (Experiments 2 and 3), and in the one five-group study, in which the sessions took place on consecutive days (Experiment 4), whereas primacy was not as predominant in the two studies (Experiments 1 and 5, one three-group and one five-group) in which days without a session were inserted between the groups of operants. Inserting additional time between first-, middle-, and last-learned groups of operant classes thus tended to diminish the predominance of primacy in some subjects.

The key question, of course, is whether a variable can be found that correlates with the group (primacy or recency) into which a particular subject's skew will fall. Noncriterial triplets used, accuracy, speed, overall number of errors, and other measures collected show no noticeable differences between the learning session data for these groups.

However, a detailed analysis of all the data did show a difference (at a statistically significant level): those subjects who later chose the most recent group of operant classes (in the test session) made significantly more errors during the first two blocks in which that last group of operant classes were initially learned than did those subjects who later chose the first-learned group. For the 12 subjects for whom recency would later predominate, an average of 2.88% of the operants they performed during those two blocks were invalid (SD 1.76), vs. 2.01% for the 28 subjects who would later show mostly primacy during the test session (SD 0.79). This difference, though small, is statistically significant, with $p = .0174$ ($t(38) = 2.188$, $p < .05$, one-tailed).

In addition, the 28 subjects who later showed primacy committed more errors during the first two blocks in which the first group of operants was learned: 2.58% (SD 2.12) vs. 1.83% for the 12 recency subjects (SD 1.09). This finding is not statistically significant ($p = .1297$) due to the higher variances. However it is possible that the higher overall levels of operant variability, which may be expected during the first few attempts at learning an entirely new skill, masked the effect. It is important to note that both effects are present only in the first two blocks spent learning the given group of operants; there is no statistically significant difference

in number of errors between recency and primacy subjects if one considers the entire learning history for the first vs. the last group of operant classes.

The median speed with which operants were executed (in seconds per operant) for all participants decreased from an average of 3.78 in the first learning session to 2.46 in the last learning session. However, there was no correlation between the speed with which given operants were executed by a particular participant, and that participant's tendency to choose those operants in terms of primacy or recency.

Discussion

The most consistent finding of the present experiments is that when different but similar operant classes were learned sequentially, when presented with a choice of three or five at a time in a final test session, the first and/or last learned ones were chosen most often. This finding is analogous (though, as discussed, not directly comparable) to that from the serial learning and memory literature cited in the introduction. However, with the operant paradigm used in the present experiments, the majority of subjects demonstrate either primacy or recency, rarely both.

The only quantifiable difference found between the subjects showing predominantly primacy and those showing predominantly recency was that each group committed slightly more errors during the first two blocks spent learning those operants they later preferred. A possible interpretation of this finding is that the difficulty encountered during the initial learning phase corresponded to greater effort and attention being exerted in learning those operants than in learning others, thus establishing them more strongly in the repertoire, with the result that they emerged preferentially in the test situation.

Although primacy, recency, or both were seen in 42 of 48 subjects, the specific effect depended on the procedure used as well as on the subject. In general, primacy was more likely to predominate in those studies in which learning sessions occurred on consecutive days. Furthermore, when five groups of operant classes were learned rather than three, both preference for and accuracy of the middle groups of operant classes were much lower in the test session. These results suggests that the effect of serial or sequential order on operant learning can be manipulated methodologically in much the same way that researchers often manipulate the traditional serial-position effect: different conditions present during learning or a different testing method can change the shape of the serial-position curve, skewing it either toward primacy or toward recency and sometimes completely wiping out one of these two effects (Conrad & Hull, 1968; Corballis, 1966; Deese & Kaufman, 1957; Glanzer & Cunitz, 1966; McGeoch & Irion, 1952; Murray, 1966; Welch & Burnett, 1924). Increasing the length of the list of items learned generally causes the point of worst performance to move closer to the middle of the series but does not change the degree to which the serial-position curve is bowed in the

middle (Hovland, 1940; Jensen, 1962; Robinson & Brown, 1926).

It is likely that the contingencies that specify criterial attributes of operants are also having a general effect on the operants' noncriterial attributes. In the current studies, noncriterial triplet variability was shown to decrease very quickly once operant learning began. The vast majority of a subject's noncriterial triplet occurrences consisted of only a few of their total unique triplets, and a significant proportion of each subject's total unique triplets were never found in any valid operants at all, occurring only during mistakes. In other words, noncriterial keystrokes, though free to vary, quickly became highly stereotyped and seemingly linked to their accompanying criterial keystrokes. These results are similar to those of Escobar & Bruner (2007), who found that noncriterial and criterial responding (rats' lever-pressing) tended to become fused in distinctive, repetitive patterns, different from rat to rat but stable when the data for individual rats are combined across several sessions. Going further, noncriterial variability, after this stability has been achieved, seems to be associated with mistakes, or breaks in the flow of behavior. These results strongly suggest that the sequence of motor behavior making up these operant classes is becoming automatized during learning, with the noncriterial keystrokes becoming fused with the criterial ones.

Some of these effects are doubtless the result of the linkages in the motor coordinations required by the criterial properties (Mechner, 1994, p. 40). With repeated executions of the operant, these linked coordinations and others may acquire strength by virtue of repetition and consequent automatization (Schmidt, 1988, p. 74; Shiffrin, 1988, pp. 740-767).

Finally, the stressors instituted in the test sessions (loss of money for longer-than-usual pauses and invalid operants, coupled with monetary reinforcement of valid ones) produced higher levels of average antiquity of noncriterial patterns (triplets) in the test session when compared to the previous learning session, and an increased incidence of new patterns. This finding raises the experimental question of whether other conditions that might be termed stressful, such as fatigue, distraction, discomfort, or sudden increases in task difficulty, or other variables that disrupt behavior, would produce similar effects.

One of the general challenges in studying variable phenomena is to discern and identify patterns in the variable elements, so that initially unpredictable events become more predictable. Thus, the parsing of noncriterial patterns (triplets) into old and new ones, and the further analysis of the old ones according to antiquity, amounts to the parsing of variability into some of its contributing components (Cleland, Foster & Temple, 2000).

The higher antiquity levels of noncriterial patterns (and the higher frequency of new patterns) in invalid operants than in valid ones may be due to invalid operants being more strongly affected by the effects of the test session stressors. Part of the explanation may also reflect the propagation of mistakes made in the first three criterial keystrokes into the noncriterial keystroke portion of the operant.

The robustness of this effect suggests that some of the errors in any skilled performance may be associated with the resurgence of previous, less efficient behavior patterns. Prior studies have shown that levels of response variability often depend on noncriterial response attributes in the individual's early learning history (Stokes, 1995; Stokes, Mechner & Balsam, 1999). The majority of the increase in average antiquity levels in invalid operants seen during the test session was caused by a greater number of "spikes," or individual noncriterial patterns with extremely high antiquity levels, rather than by slightly higher noncriterial antiquity levels across the board. This finding is consistent with that of an experiment on variability which found that extinction produced very large increases in the numbers of formerly low-probability responses while leaving the overall probability hierarchy of responses unchanged (Neuringer, Kornell & Olufs, 2001).

In addition, greater numbers of new noncriterial patterns always accompanied the increases in antiquity of noncriterial patterns observed in invalid operants during the test sessions of the present experiments (but not nearly in proportion to their available supply, as was shown previously). This finding is in line with experiments that have found that extinction -- another resurgence-inducing contingency -- not only increases variability but also increases the number of novel responses (Neuringer et al, 2001), though that study did not parse the variability increase into old response patterns and new ones.

The broader significance of the present series of experiments is the demonstration that under performance conditions, the choice of operant classes and the criterial as well as noncriterial attributes of the operant occurrences can be a function of history variables during learning. The findings of these types of experiments have practical implications for the technology and pedagogy of performance learning and practicing.

References

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

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence, & J. T. Spence (Eds.), *The Psychology of Learning and Motivation II* (pp. 89-185). Oxford, England: Academic Press, available via: http://dx.doi.org/10.1016/S0079-7421(08)60422-3

Bachá-Mendéz, G., Reid, A. K, & Mendoza-Soylovna, A. (2007). Resurgence of integrated behavioral units. *Journal of the Experimental Analysis of Behavior, 87,* 5-24.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6,* 173-189, available via: http://dx.doi.org/10.1016/0010-0285(74)90009-7

Cleland, B. S., Foster, T. M., & Temple, W. (2000). Resurgence: The role of extinction. *Behavioural Processes, 52,* 117-129, available via: http://dx.doi.org/10.1016/S0376-6357(00)00131-5

Conrad, R., & Hull, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychonomic Science, 10*, 135-136.

Corballis, M. C. (1966). Rehearsal and decay in immediate recall of visually and auditorily presented items. *Canadian Journal of Psychology, 20,* 43-51, available via: http://dx.doi.org/10.1037/h0082923

Crowder, R. G. (1976). Principles of learning and memory. Hillsdale, NJ: Erlbaum.

Crowder, R. G., & Neath, I. (1990). Schedules of presentation and temporal distinctiveness in human memory. *Journal of Experimental Psychology: Learning, Memory & Cognition, 16,* 316-327.

Deese, J., & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology, 54,* 180-187, available via: http://dx.doi.org/10.1037/h0040536

Di Lollo, V., Ensminger, W. D., & Notterman, J. M. (1965). Response force as a function of amount of reinforcement. *Journal of Experimental Psychology, 70,* 27-31.

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

Escobar, R., & Bruner, C. A. (2007). Response induction during the acquisition and maintenance of lever pressing with delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 88,* 29-49.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 351-360, available via: http://dx.doi.org/10.1016/S0022-5371(66)80044-0

Guthrie, E. R. (1959). Association by contiguity. In S. Koch (Ed.), *Psychology: A Study of a Science, Vol. 2* (pp. 158–195). New York, McGraw-Hill.

Herrnstein, R. J. (1966). Superstition: A corollary of the principles of operant condition¬ing. In W. K. Honig (Ed.), *Operant Behavior: Areas of Research and Application* (pp. 33-51). New York: Appleton-Century-Crofts.

Hovland, C. I. (1940). Experimental studies in rote-learning theory: VII. Distribution of practice with varying lengths of list. *Journal of Experimental Psychology, 27,* 271-284, available via: http://dx.doi.org/10.1037/h0063609

Hull, C. L. (1935). The conflicting psychologies of learning -- a way out. *Psychological Review, 42,* 491-516, available via: http://dx.doi.org/10.1037/h0058665

Jensen, A. R. (1962). An empirical theory of the serial-position effect. *Journal of Psychology, 53,* 127-142, available via: http://dx.doi.org/10.1080/00223980.1962.9916559

Lepley, W. M. (1934). Serial reactions considered as conditioned reactions. *Psychological Monographs, 46,* (205).

McGeoch, J. A., & Irion, A. L. (1952). *The psychology of human learning*. New York, NY: Longmans, Green.

Mechner, F. (1994). *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.* Retrieved February 10, 2006, from 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.

Murdock, B. B. (1960). The distinctiveness of stimuli. *Psychological Review, 67,* 16-31, available via: http://dx.doi.org/10.1037/h0042382

Murray, D. J. (1966). Vocalization-at-presentation and immediate recall, with varying recall methods. *Quarterly Journal of Experimental Psychology, 18,* 9-18, available via: http://dx.doi.org/10.1080/14640746608400002

Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology*: *Animal Behavior Processes, 27,* 79-94, available via: http://dx.doi.org/10.1037//0097-7403.27.1.79

Robinson, E. S., & Brown, M. A. (1926). Effect of serial position upon memorization. *American Journal of Psychology, 37,* 538-552, available via: http://dx.doi.org/10.2307/1414914

Rundus, D., & Atkinson, R. C. (1970). Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior, 9,* 99-105, available via: http://dx.doi.org/10.1016/S0022-5371(70)80015-9

Schmidt, R.A. (1988). *Motor Control and Learning: A Behavioral Emphasis.* Champaign, IL: Human Kinetics Publishers.

Schoenfeld, W. N. (1968). On the difference in resistance to extinction following regular and periodic reinforcement. *Journal of the Experimental Analysis of Behavior, 11,* 259-261.

Shiffrin, R. M. (1988). Attention. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' Handbook of Experimental Psychology, Second Edition, Volume 2: Learning and Cognition* (pp. 739-812). New York: John Wiley & Sons.

Stokes, P. D. (1995). Learned variability. *Animal Learning & Behavior, 23,* 164-176. doi:10.3758/BF03199931

Stokes, P. D., Mechner, F., & Balsam, P. D. (1999). Effects of different acquisition procedures on response variability. *Animal Learning & Behavior, 27,* 28-41, available via: http://dx.doi.org/10.3758/BF03199429

Tatham, T. A., & Wanchisen, B. A. (1998). Behavioral history: A definition and some common findings from two areas of research. *Behavior Analyst, 21,* 241-251.

Wanchisen, B. A. (1990). Forgetting the lessons of history. *Behavior Analyst, 13,* 31-37.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72,* 89-104, available via: http://dx.doi.org/10.1037/h0021797

Welch, G. B., & Burnett, C. T. (1924). Is primacy a factor in association-formation? *American Journal of Psychology, 35,* 396-401, available via: http://dx.doi.org/10.2307/1414018

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

Wright, A. A. (1998). Auditory and visual serial position functions obey different laws. *Psychonomic Bulletin & Review, 5,* 564-584, available via: http://dx.doi.org/10.3758/BF03208835